Sensors Handbook

Second Edition

SABRIE SOLOMAN

SENSORS HANDBOOK

Sabrie Soloman, **Ph.D., Sc.D., MBA, PE**

Chairman and CEO American SensoRx, Inc., USA
Professor, Founder, Advanced Manufacturing Technology,
Columbia University, USA
Chairman and CEO, SensoRx, Chongqing, Ltd., China.

Second Edition

McGraw Hill

New York Chicago San Francisco Lisbon London Madrid
Mexico City Milan New Delhi San Juan Seoul
Singapore Sydney Toronto

The McGraw-Hill Companies

ISBN: 978-0-07-160571-7

MHID: 0-07-160571-1

The material in this eBook also appears in the print version of this title: ISBN: 978-0-07-160570-0,
MHID: 0-07-160570-3.

To My Beloveds

Her hovering soul is as a lily, majestically floating on water embracing the morning dew—quenching my thirst, collecting the morning sun in the basket of her illuminating righteousness enlightening my eyes—to the noble one, and to her cherished offspring. The work of this handbook is dedicated, to my beloveds—wife and son Elvira and Stephan.

To My Mentors

The humble work of this handbook is contributed to my mentors in my childhood and adulthood alike. During my childhood, both my father and mother—Dr. Barsoum and Mrs. Zahia Soloman, who departed this earth in perfect glory and full honor—selflessly gave themselves unceasingly to my modest beginning, and their memory never fails to seize the uppermost of my intellect. The echo of their voice still resonates through my soul, remembering their words of wisdom and courage. Although a great deal of their teachings were dedicated toward acquiring the knowledge for science and innovation, yet their few divine words about Christ, the founder of science, and the creator of the universe made me embrace the ideology of creation and accept the Christian faith eternally.

Similarly, my brother Dr. Nasier Soliman and my sister-in-law Mrs. Anne Soliman have permeated in me the spirit of permutation and brought forth the hidden wealth of knowledge that I unknowingly possessed. My parents had instilled in me the desire to follow their footsteps for a short period, only to teach me how to bypass them. Similarly, and in the fullness of time, I found my brother Nasier and my sister-in-law Anne, are selflessly paving the path of my success by their dedicated advice and waving care. Their perpetual encouragements have set the course of my advancements to leap forward in spite of limitations.

Indeed, the loving living example of my parents will always live on, through me, my corporate staff, my students, and my students' students.

ABOUT THE AUTHOR

Sabrie Soloman, Ph.D., Sc.D., MBA, PE, is the Founder, Chairman, and CEO of American SensoRx, Inc., and SensoRx, Chongqing, China, Ltd. He is a professor and founder of advanced manufacturing technology at Columbia University. Dr. Soloman is considered an international authority on advanced manufacturing technology and automation in the microelectronics, automotive, pharmaceuticals, and food industries. He has been and continues to be instrumental in developing and implementing several industrial modernization programs through the United Nations to various European and African countries. Dr. Soloman is the first to introduce and implement Nano-MEMS technology in sensors manufacturing and applications. He introduced and implemented unmanned flexible synchronous/asynchronous systems to the microelectronic and pharmaceutical industries and was the first to incorporate advanced vision technology to a wide array of robot manipulators. Dr. Soloman invented the SpectRx NIR sensing technology for the medical and pharmaceuticals industry, as he correlated tablet hardness and disintegration with tablet NIR spectrum.

CONTENTS

Chapter 19. Smart Civil Structures, Intelligent Structural Systems 491

Chapter 20. True Online Color Sensing and Recognition 497

Chapter 21. Fundamentals of Solid-State Presence-Sensing Technologies 505

Chapter 22. Design and Application of Robust Instrumentation Sensors in Extreme Environments 533

Chapter 53. Enabling Technologies for Low-Cost High-Volume Pressure Sensors 973

Chapter 54. A Two-Chip Approach to Smart Sensing 985

Chapter 55. Specifying and Selecting Semiconductor Pressure Transducers 989

Chapter 56. Introduction to Silicon Sensor Terminology 997

Chapter 57. Silicon Sensors and Microstructures: Integrating an Interdisciplinary Body of Material on Silicon Sensors 1007

Chapter 58. Understanding Silicon Processing and Micromachining 1041

This page intentionally left blank

FOREWORD

We have entered the third millennium …!

Relations between nations, institutions, and even between individuals are increasingly characterized by a comprehensive and rapid exchange of information. Communication between scientists and engineers, bankers and brokers, manufacturers and consumers proceeds at an ever quickening pace. Exchanging ideas and thoughts with others is no longer a matter of weeks, months, or years. Humankind has now reached the point of being able to distribute large volumes of information to any number of addresses within the blink of an eye.

Human intelligence is thus nourished from many different sources, producing globally useful scientific, technical, and economic improvements within only a short time. The question of whether the globalization of our thoughts is an advantage or a disadvantage for humankind still remains to be answered, however.

Dr. Sabrie Soloman devotes his scientific work to the challenging area of translating scientific ideas into technically applicable solutions. He consolidates thoughts and knowledge within controlled systems and participates to ensure, during the advent of this third millennium, that technical progress is not conceived as a risk but rather as an opportunity.

Dr. Soloman's new handbook introduces us into the world of advanced sensor technology, starting with information in a simple form, and, with increasing complexity, offers sophisticated solutions for problems that to date were considered insoluble. Profound knowledge and understanding of all phases of the processes within complex systems are the prerequisites to securing control and quality. This is clearly illustrated by a large number of applications and examples from chemistry, physics, medicine, military, and other allied subjects covered in Dr. Soloman's handbook.

Adaptive systems are increasingly being called for in the processing of complex information. These systems must react quickly and adequately to expected, and even unforeseeable, developments.

Uhlmann, as a producer of highly sophisticated packaging systems for the pharmaceutical industry worldwide, stands for a high level of reliability thanks to its mechanical, electrical, electronic, and optical modules permitting quality and efficiency in pharmaceutical packaging to be combined.

Reflecting on almost 50 years of tradition as a supplier to the pharmaceutical industry, forward thinking and the ambition to combine advanced science with state-of-the-art technology in our packaging systems have been the cornerstones of our success. Developments for the future include the application of advanced sensor technology and fiber-optic transmission techniques, as well as adaptive and decision-making computers on our machinery. Thus, we contribute our share to ensure that medication, whether in the form of liquids, tablets, or coated tablets, is supplied to the consumer in perfect quality as far as Uhlmann is responsible for the packaging process.

Dr. Soloman's handbook offers the user a large variety of references in regard to target-oriented sensor technology—the analysis of the subsequent process information transformed into signals that secure function and quality. *Sensors Handbook* will certainly provide impulses for the development of highly sophisticated systems, continuing to challenge us in the future.

HEDWIG UHLMANN

Advisory—Board of Directors Uhlmann Pac-Systeme GmbH & Co. KG

This page intentionally left blank

PREFACE

SETTING THE STAGE FOR ADVANCED SENSORS

Advanced sensory and control technology, discussed in this handbook, is more than an implementation of new sensing technologies. It is a long-range strategy that allows the entire manufacturing and research operation to work together to achieve the business qualitative and quantitative goals. It must have the top management commitment. It may entail changing the mind-set of people in the organization and managing the change. The major success of this manufacturing strategy is largely credited to the success of implementing the advanced technology of sensory and control systems.

This handbook deals with setting up relatively small devices—often called sensors—designed to sense and measure an object's physical characteristics such as size, speed, acceleration, color, temperature, pressure, volume, flow rate, altitude, latitude, shape, orientation, quantity, deformation, homogeneity, topography, viscosity, electric voltage, electric current, electric resistance, surface textures, microcracks, vibrations, noise, acidity, contamination, active ingredient, assay concentration, chemical composition of pharmaceutical drugs, and blood viruses.

MANUFACTURING OF ARTIFICIAL ORGANS

The control of diabetes with insulin shots may fail to maintain adequate function of kidney. The concept of one's organs living in another's body is rarely realized. In the third-century legend of Saints Cosmos and Damian, the leg of a recently deceased Moorish servant is transplanted onto a Roman cleric whose own limb has just been amputated. The cleric's life hangs in the balance, but the transplant takes, and the cleric lives. The miraculous cure is attributed to the intervention of the saintly brothers, both physicians, who were martyred in A.D. 295.

What was considered miraculous in one era may become merely remarkable in another. Surgeons have been performing reimplantation of severed appendages for almost four decades now, and transplants of organs such as the heart, liver, and kidney are common—so common, in fact, that the main obstacle to transplantation lies not in surgical technique but in an ever-worsening shortage of the donated organs themselves. In the next three decades, medical science will move beyond the practice of transplantation and into the era of fabrication. The idea is to make organs rather than simply to move them.

"Bridging" Technologies of Sensors and Medicine

The advent of advanced sensor and control technology, has caused an advancement in cell biology and plastic manufacture. These have already enabled researchers to construct artificial tissues that look and function like their natural counterparts. Genetic engineering may produce universal donor cells—cells that do not provoke rejection by the immune system—for use in these engineered tissues. "Bridging" technologies of sensors and medicine may serve as intermediate steps before such fabrication becomes commonplace. Transplantation of organs from animals, for example, may help alleviate the problem of organ shortage. Several approaches under investigation involve either breeding animals, whose tissues will be immunologically accepted in humans, or developing drugs to allow the acceptance of these tissues. Alternatively, microelectronics may help bridge the gap between the new technologies and the old. The results will bring radical changes in the treatment of a host of devastating conditions. Engineering artificial tissue is the natural successor to treatments for injury and disease.

Millions of people suffer organ and tissue loss every year from accidents, birth defects, and diseases such as cancer and diabetes. In the last quarter of the 20th century, innovative drugs, surgical procedures, and medical devices have greatly improved the care of these patients. Immunosuppressive drugs such as cyclosporine and tacrolimus (Prograf) prevent rejection of transplanted tissue; minimally invasive surgical techniques such as laparoscopy have reduced trauma; dialysis and heart-lung machines sustain patients whose conditions would otherwise be fatal.

Yet these treatments are imperfect and often impair the quality of life. The control of diabetes with insulin shots, for example, is only partly successful. Injection of the hormone insulin once or several times a day helps the cells of diabetics to take up the sugar glucose (a critical source of energy) from the blood. But the appropriate insulin dosage for each patient may vary widely from day to day and even hour to hour. Often amounts cannot be determined precisely enough to maintain blood sugar levels in the normal range and thus prevent complications of diabetes—such as blindness, kidney failure, and heart disease—later in life.

Innovative research in biosensor design and drug delivery, will someday make insulin injections obsolete. In many diabetics, the disease is caused by the destruction in the pancreas of so-called islet tissue, which produces insulin. In other people, the pancreas makes insulin, but not enough to meet the body's demands. It is possible to envision a sensor-controlled device that would function like the pancreas, continuously monitoring glucose levels and releasing the appropriate amount of insulin in response. The device could be implanted or worn externally.

THE SPECTRx

Much of the technology for an external glucose sensor that might be worn like a watch already exists. Recent studies at the Massachusetts Institute of Technology, the University of California at San Francisco, and elsewhere have shown that the permeability of the skin can temporarily be increased by electric fields or low-frequency ultrasonic waves, allowing molecules such as glucose to be drawn from the body. The amount of glucose extracted in this way can be measured by reaction with an enzyme such as glucose oxidase; or infrared sensors, such as the SpectRx, could detect the level of glucose in the blood.

These sensing devices could be coupled via microprocessors to a power unit that would pass insulin through the skin and into the bloodstream by the same means that the sugar was drawn out. The instrument would release insulin in proportion to the amount of glucose detected.

An implantable device made of a semipermeable plastic could also be made. The implant, which could be inserted at any of several different sites in the body, would have the form of a matrix carrying reservoirs of insulin and glucose oxidase. As a patient's glucose level rose, the sugar would diffuse into the matrix and react with the enzyme, generating an acidic breakdown product. The increase in acidity would alter either the permeability of the plastic or the solubility of the hormone stored within it, resulting in a release of insulin proportional to the rise in glucose. Such an implant could last a lifetime, but its stores of glucose oxidase and insulin would have to be replenished.

The ideal implant would be one made of healthy islet cells that would manufacture insulin themselves. Investigators are working on methods to improve the survival of the tissue, but supply remains a problem. As is the case with all transplantable organs, the demand for human pancreas tissue far out strips the availability. Consequently, researchers are exploring ways to use islets from animals. They are also attempting to create islet tissue, not quite from scratch, but from cells taken from the patient, a close relative, or a bank of universal donor cells. The cells could be multiplied outside the body and then returned to patient.

SPINNING PLASTIC INTO TISSUE

Many strategies in the field of tissue engineering depend on the manipulation of ultrapure, biodegradable plastics or polymers suitable to be used as substrates for cell culture and implementation. These polymers possess both considerable mechanical strength and a high surface-to-volume ratio. Many are descendants of the degradable sutures introduced three decades ago. Using computer-aided manufacturing methods, researchers design and manipulate plastics into intricate scaffolding beds that mimic the structure of specific tissues and even organs. The scaffolds are treated with compounds that help cells adhere and multiply, then "seeded" with cells. As the cells divide and assemble, the plastic degrades. Finally, only coherent tissue remains. The new, permanent tissue can then be implanted in the patient.

This approach has already been demonstrated in animals, most recently in engineered heart valves in lambs; these valves were created from cells derived from the animals' blood vessels. During the past several years, human skin grown on polymer substrates has been grafted onto burn patients and foot ulcers of diabetic patients with some success. The epidermal layer of the skin may be rejected in certain cases, but the development of universal donor epidermal cells will eliminate that problem.

Eventually, whole organs such as kidneys and liver will be designed, fabricated, and transferred to patients. Although it may seem unlikely that a fully functional organ could grow from a few cells on a polymer frame, research with heart valves suggests that cells are remarkably adept at organizing the regeneration of their tissue of origin. They are able to communicate in three-dimensional culture using the same extracellular signals that guide the development of organs in utero. We have good reason to believe that, given the appropriate initial conditions, the cells themselves will carry out the subtler details of organ reconstruction. Surgeons will need only to orchestrate the organs' connections with patients' nerves, blood vessels, and lymph channels.

Similarly, engineered structural tissue will replace the plastic and metal prostheses used today to repair damage to bones and joints. These living implants will merge seamlessly with the surrounding tissue, eliminating problems such as infection and loosening at the joint that plague contemporary prostheses. Complex, customized shapes such as noses and ears can be generated by constructed computer-aided contour mapping and the loading of cartilage cells onto polymer constructs; indeed, these forms have been made and implanted in laboratory animals. Other structural tissues, ranging from urethral tubes to breast tissue, can be fabricated according to the same principle. After mastectomy, cells that are grown on biodegradable polymers would be able to provide a completely natural replacement for the breast.

Ultimately, tissue engineering will produce complex body parts such as hands and arms. The structure of these parts can already be duplicated in polymer scaffolding, and most of the relevant tissue types—muscle, bone, cartilage, tendon, ligaments, and skin—grow readily in culture. A mechanical bioreactor system could be designed to provide nutrients, exchange gases, remove waste, and modulate temperature while the tissue matures. The only remaining obstacle to such an accomplishment is the resistance of nervous tissue to regeneration. So far no one has succeeded in growing human nerve cells. But a great deal of research is being devoted to this problem, and many investigators are confident that it will be overcome.

INNOVATIVE MICROELECTRONICS

In the meantime, innovative microelectronic devices may substitute for implants of engineered nervous tissue. For example, a microchip implant may someday be able to restore some vision to people who have been blinded by diseases of the retina, the sensory membrane that lines the eye. In two of the more common retinal diseases, retinitis pigmentosa and macular degeneration, the light-receiving ganglion cells of the retina are destroyed, but the underlying nerves that transmit images from those cells to the brain remain intact and functional.

An ultrathin chip placed surgically at the back of the eye, could work in conjunction with a miniature camera to stimulate the nerves that transmit images. The camera would fit on a pair of eyeglasses; a laser attached to the camera would both power the chip and send it visual information via an infrared beam. The microchip would then excite the retinal nerve endings much as healthy cells do, producing the sensation of sight. At MIT and the Massachusetts Eye and Ear Infirmary, recent experiments in rabbits with a prototype of this "vision chip" have shown that such a device can stimulate the ganglion cells, which then send signals to the brain. Researchers will have to wait until the chip has been implanted in humans to know whether those signals approximate the experience of sight. Mechanical devices will also continue to play a part in the design of artificial organs, as they have in this century. They will be critical components in, say, construction of the so-called artificial womb. In the past few decades, medical science has made considerable progress in the care of premature infants. Current life support systems can sustain babies at 24 weeks of gestation; their nutritional needs are met through intravenous feeding, and ventilators help them to breathe.

Younger infants cannot survive, primarily because their immature lungs are unable to breathe air. A sterile, fluid-filled artificial womb would improve survival rates for these newborns. The babies would breathe liquids called perfluorocarbons, which carry oxygen and carbon dioxide in high concentrations. Perfluorocarbons can be inhaled and exhaled just as air is. A pump would maintain continuous circulation of the fluid, allowing for gas exchange. The uterine environment is more closely approximated by liquid breathing than by traditional ventilators, and liquid breathing is much easier on the respiratory tract. Indeed, new work on using liquid ventilation in adults with injured lungs is under way. Liquid ventilation systems for older babies are currently in clinical trials. Within a decade or so, such systems will be used to sustain younger fetuses.

In addition to a gas exchange apparatus, the artificial womb would be equipped with filtering devices to remove toxins from the liquid. Nutrition would be delivered intravenously, as it is now. The womb would provide a self-contained system in which development and growth could proceed normally until the baby's second "birth." For most premature babies, such support would be enough to ensure survival. The developing child is, after all, the ultimate tissue engineer.

SELF-ASSEMBLY AS AN ACT OF CREATION

Nature abounds with examples of self-assembly. Consider a raindrop on a leaf. The liquid drop has a smooth, curved surface of just the kind required for optical lenses. Grinding a lens of that shape would be a major undertaking. Yet the liquid assumes this shape spontaneously, because molecules at the interface between liquid and air are less stable than those in the interior. The laws of thermodynamics require that a raindrop take the form that maximizes its energetic stability. The smooth, curved shape does so by minimizing the area of the unstable surface.

This type of self-assembly, known as thermodynamic self-assembly, works to construct only the simplest structures. Living organisms, on the other hand, represent the extreme in complexity. They, too, are self-assembling: cells reproduce themselves each time they divide. Complex molecules inside a cell direct its function. Complex subcomponents help to sustain cells. The construction of a cell's complexity is balanced thermodynamically by energy-dissipating structures within the cell and requires complex molecules such as ATP. An embryo, and eventually new life, can arise from the union of two cells, whether or not human beings attend to the development.

The kind of self-assembly embodied by life is called coded self-assembly because instructions for the design of the system are built into its components. The idea of designing materials with a built-in set of instructions that will enable them to mimic the complexity of life is immensely attractive. Researchers are only beginning to understand the kinds of structures and tasks that could exploit this approach. Coded self-assembly is truly a concept for the remainder of this century.

SUPER-INTELLIGENT MATERIALS

Imagine, for a moment, music in your room or car that emanates from the doors, floor, or ceiling; ladders that alert us when they are overburdened and may soon collapse under the strain; buildings and bridges that reinforce themselves during earthquakes and seal cracks of their own accord. Like living beings, these systems would alter their structure, account for damage, effect repairs, and retire—gracefully, one hopes—when age takes its toll.

Such structures may seem far-fetched. But, in fact, many researchers have demonstrated the feasibility of such "living" materials. To animate an otherwise inert substance, modern-day alchemists enlist a variety of devices: actuators and motors that behave like muscles; sensors that serve as nerves and memory; and communications and computational networks that represent the brain and spinal column. In some respects, the systems have features that can be considered superior to biological functions—some substances can be hard and strong one moment but made to act like Jell-O the next.

These so-called intelligent materials systems have substantial advantages over traditionally engineered constructs. Henry Petroski, in his book *To Engineer Is Human*, perhaps best articulated the traditional principles. A skilled designer always considers the worst-case scenario. As a result, the design contains large margins of safety, such as numerous reinforcements, redundant subunits, backup subsystems, and added mass. This approach, of course, demands more natural resources than are generally required and consumes more energy to produce and maintain. It also requires more human effort to predict those circumstances under which an engineered artifact will be used and abused.

Trying to anticipate the worst case has a much more serious and obvious flaw, one we read about in the newspapers and hear about on the evening news from time to time: that of being unable to foresee all possible contingencies. Adding insult to injury is the costly litigation that often ensues.

Intelligent materials systems, in contrast, would avoid most of these problems. Made for a given purpose, they would also be able to modify their behavior under dire circumstances. As an example, a ladder that is overloaded with weight could use electrical energy to stiffen and alert the user of the problem. The overload response would be based on the actual life experience of the ladder, to account for aging or damage. As a result, the ladder would be able to evaluate its current health; when it could no longer perform even minimal tasks, the ladder would announce its retirement. In a way, then, the ladder resembles living bone, which remodels itself under changing loads. But unlike bone, which begins to respond within minutes of an impetus but may take months to complete its growth, an intelligent ladder needs less than a second to change.

ARTIFICIAL MUSCLES FOR INTELLIGENT SYSTEMS

Materials that allow structures such as ladders to adapt to their environment are known as actuators. Such substances can change shape, stiffness, position, natural frequency, and other mechanical characteristics in response to temperature or electromagnetic fields. The four most common actuator materials being used today are shape-memory alloys, piezoelectric ceramics, magnetostrictive materials, and electrorheological and magnetorheological fluids. Although none of these categories stands as the perfect artificial muscle, each can nonetheless fulfill particular requirements of many tasks.

Shape-memory alloys are metals that at a certain temperature revert back to their original shape after being strained. In the process of returning to their "remembered" shape, the alloys can generate a large force useful for actuation. Most prominent among them, perhaps, is the family of the nickel-titanium alloys developed at the Naval Surface Warfare Center (formerly the Naval Ordnance Laboratory). The material, known as Nitinol (Ni for nickel, Ti for titanium, and NOL for Naval Ordnance Lab), exhibits substantial resistance to corrosion and fatigue and recovers well from large deformations. Strains that elongate up to 8 percent of the alloy's length can be reversed by heating the alloy, typically with electric current.

THE USE OF JAPANESE NITINOL IN MICROMANIPULATORS

Japanese engineers are using Nitinol in micromanipulators and robotics actuators to mimic the smooth motions of human muscles. The controlled force exerted when the Nitinol recovers its shape allows these devices to grasp delicate paper cups filled with water. Nitinol wires embedded in composite materials have also been used to modify vibrational characteristics. They do so by altering the rigidity or state of stress in the structure, thereby shifting the natural frequency of the composite. Thus, the structure would be unlikely to resonate with any external vibrations; this process is known to be powerful enough to prevent the collapse of a bridge. Experiments have shown that embedded Nitinol can apply compensating compression to reduce stress in a structure. Other applications for these actuators include engine mounts and suspensions that control vibration.

The main drawback of shape-memory alloys is their slow rate of change. Because actuation depends on heating and cooling, they respond only as fast as the temperature can shift.

PIEZOELECTRIC DEVICES—PIERRE AND JACQUES CURIE

A second kind of actuator, one that addresses the sluggishness of the shape-memory alloys, is based on piezoelectrics. This type of material, discovered in 1880 by French physicists Pierre and Jacques Curie, expands and contracts in response to an applied voltage.

Piezoelectric devices do not exert nearly so potent a force as shape-memory alloys; the best of them recover only from less than 1 percent strain. But they act much more quickly, in thousandths of a second. Hence, they are indispensable for precise, high-speed actuation. Optical tracking devices, magnetic heads and adaptive optical systems for robots, ink-jet printers, and speakers are some examples of systems that rely on piezoelectrics. Lead zirconate titanate (PZT) is the most widely used type.

Recent research has focused on using PZT actuators to attenuate sound, dampen structural vibrations, and control stress. At Virginia Polytechnic Institute and State University, piezoelectric actuators were used in bonded joints to resist the tension near locations that have a high concentration of strain. The experiments extended the fatigue life of some components by more than an order of magnitude.

A third family of actuators is derived from magnetostrictive materials. This group is similar to piezoelectrics except that it responds to magnetic, rather than electric, fields. The magnetic domains in the substance rotate until they line up with an external field. In this way, the domains can expand the material. Terfenol-D, which contains the rare earth element terbium, expands by more than 0.1 percent. This relatively new material has been used in low-frequency, high-power sonar transducers, motors, and hydraulic actuators. Like Nitinol, Terfenol-D is being investigated for use in the active damping of vibrations.

The fourth kind of actuator for intelligent systems is made of special liquids called electrorheological and magnetorheological fluids. These substances contain micrometer-sized particles that form chains when placed in an electric or magnetic field, resulting in increases in apparent viscosity of up to several orders of magnitude in milliseconds. Applications that have been demonstrated with these fluids include tunable dampers, vibration-isolation systems, joints for robotic arms, and frictional devices such as clutches, brakes, and resistance controls on exercise equipment. Still, several problems such as abrasiveness and chemical instability plague these fluids, and much recent work to improve these conditions is aimed at the magnetic substances.

FIBEROPTICS

Providing the actuators with information are the sensors, which describe the physical state of the materials system. Advances in micromachining, contributed largely by American electronic industries and research institutes have created a wealth of promising electromechanical devices that can serve as sensors. The main focus is on two types that are well developed now and are the most likely to be incorporated in intelligent systems: optical fibers and piezoelectric materials.

Optical fibers embedded in a “smart” material can provide data in two ways. First, they can simply provide a steady light signal to a sensor; breaks in the light beam indicate a structural flaw that has snapped the fiber. The second, more subtle, approach involves looking at key characteristics of the light intensity, phase, polarization, or similar feature. The National Aeronautics and Space Administration and other research centers have used such a fiber-optic system to measure the strain in composite materials. Fiber-optic sensors can also measure magnetic fields, deformations, vibrations, and acceleration. Resistance to adverse environments and immunity to electrical or magnetic noise are among the advantages of optical sensors.

In addition to serving as actuators, piezoelectric materials make good sensors. Piezoelectric polymers, such as polyvinylidene fluoride (PVDF), are commonly exploited for sensing because they can be formed in thin films and bonded to many kinds of surfaces. The sensitivity of PVDF to pressure has proved suitable for sensors that can read braille and distinguish grades of sandpaper. Ultrathin PVDF films, perhaps 200 to 300 μm thick, have been proposed for use in robotics. Such a sensor might be used to replicate the capability of human skin, detecting temperature and geometric features such as edges and corners, or distinguishing between different fabrics.

Actuators and sensors are crucial elements in an intelligent materials system but the essence of this new design philosophy in the manifestation of the most critical of life functions, intelligence—the extent to which the material should be smart or merely adaptive—is debatable. At a minimum, there must be an ability to learn about the environment and live within it.

The thinking features that the intelligent materials community is trying to create have constraints that the engineering world has never experienced before. Specifically, the vast number of sensors and actuators and their associated power sources would argue against feeding all these devices into a central processor. Instead designers have taken clues from nature. Neurons are not nearly so fast as modern-day silicon chips, but they can nonetheless perform complex tasks with amazing speed because they are networked efficiently.

The key appears to be hierarchical architecture. Signal processing and the resulting action can take place at levels below and far removed from the brain. The reflex of moving your hand away from a hot stove, for example, is organized entirely within the spinal cord. Less automatic behaviors are organized by successively higher centers within the brain. Besides being efficient, such an organization is fault-tolerant: unless there is some underlying organic reason, we rarely experience a burning sensation when holding an iced drink.

The brains behind an intelligent materials system follow a similar organization. In fact, investigators take their cue from research into artificial life, an outgrowth of the cybernetics field. Among the trendiest control concepts is the artificial neural network, which is computer programming that mimics the functions of real neurons. Such software can learn, change in response to contingencies, anticipate needs, and correct mistakes—more than adequate functions for intelligent materials systems. Ultimately, computational hardware and the processing algorithms will determine how complex these systems can become—that is, how many sensors and actuators we can use.

ENGINEERING MICROSCOPIC MACHINES

The electronics industry relies on its ability to double the number of transistors on a microchip every 18 months a trend that drives the dramatic revolution in electronics. Manufacturing millions of microscopic elements in an area no larger than a postage stamp has now begun to inspire technology that reaches beyond the field that produced the pocket telephone and the personal computer.

Using the materials and processes of microelectronics, researchers have fashioned microscopic beams, pits, gears, membranes, and even motors that can be deployed to move atoms or to open and close valves that pump microliters of liquid. The size of these mechanical elements is measured in micrometers—a fraction of the width of a human hair. And, like transistors, millions of these elements can be fabricated at one time.

In the next 50 years, this structural engineering of silicon may have as profound an impact on society as did the miniaturization of electronics in preceding decades. Electronic computing and memory circuits, as powerful as they are, do nothing more than switch electrons and route them on their way over tiny wires. Micromechanical devices will supply electronic systems with a much-needed window to the physical world, allowing them to sense and control motion, light, sound, heat, and other physical forces.

The coupling of micromechanical and electronic systems will produce dramatic technical advances across diverse scientific and engineering disciplines. Thousands of beams with cross sections of less than a micrometer will move tiny electrical scanning heads that will read and write enough data to store a small library of information on an area the size of a microchip. Arrays of valves will release drug dosages into the bloodstream at precisely timed intervals. Inertial guidance systems on chips will aid in locating the positions of military combatants and direct munitions precisely at targets.

Nano-microelectromechanical systems (NANO-MEMS) is the name given to the practice of making and combining miniaturized mechanical and electronic components. NANO-MEMS devices are made using manufacturing processes that are similar, and in some cases identical, to those for electronic components.

SURFACE MICROMACHINING

One technique, called surface micromachining, parallels electronics fabrication so closely that it is essentially a series of steps added to the making of a microchip. Surface micromachining acquired its name because the small mechanical structures are "machined" onto the surface of a silicon disk known as a wafer. The technique relies on photolithography as well as other staples of the electronic manufacturing process that deposit or etch away small amounts of material on the chip.

Photolithography creates a pattern on the surface of a wafer, marking off an area that is subsequently etched away to build up micromechanical structures such as a motor or a freestanding beam. Manufacturers start by patterning and etching a hole in a layer of silicon dioxide deposited on the wafer. A gaseous vapor reaction then deposits a layer of polycrystalline silicon, which coats both the hole and the remaining silicon dioxide material. The silicon deposited into the hole becomes the base of the beam, and the same material that overlays the silicon dioxide forms the suspended part of the beam structure. In the final step, the remaining silicon dioxide is etched away, leaving the polycrystalline silicon beam free and suspended above the surface of the wafer.

Such miniaturized structures exhibit useful mechanical properties. When stimulated with an electrical voltage, a beam with a small mass will vibrate more rapidly than a heavier device, making it a more sensitive detector of motion, pressure, or even chemical properties. For instance, a beam could adsorb a certain chemical (adsorption occurs when thin layers of a molecule adhere to a surface). As more of the chemical is adsorbed, the weight of the beam changes, altering the frequency at which it would vibrate when electrically excited. This chemical sensor could therefore operate by detecting such changes in vibrational frequency. Another type of sensor that employs beams manufactured with surface micromachining functions on a slightly different principle. It changes the position of suspended parallel beams that make up an electrical capacitor—and thus alters the amount of stored electrical charge—when an automobile goes through the rapid deceleration of a crash. Analog Devices, a Massachusetts-based semiconductor company, manufactures this acceleration sensor to trigger the release of an air bag. The company has sold more than 25 million of these sensors to automobile makers over the past 15 years.

This air bag sensor may one day be looked back on as the microelectromechanical equivalent of the early integrated electronics chips. The fabrication of beams and other elements of the motion sensor on the surface of a silicon microchip has made it possible to produce this device on a standard integrated circuit fabrication line.

The codependence link of machines and sensors demonstrates that integrating more of these devices with electronic circuits will yield a window to the world of motion, sound, heat, and other physical forces.

The structures that serve as part of an acceleration sensor for triggering air bags are made by first depositing layers of silicon nitride (an insulating material) and silicon dioxide on the surface of a silicon substrate. Holes are lithographically patterned and etched into the silicon dioxide to form anchor points for the beams. A layer of polycrystalline silicon is then deposited. Lithography and etching form the pattern of the beams. Finally, the silicon dioxide is etched away to leave the freestanding beams.

In microelectronics the ability to augment continually the number of transistors that can be wired together has produced truly revolutionary developments: the microprocessors and memory chips that made possible small, affordable computing devices such as the personal computer. Similarly, the worth of MEMS may become apparent only when thousands or millions of mechanical structures are manufactured and integrated with electronic elements.

The first examples of mass production of microelectromechanical devices have begun to appear, and many others are being contemplated in research laboratories all over the world. An early prototype demonstrates how NANO-MEMS may affect the way millions of people spend their leisure time in front of the television set. Texas Instruments has built an electronic display in which the picture elements, or pixels, that make up the image are controlled by microelectromechanical structures. Each pixel consists of a 16-μm-wide aluminum mirror that can reflect pulses of colored light onto a screen. The pixels are turned off or on when an electric field causes the mirrors to tilt 10° to one side or the other. In one direction, a light beam is reflected onto the screen to illuminate the pixel. In the other, it scatters away from the screen, and the pixel remains dark.

MICROMIRROR DISPLAY

This micromirror display could project the images required for a large-screen television with a high degree of brightness and resolution of picture detail. The mirrors could compensate for the inadequacies encountered with other technologies. Display designers, for instance, have run into difficulty in making liquid-crystal screens large enough for a wall-sized television display.

The future of NANO-MEMS can be glimpsed by examining projects that have been funded during the past three years under a program sponsored by the U.S. Department of Defense's Advanced Research Projects Agency. This research is directed toward building a number of prototype microelectromechanical devices and systems that could transform not only weapons but also consumer products.

A team of engineers at the University of California at Los Angeles and the California Institute of Technology wants to show how NANO-MEMS may eventually influence aerodynamic design. The group has outlined its ideas for a technology that might replace the relatively large moving surfaces of a wing—the flaps, slats, and ailerons—that control both turning and ascent and descent. It plans to line the surface of a wing with thousands of 150-μm-long plates that, in their resting position, remain flat on the wing surface. When an electrical voltage is applied, the plates rise from the surface at up to a 90° angle. Thus activated, they can control the vortices of air that form across selected areas of the wing. Sensors can monitor the currents of air rushing over the wing and send a signal to adjust the position of the plates.

These movable plates, or actuators, function similarly to a microscopic version of the huge flaps on conventional aircraft. Fine-tuning the control of the wing surfaces would enable an airplane to turn more quickly, stabilize against turbulence, or burn less fuel because of greater flying efficiency. The additional aerodynamic control achieved with this "smart skin" could lead to radically new aircraft designs that move beyond the cylinder-with-wings appearance that has prevailed for 70 years. Aerospace engineers might dispense entirely with flaps, rudders, and even the wing surface, called a vertical stabilizer. The aircraft would become a kind of "flying wing," similar to the U.S. Air Force's Stealth bomber.

An aircraft without a vertical stabilizer would exhibit greater maneuverability—a boon for fighter aircraft and perhaps also one day for high-speed commercial airliners that must be capable of changing direction quickly to avoid collisions.

MICRO-MICROSCOPES

The engineering of small machines and sensors allows new uses for old ideas. For a decade, scientists have routinely worked with scanning probe microscopes that can manipulate and form images with individual atoms. The most well known of these devices is the scanning tunneling microscope (STM).

The STM, an invention for which Gerd Binnig and Heinrich Rohrer of IBM won the Nobel Prize in Physics in 1986, caught the attention of micromechanical specialists in the early 1980s. The fascination of the engineering community stems from calculations of how much information could be stored if STMs were used to read and write digital data. A trillion bits of information—equal to the text of 500 Encyclopedia Britannicas—might be fit into a square centimeter on a chip by deploying an assembly of multiple STMs.

The STM is a needle-shaped probe, the tip of which consists of a single atom. A current that "tunnels" from the tip to a nearby conductive surface can move small groups of atoms, either to create holes or to pile up tiny mounds on the silicon chip. Holes and mounds correspond to the zeros and ones required to store digital data. A sensor, perhaps one constructed from a different type of scanning probe microscope, would "read" the data by detecting whether a nanometer-sized plot of silicon represents a zero or a one.

Only beams and motors a few micrometers in size, and with a commensurately small mass, will be able to move an STM quickly and precisely enough make terabit (trillionbit) data storage on a chip practicable. With MEMS, thousands of STMs could be suspended from movable beams built on the surface of a chip, each one reading or writing data in an area of a few square micrometers. The storage medium, moreover, could remain stationary, which would eliminate the need for today's spinning media disk drives.

Noel C. MacDonald, an electrical engineering professor at Cornell University, has taken a step toward fulfilling the vision of the pocket research library. He has built an STM-equipped microbeam that can be moved in the vertical and horizontal axes or even at an oblique angle. The beam hangs on a suspended frame attached to four motors, each of which measures only 200 μm (two hair widths) across. These engines push or pull on each side of the tip at speeds as high a million times a second. MacDonald's next plan is to build an array of STMS.

THE PERSONAL SPECTROPHOTOMETER

The Lilliputian infrastructure afforded by MEMS might let chemists and biologists perform their experiments with instruments that fit in the palm of the hand. Westinghouse Science and Technology Center is in the process of reducing to the size of a calculator a 50-lb benchtop spectrometer used for measuring the mass of atoms or molecules. A miniaturized mass spectrometer presages an era of inexpensive chemical detectors for do-it-yourself toxic monitoring.

In the same vein, Richard M. White, a professor at the University of California at Berkeley, contemplates a chemical factory on a chip. White has begun to fashion millimeter-diameter wells each of which holds a different chemical, in a silicon chip. An electrical voltage causes liquids or powders to move from the wells down a series of channels into a reaction chamber. These materials are pushed there by micropumps made of piezoelectric materials that constrict and then immediately release sections of the channel. The snakelike undulations create a pumping motion. Once the chemicals are in the chamber, a heating plate causes them to react. An outlet channel from the chamber then pumps out what is produced in the reaction.

POCKET CALCULATOR CHEMICAL FACTORY

A pocket-calculator-sized chemical factory could thus reconstitute freeze-dried drugs, perform DNA testing to detect waterborne pathogens, or mix chemicals that can then be converted into electrical energy more efficiently than can conventional batteries. MEMS gives microelectronics an opening to the world beyond simply processing and storing information. Automobiles, scientific laboratories, televisions, airplanes, and even the home medicine cabinet will never be the same.

"THERE IS NO NEW THING UNDER THE SUN!"

—Ecclesiastes 1:3–10, 2:11–13, 12:12–13

All the rivers run into the sea; yet the sea is not full; unto the place from whence the rivers come, thither they return again. All things are full of labour, man cannot utter it: the eye is not satisfied with seeing, nor the ear filled with hearing. The thing that hath been, it is that which shall be; and that which is done is that which shall be done: and there is no new thing under the sun. Is there any thing whereof it may be said, See, this is new? It hath been already of old time, which was before us....I looked on all the works that my hands had wrought, and on the labour that I had laboured to do: and, behold, all was vanity and vexation of spirit, and there was no profit under the sun. My son, be admonished: of making many books there is no end; and much study is a weariness of the flesh. Let us hear the conclusion of the whole matter: Fear God, and keep his commandments: for this is the whole duty of man.

Sabrie Soloman, Ph.D., Sc.D., MBA, PE

Chairman and CEO American SensoRx, Inc., USA
Chairman and CEO, SensoRx, Chongqing, Ltd., China.
Professor, Founder, Advanced Manufacturing Technology,
Columbia University, USA

ACKNOWLEDGMENTS

Mere thanks is insufficient to Harleen Chopra, project manager; Michael McGee, copyeditor; P. Prasanna, compositor; and Kritika Gupta and Mitali Srivastav, proofreaders for their relentless efforts, and unfaltering diligence to bring the work of this handbook to a high standard of world-class scientific and technical publication.

Furthermore, Mr. Taisuke Soda, Senior Editor, McGraw-Hill Professional, New York, has my sincere gratitude for his insight and distinctive wisdom in guiding me through the composition of this handbook.

Also, this handbook was made possible by the efforts of my colleagues and friends in various universities and industries, and by the encouragement, and the assistance of my staff at Columbia University, New York, American SensoRx, Inc., USA, and SensoRx, Chongqing, Ltd., China.

Abundant appreciation is not enough to my beloved wife, Elvira Soloman, for her immeasurable efforts assisting me to compile the most advanced Sensors Handbook technologies in numerous fields and applications; varying from industrial, medical, biological, and military fields, in addition to the spectroscopy and vision sensors applications; in particularly, the SpectRx™ near infrared and the InspectRx® vision sensing systems. Elvira comforted my mind, and provided me the tranquility to write and compile the information presented in this handbook. Her sacrifices and devotions solely made this handbook possible to the world.

Also, it was an exceedingly humbling experience to watch my little child, Stephan Soloman, attempting to climb my knees hoping to have a glimpse of freshly composed sentences, while I was working on this handbook within the past 2 years. Even prior to his bedtime, he was often moved to raise innocent and diligent prayers of encouragement and for the success of my work on this handbook.

This page intentionally left blank

INTRODUCTION

Integrated sensors and control systems are the way of the future. In times of disaster, even the most isolated outposts can be linked directly into the public telephone network by portable versions of satellite earth stations called *very small aperture terminals* (VSATs). They play a vital role in relief efforts such as those for the eruption of Mount Pinatubo in the Philippines, the massive oil spill in Valdez, Alaska, the 90,000-acre fire in the Idaho forest, and Hurricane Andrew's destruction in south Florida and the coast of Louisiana.

LIDAR (light detection and ranging) is an optical remote sensing technology that measures properties of scattered light to find range and/or other information of a distant target. The prevalent method to determine distance to an object or surface is to use laser pulses. Like the similar radar technology, which uses radio waves instead of light, the range to an object is determined by measuring the time delay between transmission of a pulse and detection of the reflected signal. LIDAR technology has application in archaeology, geography, geology, geomorphology, seismology, remote sensing, and atmospheric physics.

VSATs are unique types of sensors and control systems. They can be shipped and assembled quickly and facilitate communications by using more powerful antennas that are much smaller than conventional satellite dishes. These types of sensors and control systems provide excellent alternatives to complicated conventional communication systems, which in disasters often experience serious degradation because of damage or overload.

Multispectral sensors and control systems will play an expanding role to help offset the increasing congestion on America's roads by creating "smart" highways. At a moment's notice, they can gather data to help police, tow trucks, and ambulances respond to emergency crises. Understanding flow patterns and traffic composition would also help traffic engineers plan future traffic control strategies. The result of less congestion will be billions of vehicle hours saved each year.

In Fig. I.1, the Magellan spacecraft is close to completing its third cycle of mapping the surface of planet Venus. The key to gathering data is the development of a synthetic aperture radar as a sensor and information-gathering control system, the sole scientific instrument aboard Magellan. Even before the first cycle ended, in mid-1991, Magellan had mapped 84 percent of Venus' surface, returning more digital data than all previous U.S. planetary missions combined, with resolutions ten times better than those provided by earlier missions. To optimize radar performance, a unique and simple computer software program was developed, capable of handling nearly 950 commands per cycle. Each cycle takes a Venusian day, the equivalent of 243 Earth days.

Manufacturing organizations in the United States are under intense competitive pressure. Major changes are being experienced with respect to resources, markets, manufacturing processes, and product strategies. As a result of international competition, only the most productive and cost-effective industries will survive.

Today's sensors, remote sensors, and control systems have explosively expanded beyond their traditional production base into far-ranging commercial ventures. They will play an important role in the survival of innovative industries. Their role in information

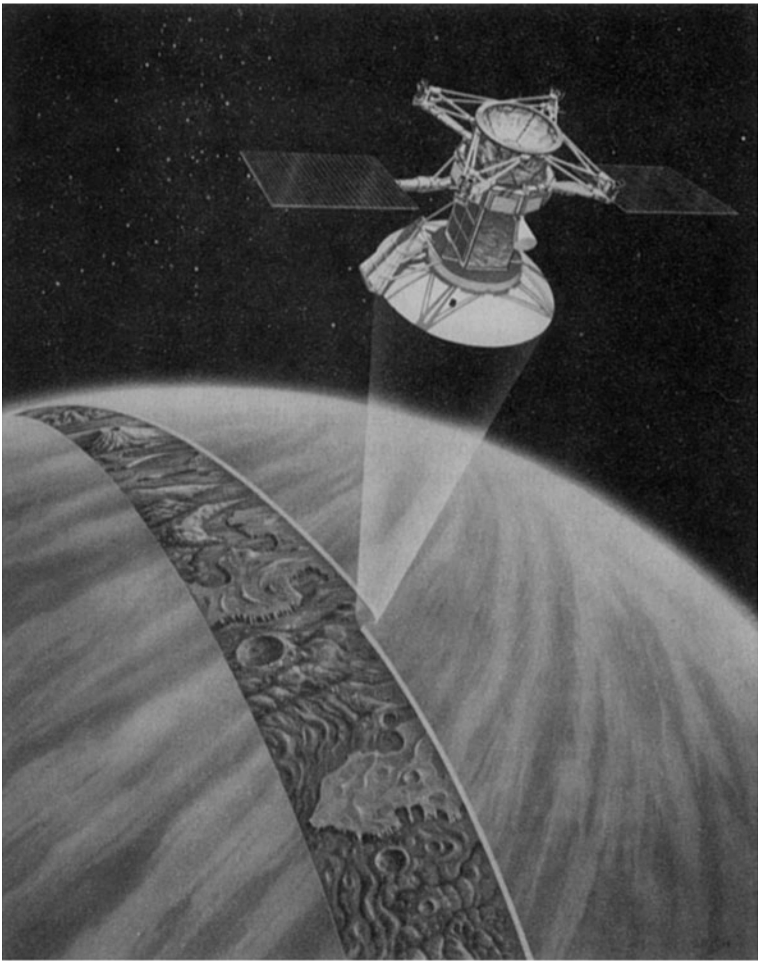


FIGURE I.1 The Magellan Spacecraft. (*Courtesy Hughes Corp.*)

assimilation, and control of operations to maintain an error-free production environment, will help enterprises to stay effective on their competitive course.

ESTABLISHING AN AUTOMATION PROGRAM

Manufacturers and vendors have learned the hard way that technology alone does not solve problems. A prime example is the gap between the information and the control worlds, which caused production planners to set their goals according to dubious assumptions concerning plant-floor activities, preventing plant supervisors from isolating production problems until well after they had arisen.

The problem of creating effective automation for an error-free production environment has drawn a long list of solutions. Some are as old as the term *computer-integrated* *manufacturing* (CIM) itself. However, in many cases, the problem turned out not to be technology but the ability to integrate equipment, information, and people.

The debate over the value of agile manufacturing and computer-integrated manufacturing technology has been put to rest, although executives at every level in almost every industry are still questioning the cost of implementing CIM solutions. Recent economic belt tightening has forced industry to justify every capital expense, and CIM has drawn fire from budget-bound business people in all fields.

Too often, the implementations of CIM have created a compatibility nightmare in today's multivendor factory-floor environments. Too many end users have been forced to discard previous automation investments and/or spend huge sums on new equipment, hardware, software, and networks in order to effectively link together data from distinctly dissimilar sources. The expense of compatible equipment and the associated labor cost for elaborate networking are often prohibitive.

The claims of CIM open systems are often misleading. This is largely due to proprietary concerns, a limited-access database, and operating system compatibility restrictions. The systems fail to provide the transparent integration of process data and plant business information that makes CIM work.

In order to solve this problem, it is necessary to establish a clearly defined automation program. A common approach is to limit the problem description to a workable scope, eliminating the features that are not amenable to consideration. The problem is examined in terms of a simpler workable model. A solution can then be based on model predictions.

The danger associated with this strategy is obvious: If the simplified model is not a good approximation of the actual problem, the solution will be inappropriate and may even worsen the problem.

Robust automation programs can be a valuable asset in deciding how to solve production problems. Advances in sensor technology have provided the means to make rapid large-scale improvements in problem solving and have contributed in essential ways to today's manufacturing technology.

The infrastructure of an automation program must be closely linked with the use and implementation of sensors and control systems within the framework of the organization. The problem becomes more difficult whenever it is extended to include the organizational setting. Organization theory is based on a fragmented and partially developed body of knowledge, and can provide only limited guidance in the formation of problem models. Managers commonly use their experience and instinct in dealing with complex production problems that include organizational aspects. As a result, creating a competitive manufacturing enterprise—one involving advanced automation technology utilizing sensors and control systems and organizational aspects—is a task that requires an understanding of both how to establish an automation program and how to integrate it with a dynamic organization.

In order to meet the goals of integrated sensory and control systems, an automated manufacturing system has to be built from compatible and intelligent subsystems. Ideally, a manufacturing system should be computer-controlled and should communicate with controllers and materials-handling systems at higher levels of the hierarchy, as shown in Fig. I.2.

UNDERSTANDING FLEXIBLE WORKSTATIONS, FLEXIBLE WORK CELLS, AND FLEXIBLE WORK CENTERS

Flexible workstations, flexible work cells, and flexible work centers represent a coordinated cluster of a production system. A production machine with several processes is considered a workstation. A machine tool is also considered a workstation. Integrated workstations form

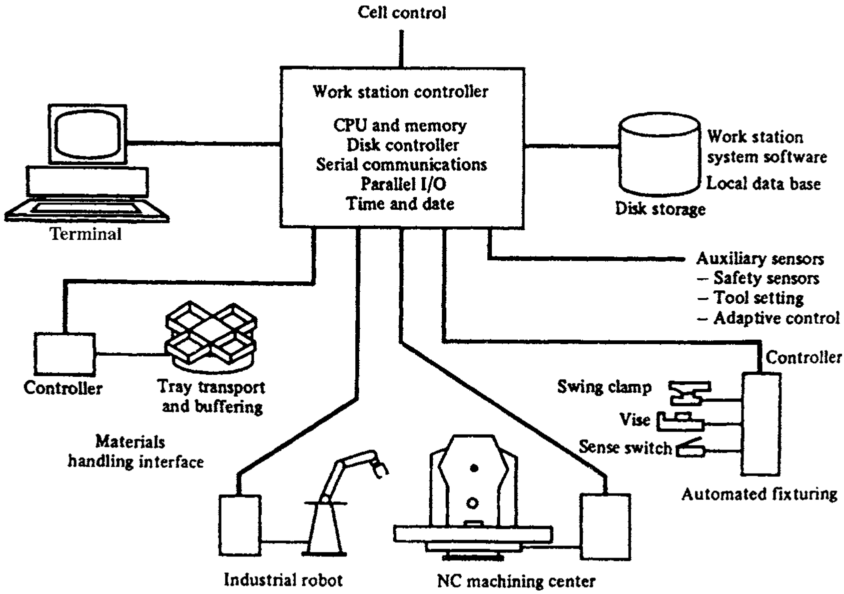


FIGURE I.2 Computer-controlled manufacturing system.

a work cell. Several complementary workstations may be grouped together to construct a work cell. Similarly, integrated work cells may form a work center. This structure is the basic concept in modeling a flexible manufacturing system. The flexible manufacturing system is also the cornerstone of the computer-integrated manufacturing strategy (Fig. I.3).

The goal is to provide the management and project development team with an overview of major tasks to be solved during the planning, design, implementation, and operation phases of computer-integrated machining, inspection, and assembly systems. Financial and technical disasters can be avoided if a clear understanding of the role of sensors and control systems in the computer-integrated manufacturing strategy is asserted.

Sensors are largely applied within the workstations, and are the only practical means of operating a manufacturing system and tracking its performance continuously.

Sensors and control systems in manufacturing provide the means of integrating different, properly defined processes as input to create the expected output. Input may be raw material and/or data that have to be processed by various auxiliary components such as tools, fixtures, and clamping devices. Sensors provide the feedback data to describe the status of each process. The output may also be data and/or materials, which can be processed by further cells of the manufacturing system. A flexible manufacturing system that contains workstations, work cells, and work centers and is equipped with appropriate sensors and control systems is a distributed management information system, linking together subsystems of machining, packaging, welding, painting, flame cutting, sheet-metal manufacturing, inspection, and assembly with material-handling and storage processes.

In designing various workstations, work cells, and work centers in a flexible manufacturing system within the computer-integrated manufacturing strategy, the basic task is to create a variety of sensors interconnecting different material-handling systems, such as robots, automated guided-vehicle systems, conveyers, and pallet loading and unloading carts, to allow them to communicate with data processing networks for successful integration with the system.

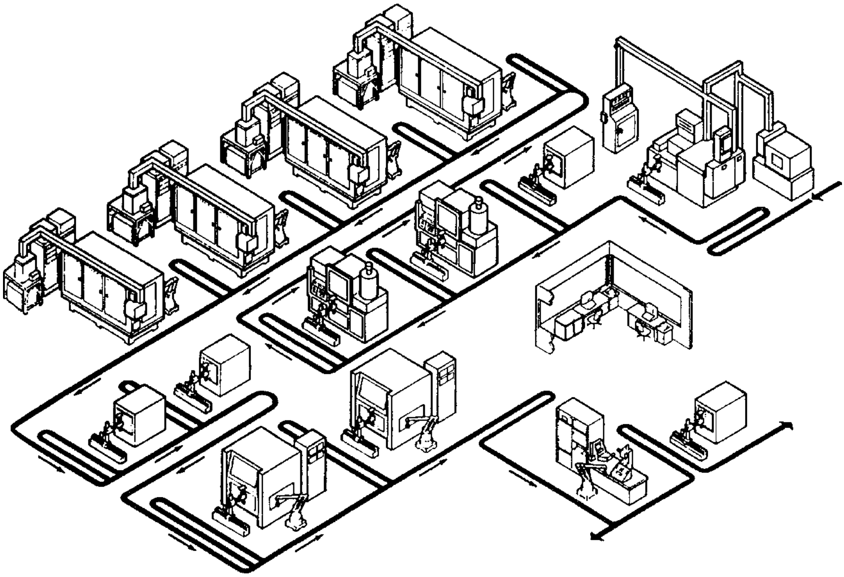


FIGURE I.3 Workstation, work cell, and work center.

Figure I.4 illustrates a cell consisting of several workstations with its input and output, and indicates its basic functions in performing the conversion process, storing workpieces, linking material-handling systems to other cells, and providing data communication to the control system.

The data processing links enable communication with the databases containing part programs, inspection programs, robot programs, packaging programs, machining data, and real-time control data through suitable sensors. The data processing links also enable communication of the feedback data to the upper level of the control hierarchy. Accordingly, the entire work-cell facility is equipped with current data for real-time analysis and for fault recovery.

A cluster of manufacturing cells grouped together for particular production operations is called a *work center*. Various work centers can be linked together via satellite communication links irrespective of the location of each center. Manufacturing centers can be located several hundred feet apart or several thousand miles apart. Adequate sensors and control systems together with effective communication links will provide practical real-time data analysis for further determination.

The output of the flexible cell is the product of the module of the flexible manufacturing system. It consists of a finished or semifinished part as well as data in a computer-readable format that will instruct the next cell on how to achieve its output requirement. The data are conveyed through the distributed communication networks. If, for example, a part is required to be surfaced to a specific datum in a particular cell, sensors will be adjusted to read the required acceptable datum during the surfacing process. Once the operation is successfully completed, the part must once again be transferred to another cell for further machining or inspection processes. The next cell is not necessarily physically adjacent; it may be the previous cell, for instance, as programmed for the required conversion process.

The primary reason for the emphasis on integrating sensors and control systems into every manufacturing operation is the worldwide exponentially increasing demand for

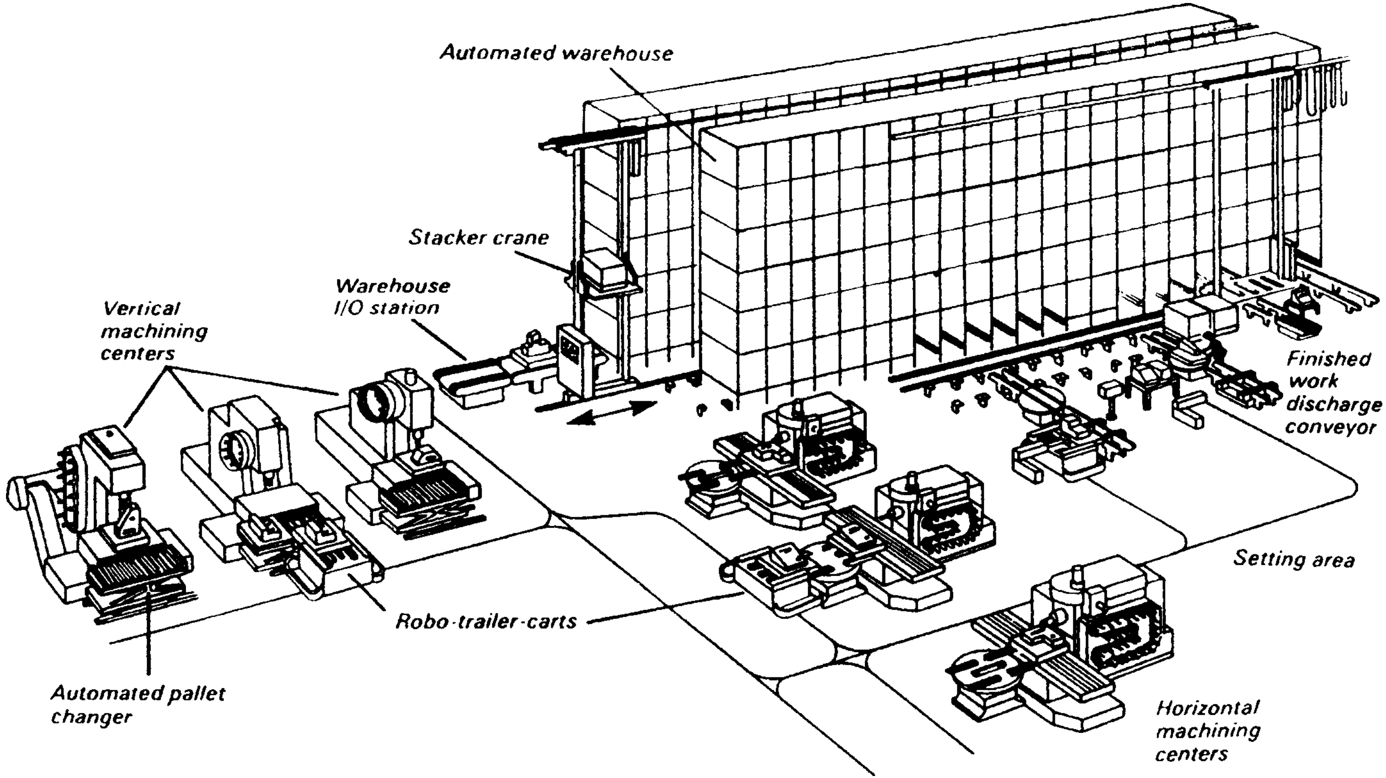


FIGURE I.4 Conversion process in a manufacturing cell.

error-free production operations. Sensors and control technology can achieve impressive results only if effectively integrated with corporate manufacturing strategy.

The following benefits can be achieved:

- *Productivity.* A greater output and a lower unit cost.
- *Quality.* Product is more uniform and consistent.
- *Production reliability.* The intelligent self-correcting sensory and feedback system increases the overall reliability of production.
- *Lead time.* Parts can be randomly produced in batches of one or in reasonably high numbers, and the lead time can be reduced by 50 to 75 percent.
- *Expenses.* Overall capital expenses are 5 to 10 percent lower. The cost of integrating sensors and feedback control systems into the manufacturing source is less than that of stand-alone sensors and feedback systems.
- *Greater utilization.* Integration is the only available technology with which a machine tool can be utilized as much as 85 percent of the time—and the time spent cutting can also be over 90 percent.

In contrast, a part, from stock to finished item, spends only 5 percent of its time on the machine tool, and actual productive work takes only 30 percent of this 5 percent. The time for useful work on stand-alone machines without integrated sensory and control systems is as little as 1 to 1.5 percent of the time available (see Tables I.1 and I.2).

To achieve the impressive results indicated in Table I.1, the integrated manufacturing system carrying the sensory and control feedback systems must maintain a high degree of flexibility. If any cell breaks down for any reason, the production planning and control system can reroute and reschedule the production or, in other words, reassign the system

TABLE I.1 Time Utilization of Integrated Manufacturing Center Carrying Sensory and Control Systems

	Active, %	Idle, %
Tool positioning and tool changing	25	
Machining process	5	
Loading and inspection	15	
Maintenance	20	
Setup	15	
Idle time		15
Total	85	15

TABLE I.2 Productivity Losses of Stand-alone Manufacturing Center Excluding Sensory and Control Systems

	Active, %	Idle, %
Machine tool in wait mode		35
Labor control		35
Support services		15
Machining process	15	
Total	15	85

environment. This can be achieved only if both the processes and the routing of parts are programmable. The sensory and control systems will provide instantaneous descriptions of the status of parts to the production and planning system.

If different processes are rigidly integrated into a special-purpose highly productive system, such as a transfer line for large batch production, then neither modular development nor flexible operation is possible.

However, if the cells and their communication links to the outside world are programmable, much useful feedback data may be gained. Data on tool life, measured dimensions of machined surfaces by in-process gauging and production control, and fault recovery derived from sensors and control systems can enable the manufacturing system to increase its own productivity, learn its own limits, and inform the part programmers of them. The data may also be very useful to the flexible manufacturing system designers for further analysis. In non-real-time control systems, the data cannot usually be collected, except by manual methods, which are time-consuming and unreliable.

CHAPTER 1
TYPES AND CLASSIFICATIONS OF SENSORS AND CONTROL SYSTEMS

CLASSIFICATION OF CONTROL PROCESSES

An *engineering integrated system* can be defined as a machine responsible for certain production output, a controller to execute certain commands, and sensors to determine the status of the production processes. The machine is expected to provide a certain product as an output, such as computer numerical control (CNC) machines, packaging machines, and high-speed press machines. The controller provides certain commands arranged in a specific sequence and designed for a particular operation. The controller sends its commands in the form of signals, usually electric pulses. The machine is equipped with various devices, such as solenoid valves and step motors, that receive the signals and respond according to their functions. The sensors provide a clear description of the status of the machine performance. They give detailed accounts of every process in the production operation (Fig. 1.1).

Once a process is executed successfully, according to a specific sequence of operations, the controller can send additional commands for further processes until all processes are executed. This completes one cycle. At the end of each cycle, a command is sent to begin a new loop until the production demand is met.

In an automatic process, the machine, the controller, and the sensors interact with one another to exchange information. Mainly, two types of interaction occur between the controller and the rest of the system: through either an open-loop control system or a closed-loop control system.

An *open-loop control system* (Fig. 1.2) can be defined as a system in which there is no feedback. Motor motion is expected to faithfully follow the input command. Stepping motors are an example of open-loop control.

A *closed-loop control system* (Fig. 1.3) can be defined as a system in which the output is compared to the command, with the result being used to force the output to follow the command. Servo systems are an example of closed-loop control.

OPEN- AND CLOSED-LOOP CONTROL SYSTEMS

In an open-loop control system, the actual value in Fig. 1.2 may differ from the reference value in the system. In a closed-loop system, the actual value is constantly monitored against the reference value described in Fig. 1.3.

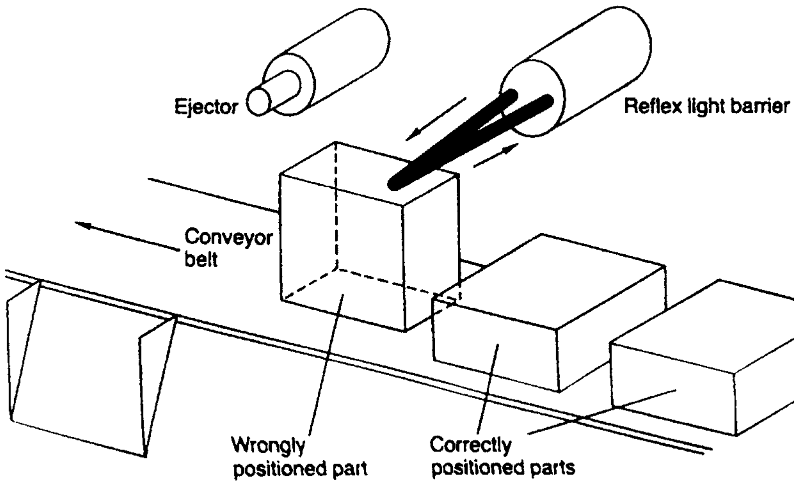


FIGURE 1.1 Sensors providing machine status.

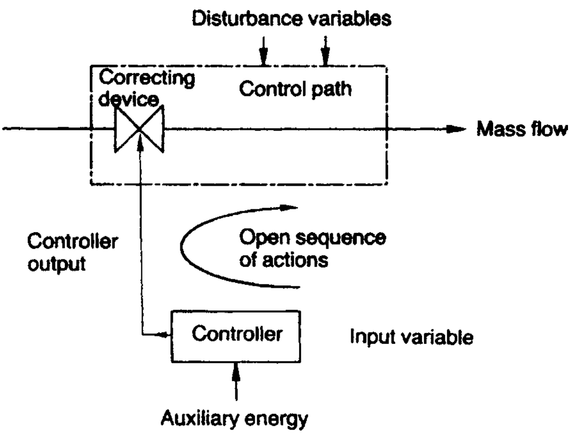


FIGURE 1.2 An open-loop control system.

The mass flow illustrated in Fig. 1.4 describes the amount of matter per unit time flowing through a pipeline that must be regulated. The current flow rate can be recorded by a measuring device, and a correcting device such as a valve may be set to a specific flow rate. The system, if left on its own, may suffer fluctuations and disturbances which will change the flow rate. In such an open-loop system, the reading of the current flow rate is the *actual value,* and the *reference value* is the desired value of the flow rate. The reference value may differ from the actual value, which then remains unaltered.

If the flow rate falls below the reference value because of a drop in pressure, as illustrated in Fig. 1.5, the valve must be opened further to maintain the desired actual value. Where disturbances occur, the course of the actual value must be continuously observed. When adjustment is made to continuously regulate the actual value, the loop of action governing measurement, comparison, adjustment, and reaction within the process is called a *closed loop*.

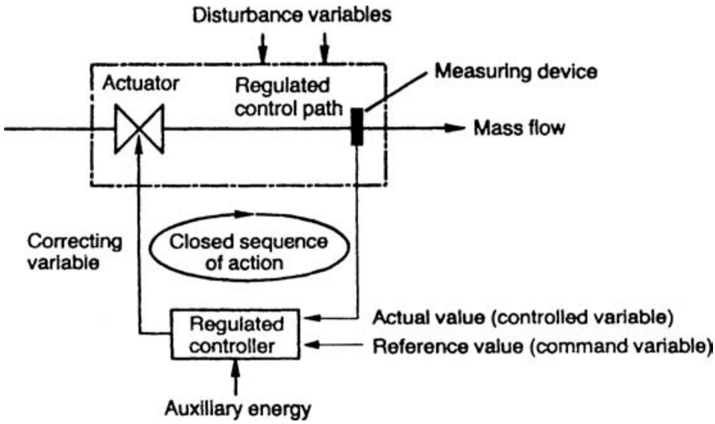


FIGURE 1.3 A closed-loop control system.

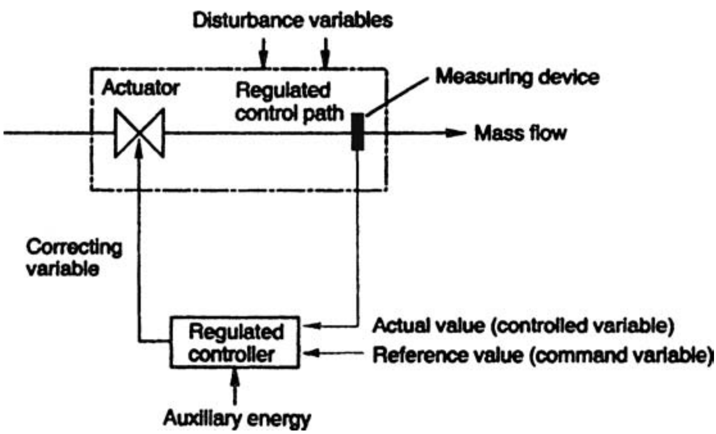


FIGURE 1.4 Regulation of mass flow.

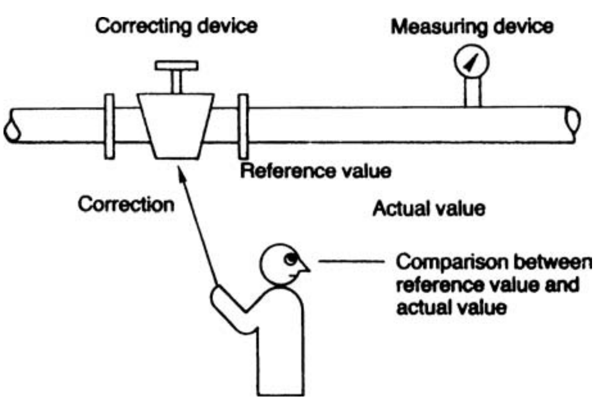


FIGURE 1.5 Reference value.

UNDERSTANDING PHOTOELECTRIC SENSORS

In order to successfully automate a process, it is necessary to obtain information about its status. The sensors are the integral part of the control system. The control system is responsible for collecting and preparing process status data and then passing it onto a processor (Fig. 1.6).

Principles of Operation

Photoelectric controls use light to detect the presence or absence of an object. All photoelectric controls consist of a sensor, a control unit, and an output device. A logic module or other accessories can be added to the basic control to add versatility. The sensor consists of a source and a detector. The source is a light-emitting diode (LED) that emits a powerful beam of light either in the infrared or visible light spectrum. The detector is typically a photodiode that senses the presence or absence of light. The detection amplifier in all photoelectric controls is designed so it responds to the light emitted by the source; ambient light, including sunlight up to 3100 metercandles, does not affect operation.

The source and detector may be separated or may be mounted in the same sensor head, depending on the particular series and application (Fig. 1.7).

The control unit modulates and demodulates the light sent and received by the source and detector. This assures that the photoelectric control responds only to its light source. The control unit also controls the output device in *self-contained* photoelectric controls; the control unit and sensor are built into an integral unit.

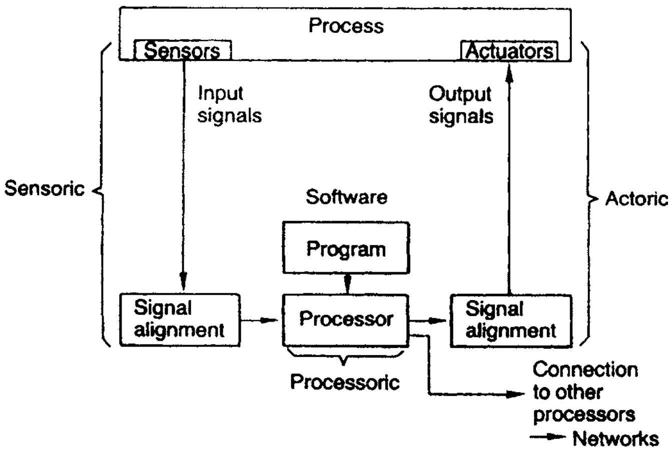


FIGURE 1.6 Components of controlled process.

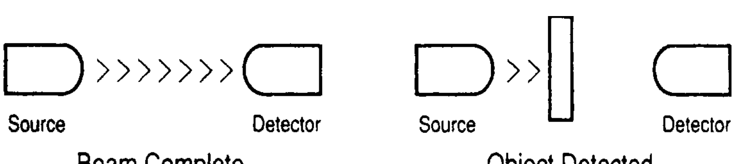


FIGURE 1.7 Components of photoelectric control.

Controls can be configured to operate as light-actuated devices. The output is triggered when the detector sees light. They can also be dark-actuated devices, where the output is triggered when the detector does not see light.

Output devices may include relays such as *double pole, double throw* (DPDT) and *single pole, double throw* (SPDT). Output devices may also include a triac or other high-current device and may be programmable-controller–compatible.

Logic modules are optional devices that allow the addition of logic functions to a photoelectric control. For example, instead of providing a simple ON/OFF signal, a photoelectric control can (with a logic module) provide time-delay, one-shot, retriggerable one-shot, motion-detection, and counting functions.

Manufacturing Applications of Photodetectors

The following applications of photoelectric sensors are based on normal practices at home, at the workplace, and in various industries. The effective employment of photoelectric sensors can lead to successful integration of data in manufacturing operations to maintain an error-free environment and assist in obtaining instantaneous information for dynamic interaction.

A photoelectric sensor is a semiconductor component that reacts to light or emits light. The light may be either in the visible range or the invisible infrared range. These characteristics of photoelectric components have led to the development of a wide range of photoelectric sensors.

A photoelectric reflex sensor equipped with a time-delay module set for *delay dark* ignores momentary beam breaks. If the beam is blocked longer than the predetermined delay period, the output energizes to sound an alarm or stop the conveyer (Fig. 1.8).

A set of photoelectric through-beam sensors can determine the height of a scissor lift, as illustrated in Fig. 1.9. For example, when the control is set for *dark-to-light* energizing, the lift rises after a layer has been removed and stops when the next layer breaks the beam again.

Cans on a conveyer are diverted to two other conveyers controlled by a polarized photoelectric reflex sensor with a divider module (Fig. 1.10). Items can be counted and diverted in groups of 2, 6, 12, or 24. A polarized sensor is used so that shiny surfaces may not falsely trigger the sensor control.

Two photoelectric control sensors can work together to inspect a fill level in cartons on a conveyer (Fig. 1.11). A reflex photoelectric sensor detects the position of the carton and energizes another synchronized photoelectric sensor located above the contents. If the photoelectric sensor located above the carton does not "see" the fill level, the carton does not pass inspection.

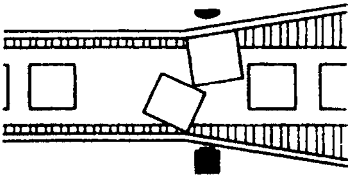


FIGURE 1.8 Jam detection with photoelectric sensor.

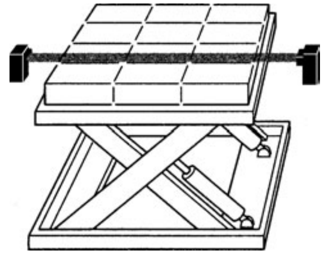


FIGURE 1.9 Stack height measurement with photoelectric sensor.

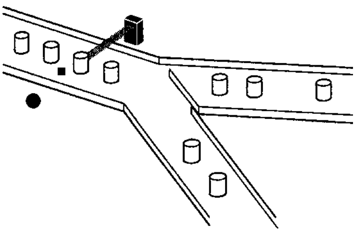


FIGURE 1.10 Batch counting and diverting with photoelectric sensor.

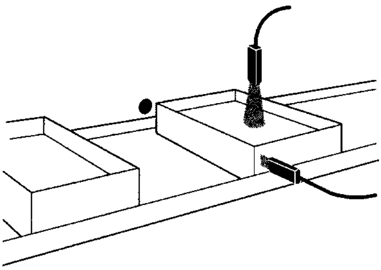


FIGURE 1.11 Measuring carton fill with photoelectric sensor.

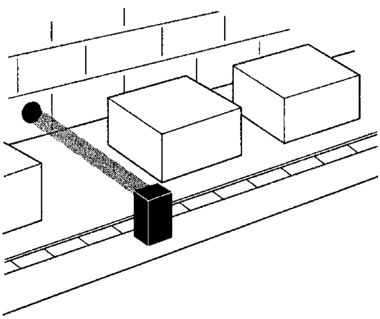


FIGURE 1.12 Box counting with photoelectric sensor.

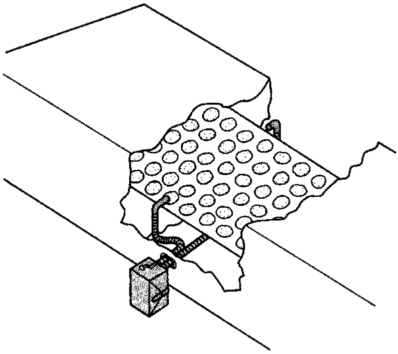


FIGURE 1.13 Detecting proper cookie arrangement.

A single reflex photoelectric sensor detects boxes anywhere across the width of a conveyer. Interfacing the sensor with a programmable controller provides totals at specific time intervals (Fig. 1.12).

High-temperature environments are accommodated by the use of fiber optics. The conveyer motion in a 450°F cookie oven can be detected as shown in Fig. 1.13. If the motion stops, the one-shot logic module detects light or dark that lasts too long, and the output device shuts the oven down.

Placing the photoelectric sensor to detect a saw tooth (Fig. 1.14) enables the programmable controller to receive an input signal that rotates the blade into position for sharpening of the next tooth.

A through-beam photoelectric sensor is used to time the toll gate in Fig. 1.15. To eliminate toll cheating, the gate lowers the instant the rear of the paid car passes the control. The rugged sensor can handle harsh weather, abuse, and 24-hour operation.

A safe and secure garage is achieved through the use of a through-beam photoelectric sensor interfaced to the door controller. The door shuts automatically after a car leaves, and if the beam is broken while the door is lowering, the motor reverses direction and raises the door again (Fig. 1.16).

A photoelectric sensor that generates a "curtain of light" detects the length of a loop on a web drive system by measuring the amount of light returned from an array of retroreflectors. With this information, the analog control unit instructs a motor controller to speed up or slow down the web drive (Fig. 1.17).

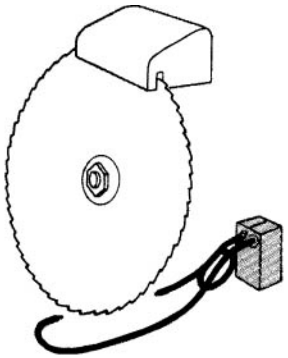


FIGURE 1.14 Sawtooth inspection.

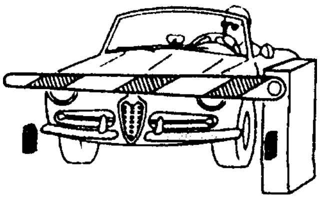


FIGURE 1.15 Toll-booth control with photoelectrical sensor.

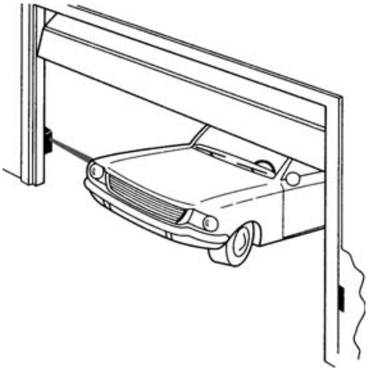


FIGURE 1.16 Garage door control with photoelectric control.

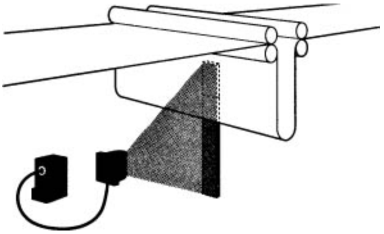


FIGURE 1.17 Web loop control.

Small objects moving through a curtain of light are counted by a change in reflected light. A low-contrast logic module inside the photoelectric sensor unit responds to slight but abrupt signal variations while ignoring slow changes such as those caused by dust buildup (Fig. 1.18).

A pair of through-beam photoelectric sensors scan over and under multiple strands of thread. If a thread breaks and passes through one of the beams, the low-contrast logic module detects the sudden changes in signal strength and energizes the output. Because this logic module does not react to slow changes in signal strength, it can operate in a dusty environment with little maintenance (Fig. 1.19).

A remote photoelectric source and detector pair looks for the passage of light through a hypodermic needle (Fig. 1.20). The small waterproof stainless-steel housing is appropriate for crowded machinery spaces and frequent wash-downs. High signal strength allows quality inspection of hole sizes down to 0.015 mm.

Index marks on the edge of a paper are detected by a fiber-optic photoelectric sensor to control a cutting shear on production lines (Fig. 1.21).

Liquids are monitored in a clear tank through beam sensors and an analog control. Because the control produces a voltage signal proportional to the amount of detected light, liquid mixtures and densities can be controlled (Fig. 1.22).

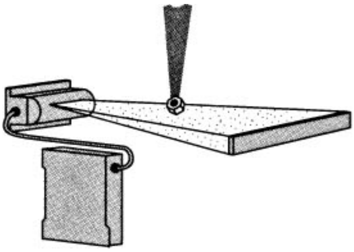


FIGURE 1.18 Small parts detection.

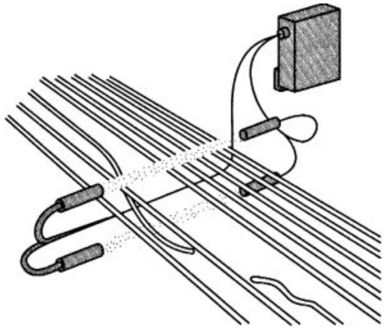


FIGURE 1.19 Broken-thread detection.

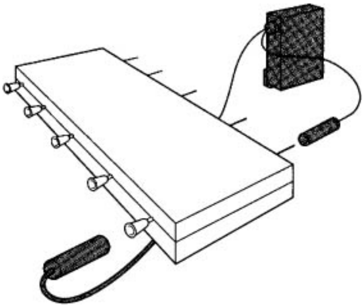


FIGURE 1.20 Hypodermic needle quality assurance.

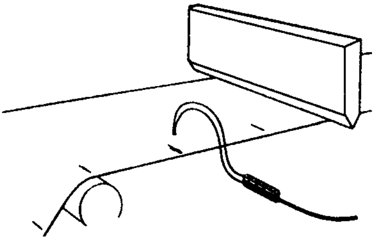


FIGURE 1.21 Indexing mark detection.

Remote photoelectric sensors inspect for the presence of holes in a metal casting (Fig. 1.23). Because each hole is inspected, accurate information is recorded. A rugged sensor housing and extremely high signal strength handle dirt and grease with minimum maintenance. The modular control unit allows for dense packaging in small enclosures.

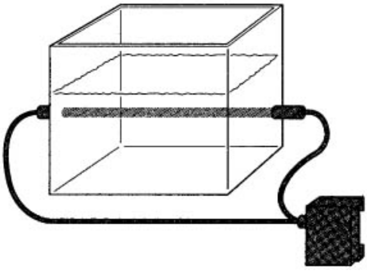


FIGURE 1.22 Liquid clarity control.

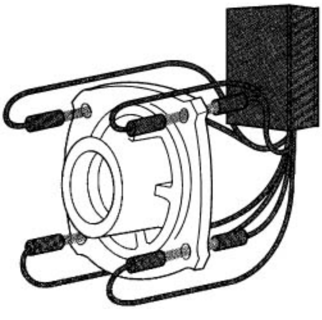


FIGURE 1.23 Multihole casting inspection.

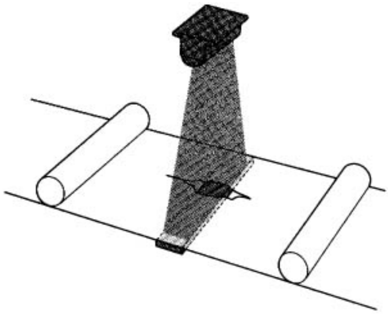


FIGURE 1.24 Web flaw detection.

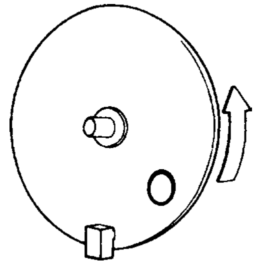


FIGURE 1.25 Over/underspeed of rotating disk.

In a web flaw detection application, a web passes over an array of retroreflectors (Fig. 1.24). When light is returned to the sensor head, the output is energized and the web shuts down. High web speeds can be maintained because of the superior response time of the control unit.

A reflex photoelectric sensor with a motion control module counts the revolutions of a wheel to monitor over/underspeed of a rotating object. Speed is controlled by a programmable controller. The rate ranges from 2.4 to 12,000 counts per minute (Fig. 1.25).

When the two through-beam photoelectric sensors in Fig. 1.26 observe the same signal strength, the output is zero. When the capacity of the web changes, as in a splice, the signal strengths are thrown out of balance and the output is energized. This system can be used on webs of different colors and opacities with no system reconfiguration.

Understanding the environment is important to effective implementation of an error-free environment. An awareness of the characteristics of photoelectric controls and the different ways in which they can be used will establish a strong foundation. This understanding also will allow the user to obtain a descriptive picture of the condition of each manufacturing process in the production environment.

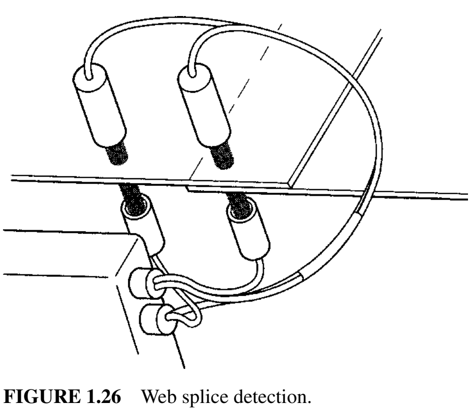


FIGURE 1.26 Web splice detection.

TABLE 1.1 Key Characteristics of Sensors

	Key point	Consideration
1.	Range	How far is the object to be detected?
2.	Environment	How dirty or dark is the environment?
3.	Accessibility	What accessibility is there to both sides of the object to be detected?
4.	Wiring	Is wiring possible to one or both sides of the object?
5.	Size	What size is the object?
6.	Consistency	Is object consistent in size, shape, and reflectivity?
7.	Requirements	What are the mechanical and electrical requirements?
8.	Output signal	What kind of output is needed?
9.	Logic functions	Are logic functions needed at the sensing point?
10.	Integration	Is the system required to be integrated?

Table 1.1 highlights key questions the user must consider.

DETECTION METHODS

Three modes of detection are used by photoelectric sensors:

- Through-beam detection
- Reflex detection
- Proximity detection

The Through-Beam Detection Method

The through-beam method requires that the source and detector are positioned opposite each other and the light beam is sent directly from source to detector (Fig. 1.27). When an object passes between the source and detector, the beam is broken, signaling detection of the object.

Through-beam detection generally offers the longest range of the three operating modes and provides high power at shorter range to penetrate steam, dirt, or other contaminants between the source and detector. Alignment of the source and detector must be accurate.

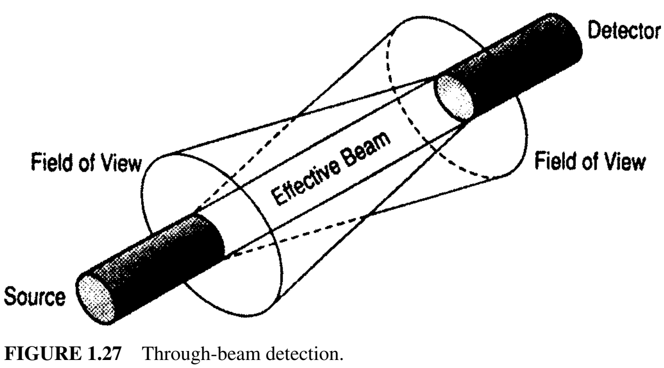


FIGURE 1.27 Through-beam detection.

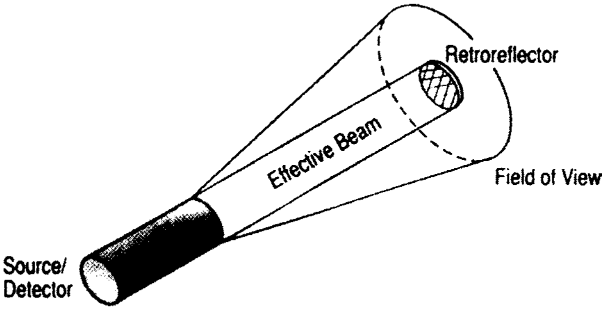


FIGURE 1.28 Reflex detection.

The Reflex Detection Method

The reflex method requires that the source and detector be installed on the same side of the object to be detected (Fig. 1.28). The light beam is transmitted from the source to a retroreflector that returns the light to the detector. When an object breaks a reflected beam, the object is detected.

The reflex method is widely used because it is flexible and easy to install and provides the best cost-performance ratio of the three methods. The object to be detected must be less reflective than the retroreflector.

The Proximity Detection Method

The proximity method requires that the source and detector are installed on the same side of the object to be detected and aimed at a point in front of the sensor (Fig. 1.29). When an object passes in front of the source and detector, light from the source is reflected from the object's surface back to the detector, and the object is detected.

Each sensor type has a specific operating range. In general, through-beam sensors offer the greatest range, followed by reflex sensors, then proximity sensors.

The maximum range for through-beam sensors is of primary importance. At any distance less than the maximum range, the sensor has more than enough power to detect an object.

The optimum range for the proximity and reflex sensors is more significant than the maximum range. The optimum range is the range at which the sensor has the most power available to detect objects. The optimum range is best shown by an excess gain chart (Fig. 1.30).

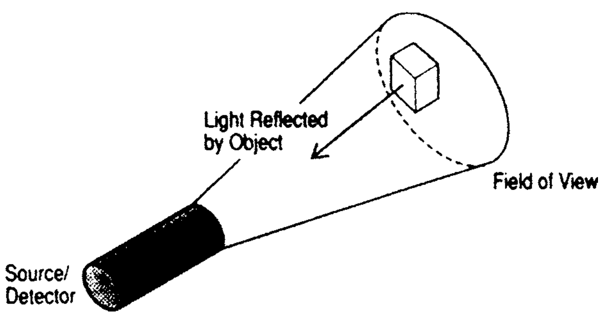


FIGURE 1.29 Proximity detection.

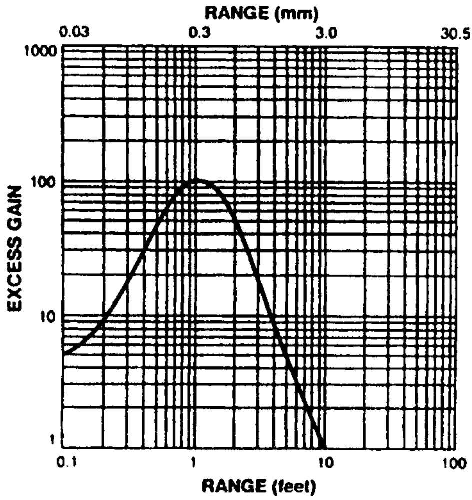


FIGURE 1.30 Photoelectric excess gain and range

Excess gain is a measure of sensing power available in excess of that required to detect an object. An excess gain of 1 means there is just enough power to detect an object, under the best conditions and without obstacles placed in the light beam. The distance at which the excess gain equals 1 is the maximum range. An excess gain of 100 means there is 100 times the power required to detect an object. Generally, the more excess gain available at the required range, the more consistently the control will operate.

For each distance within the range of sensor, there is a specific excess gain. Through-beam controls generally provide the most excess gain, followed by reflex, and then proximity sensors.

General guidelines can be provided for the amount of excess gain required for the level of contamination in an environment. Environments are classified as relatively clean, lightly dirty, dirty, very dirty, and extremely dirty. Table 1.2 illustrates the excess gain recommended for these types of environments for each sensing mode.

TABLE 1.2 Excess Gain Chart

Environment	Through-beam	Reflex	Proximity
Relatively clean	1.25 per side	1.6 per side	
Office clean	1.6 total	2.6 total	2.6 total
Lightly dirty	1.8 per side	3.2 per side	
Warehouse, post office	3.2 total	10.5 total	3.2 total
Dirty	8 per side	64 per side	
Steel mill, saw mill	64 total		64 total
Very dirty	25 per side		
Steam tunnel, painting, rubber or grinding, cutting with coolant, paper plant	626 total		
Extremely dirty	100 per side		
Coal bins or areas where thick layers build quickly	10,000 total		

Example. If, in a through-beam setup, the source is in a lightly dirty environment where excess gain is 1.8, and the detector is in a very dirty environment where excess gain is 25, the recommended excess gain is $1.8 \times 25 = 45$, from Table 1.2.

PROXIMITY SENSORS

Proximity sensing is the technique of detecting the presence or absence of an object with an electronic noncontact sensor.

Mechanical limit switches were the first devices to detect objects in industrial applications. A mechanical arm touching the target object moves a plunger or rotates a shaft, which causes an electrical contact to close or open. Subsequent signals will produce other control functions through the connecting system. The switch may be activating a simple control relay, or a sophisticated programmable logic control device, or a direct interface to a computer network. This simple activity, once done successfully, will enable varieties of manufacturing operations to direct a combination of production plans according to the computer-integrated manufacturing strategy.

Inductive proximity sensors are used in place of limit switches for noncontact sensing of metallic objects. Capacitive proximity switches are used on the same basis as inductive proximity sensors; however, capacitive sensors can also detect nonmetallic objects. Both inductive and capacitive sensors are limit switches with ranges up to 100 mm.

The distinct advantage of photoelectric sensors over inductive or capacitive sensors is their increased range. However, dirt, oil mist, and other environmental factors will hinder operation of photoelectric sensors during the vital operation of reporting the status of a manufacturing process. This may lead to significant waste and buildup of false data.

Typical Applications of Inductive Proximity Sensors

Motion position detection (Fig. 1.31)

- Detection of rotating motion
- Zero-speed indication
- Speed regulation

Motion control (Fig. 1.32)

- Shaft travel limiting
- Movement indication
- Valve open/closed

Conveyer system control (Fig. 1.33)

- Transfer lines
- Assembly line control
- Packaging machine control

Process control (Fig. 1.34)

- Product complete
- Automatic filling
- Product selection

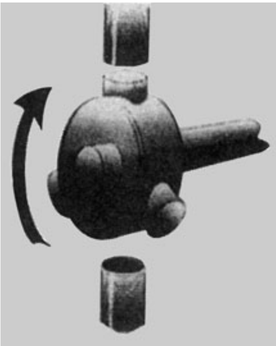


FIGURE 1.31 Motion/position detection with inductive proximity sensor.

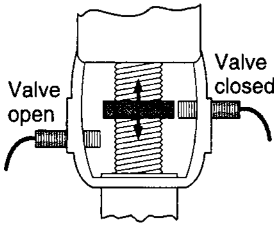


FIGURE 1.32 Motion control, inductive proximity sensor.

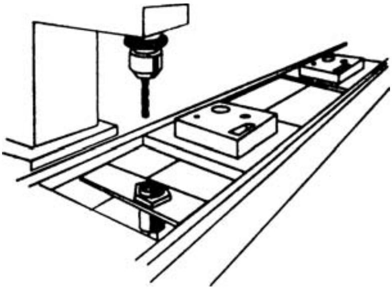


FIGURE 1.33 Conveyer system control, inductive proximity sensor.

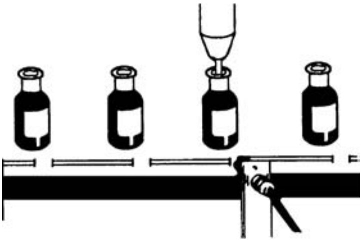


FIGURE 1.34 Process control, inductive proximity sensor.

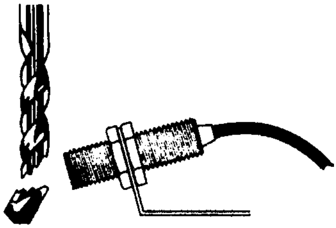


FIGURE 1.35 Machine control, inductive proximity sensor.

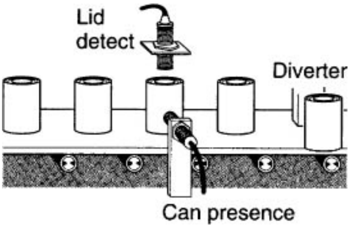


FIGURE 1.36 Verification and counting, inductive proximity sensor.

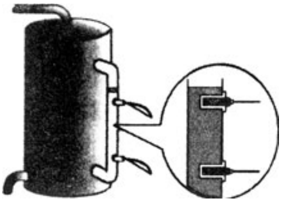


FIGURE 1.37 Liquid level detection, capacitive proximity sensor.

Machine control (Fig. 1.35)

- Fault condition indication
- Broken tool indication
- Sequence control

Verification and counting (Fig. 1.36)

- Product selection
- Return loop control
- Product count

Typical Applications of Capacitive Proximity Sensors

Liquid level detection (Fig. 1.37)

- Tube high/low liquid level
- Overflow limit
- Dry tank

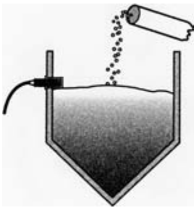


FIGURE 1.38 Bulk material level control, capacitive proximity sensor.

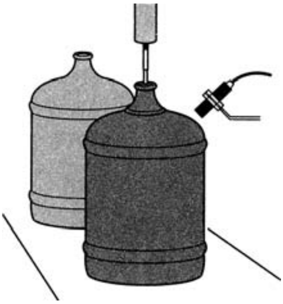


FIGURE 1.39 Process control, capacitive proximity sensor.

Bulk material level control (Fig. 1.38)

- Low level limit
- Overflow limit
- Material present

Process control (Fig. 1.39)

- Product present
- Bottle fill level
- Product count

UNDERSTANDING INDUCTIVE PROXIMITY SENSORS

Principles of Operation

An inductive proximity sensor consists of four basic elements (Fig. 1.40).

- Sensor coil and ferrite core
- Oscillator circuit
- Detector circuit
- Solid-state output circuit

The oscillator circuit generates a radio-frequency electromagnetic field that radiates from the ferrite core and coil assembly. The field is centered around the axis of the ferrite core, which shapes the field and directs it at the sensor face. When a metal target approaches and enters the field, *eddy currents* are induced into the surfaces of the target. This results in a loading effect, or "damping," that causes a reduction in amplitude of the oscillator signal (Fig. 1.41).

The detector circuit detects the change in oscillator amplitude (Fig. 1.42). The detector circuit will switch ON at a specific operate amplitude. This ON signal generates a signal to turn ON the solid-state output. This is often referred to as the *damped* condition. As the target leaves the sensing field, the oscillator responds with an increase in amplitude. As the amplitude increases above a specific value, it is detected by the detector circuit, which

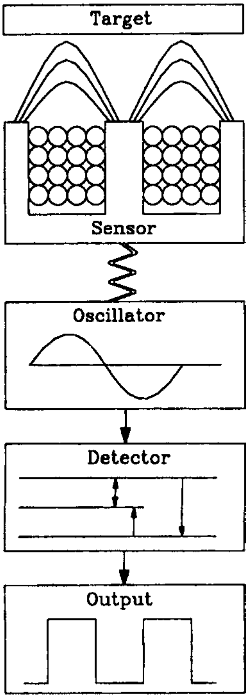


FIGURE 1.40 Operating principle of inductive proximity sensor.

switches OFF, causing the output signal to return the normal or OFF (*undamped*) state.

The difference between the operate and the release amplitude in the oscillator and corresponding detector circuit is referred to as the *hysteresis* (H) of the sensor. It corresponds to a difference in the point of target detection and release distance between the sensor face and the target surface (Fig. 1.43).

Inductive Proximity Sensing Range

The *sensing range* of an inductive proximity sensor refers to the distance between the sensor face and the target. It also includes the shape of the sensing field generated through the coil and core. Several mechanical and environmental factors can affect the sensing range:

Mechanical factors	Environmental factors
Core size	Ambient temperature
Core shield	Surrounding electrical conditions
Target material	Surrounding mechanical conditions
Target size	Variation between devices
Target shape	

The geometry of the sensing field can be determined by the construction factor of the core and coil. An open coil with no core produces an omnidirectional field. The geometry of an air-core is a toroid. Such sensors could be actuated by a target approaching from any direction, making them undesirable for practical industrial applications (Fig. 1.44).

Ferrite material in the shape of a cup core is used to shape the sensing field. The ferrite material absorbs the magnetic field, but enhances the field intensity and directs the field out of the open end of the core (Fig. 1.45).

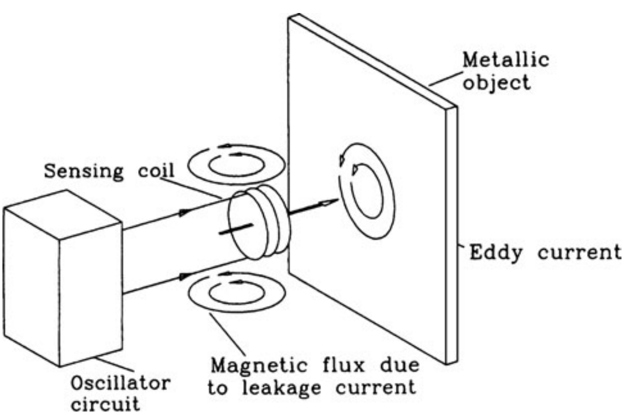


FIGURE 1.41 Induced eddy current.

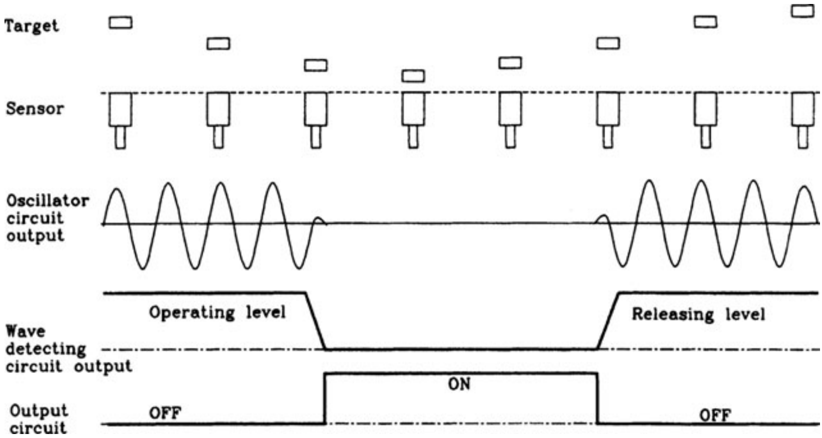


FIGURE 1.42 Detection cycle.

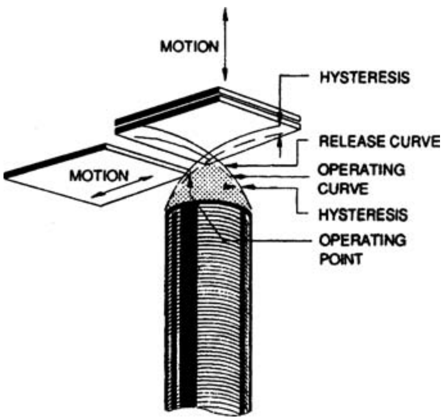


FIGURE 1.43 Core assembly.

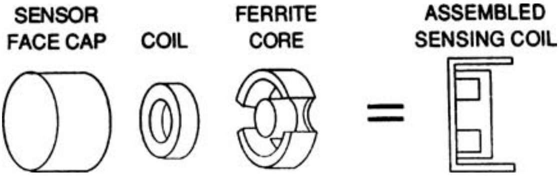


FIGURE 1.44 Open coil without core.

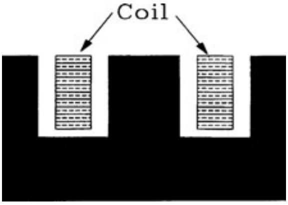


FIGURE 1.45 Cup-shaped coil/core assembly.

A standard field range sensor is illustrated in Fig. 1.46. It is often referred to as *shielded* sensing coil. The ferrite contains the field so that it emanates straight from the sensing face. Figure 1.47 shows the typical standard-range sensing-field plot.

An extended range coil and core assembly does not have the ferrite around the perimeter of the coil (Fig. 1.48). This unshielded device accordingly has an extended range. Figure 1.49 illustrates a typical extended sensing-field plot.

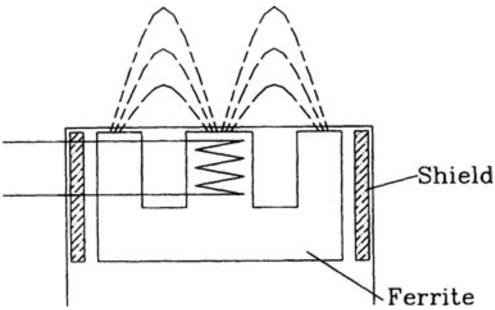


FIGURE 1.46 Standard range core coil.

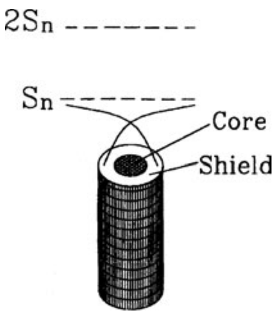


FIGURE 1.47 Standard range field plot.

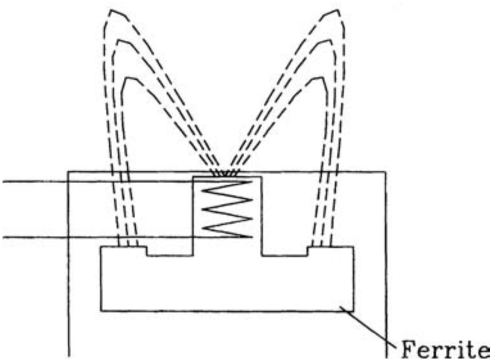


FIGURE 1.48 Extended range core and coil.

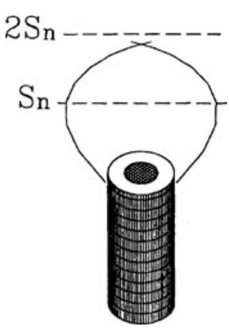


FIGURE 1.49 Extended range field plot.

Sensing Distance

The electromagnetic field emanates from the coil and core at the face of the sensor and is centered around the axis of the core. The nominal sensing range is a function of the coil diameter and the power that is available to operate the electromagnetic field.

The sensing range is subject to manufacturing tolerances and circuit variations. Typically it varies by 10 percent. Similarly, temperature drift can affect the sensing range by 10 percent. Applied to the nominal sensing switch, these variations mean that the sensing range can be as much as 120 percent or as little as 81 percent of the nominal stated range (Fig. 1.50).

$$S_r = 0.9 < S_n < 1.1$$
$$S_n = 0.9 < S_r < 1.1 \qquad (1.1)$$

where S_n = nominal sensing range
S_r = effective sensing range
S_u = usable sensing range

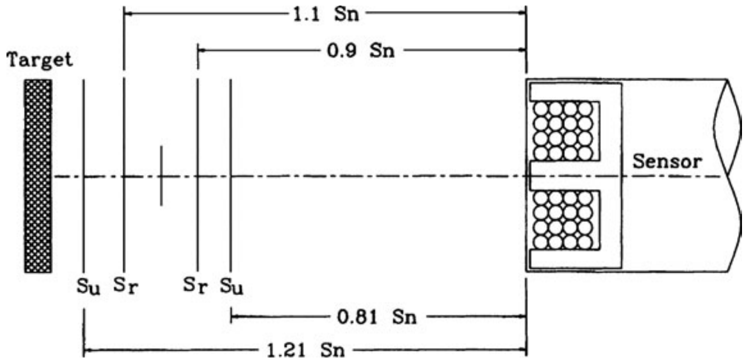


FIGURE 1.50 Sensing distance tolerances.

Target Material and Size

As a target approaches the sensing field, eddy currents are induced in the target. In order to ensure that the target has the desired damping affect on the sensor, the target must be of appropriate size and material. Metallic targets can be defined as:

- *Ferrous.* Containing iron, nickel, or cobalt.
- *Nonferrous.* All other metallic materials, such as aluminum, copper, and brass. Eddy currents induced in ferrous targets are stronger than in nonferrous targets, as illustrated in Table 1.3.

An increase in target size will not produce an increase in sensing range. However, a decrease in target size will produce a decrease in sensing range, and may also increase response time. Figure 1.51 illustrates the relationship of target size and target material to the sensing range of a limit-switch-type sensor with a nominal 13-mm sensing range.

Table 1.4, shows correction factors by which the rated nominal sensing range of most inductive proximity sensors can be multiplied. This will determine the effective sensing range for a device sensing a stated target material of standard target size.

Example.	18-mm tubular extended range	8 mm
	Copper target correction factor	× 0.35
	Sensing range detecting copper standard target	2.80 mm

TABLE 1.3 Standard Target

Device	Standard target dimensions	Standard target material
Modular limit switch type	45 mm square × 1 mm thick	Mild steel
8 mm tubular	8 mm square × 1 mm thick	Mild steel
12 mm tubular	12 mm square × 1 mm thick	Mild steel
18 mm tubular	18 mm square × 1 mm thick	Mild steel
30 mm tubular	30 mm square × 1 mm thick	Mild steel

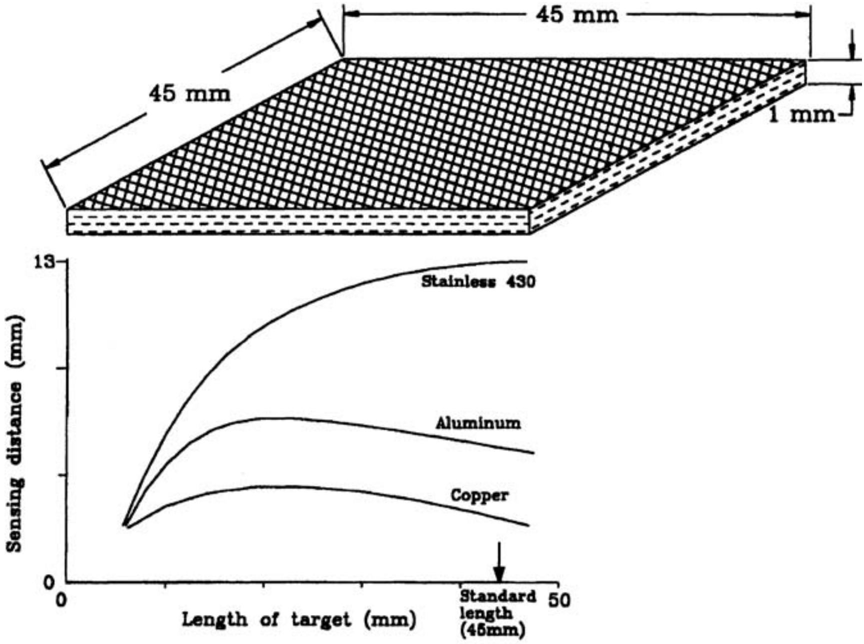


FIGURE 1.51 Sensing range correction factor.

Target Shape

Standard targets are assumed to be of a flat square shape with the stated dimensions. Targets of round shape or with a pocketed surface have to be of adequate dimensions to cause the necessary dampening effect on the sensor. Allowing the sensor-to-target distance less than the nominal range will help assure proper sensing. In addition, using the next largest size or an extended-range sensor will also minimize problems with other-than-standard target dimensions or shapes. Figure 1.52 illustrates the axial (head-on) approach, indicating that the target approaches the face of the sensor on the axis of the coil core. When the target approaches axially, the sensor should not be located such that it becomes an end stop. If axial operation is considered, good application practice is to allow for 25 percent overtravel.

TABLE 1.4 Target Material Correction

Target material	Limit-switch type	Pancake type	Tubular, mm			
			8	12	18	30
Steel (1020)	1.0	1.0	1.0	1.0	1.0	1.0
Stainless steel (400)	1.03	0.90	0.90	0.90	1.0	1.0
Stainless steel (300)	0.85	0.70	0.60	0.70	0.70	0.65
Brass	0.50	0.54	0.35	0.45	0.45	0.45
Aluminum	0.47	0.50	0.35	0.40	0.45	0.40
Copper	0.40	0.46	0.30	0.25	0.35	0.30

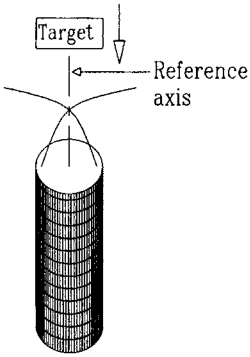


FIGURE 1.52 Axial approach.

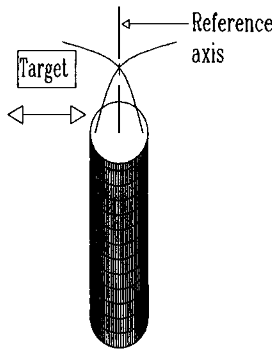


FIGURE 1.53 Lateral approach.

Lateral (*side-by*) *approach* means the target approaches the face of the sensor perpendicular to the axis of the coil core (Fig. 1.53). Good application practice (GAP), a terminology often used in "world-class" manufacturing strategies, dictates that the tip of the sensing field envelope should not be used. That is the point where sensing range variations start to occur. Therefore, it is recommended that the target pass not more than 75 percent of the sensing distance *D* from the sensor face. Also, the target should not pass any closer than the basic tolerance incorporated in the machine design, to prevent damage to the sensor. Hysteresis can be greater for an axial approach (Fig. 1.54).

Variation Between Devices

Variations of sensing range between sensors of the same type often occur. With modern manufacturing technologies and techniques, variations are held to a minimum. The variations can be attributed to collective tolerance variations of the electrical components in the sensor circuit and to subtle differences in the manufacturing process from one device to the next; 5 percent variation is typical (Fig. 1.55).

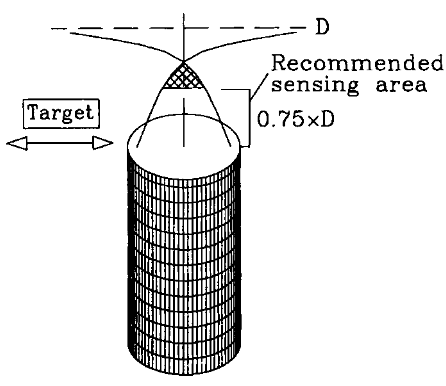


FIGURE 1.54 Lateral approach—recommended sensing distance.

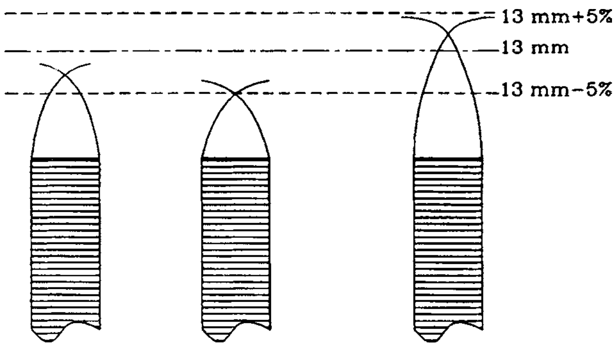


FIGURE 1.55 Sensing range variation.

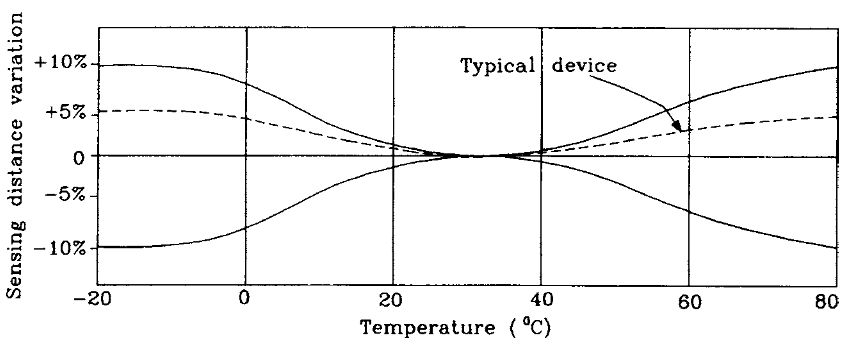


FIGURE 1.56 Sensing range variation—with temperature.

The sensing distance also will vary from one temperature extreme to the other because of the effect of temperature change on the components of the sensor. Typical temperature ranges are −25°C (−3°F) to +70°C (+180°F). Figure 1.56 illustrates sensing range variation with temperature.

Surrounding Conditions

Several environmental factors must also be considered in order to obtain reliable information from inductive proximity sensors. These surrounding factors are:

- *Embeddable mounting.* The shielded sensor in Fig. 1.57 is often referred to as a *flush-mounted* sensor. Shielded sensors are not affected by the surrounding metal.
- *Flying metal chips.* A chip removed from metal during milling and drilling operations may affect the sensor performance depending on the size of the chip, its location on the sensing face, and type of material. In these applications, the sensor face should be oriented so that gravity will prevent chips from accumulating on the sensor face. If this is not possible, then coolant fluid should wash the sensor face to remove the chips. Generally, a chip does not have sufficient surface area to cause a sensor turn on. If a chip lands on the center of the sensor face, it will have a negligible effect, but elsewhere on the sensor face it will extend the range of the sensor.

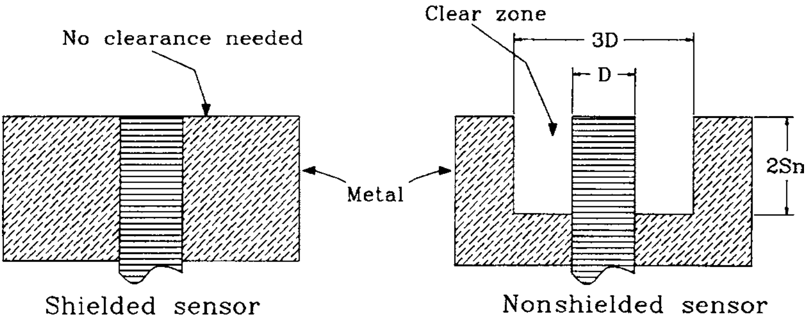


FIGURE 1.57 Embeddable and nonembeddable sensors.

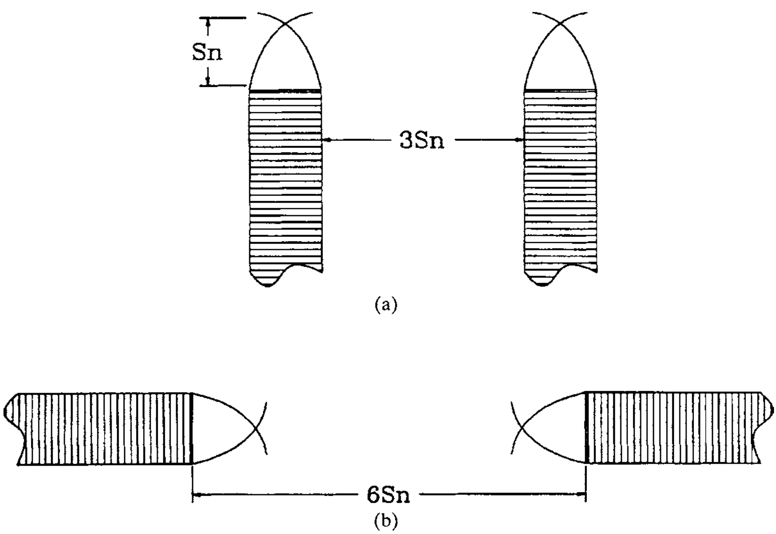


FIGURE 1.58 Adjacent sensors.

- *Adjacent sensors.* When two similar sensors are located adjacent to or opposite each other, the interaction of their fields can affect operation. Figure 1.58 provides the guidelines for placing two similar sensors adjacent to each other. *Alternate-frequency* heads will allow adjacent mounting of sensors without interaction of their sensing fields.
- *Magnetic fields.* Electrical wiring in the vicinity of the sensor face may affect sensor operation. If the magnetic field around the electrical wiring reaches an intensity that would saturate the sensor ferrite or coil, the sensor will not function properly. Use of inductive sensors in the presence of high-frequency radiation can also unexpectedly affect their operation. Sensors specially designed for welding application can be used with programmable logic control (PLC). The PLC can be programmed to ignore the signal from the sensor for the period that the high-frequency welder is operated. A slight OFF-time delay assures proper operation of the sensor.
- *Radio-frequency interference* (RFI). Radio transceivers, often called *walkie-talkie* devices, can produce a signal that can cause an inductive proximity sensor to operate falsely. The

radio transceiver produces a radio-frequency signal similar to the signal produced by the oscillator circuit of the sensor. The effect that RFI has on an inductive proximity switch can vary. The factors that determine this variation are as follows:

- *Distance between the RFI source and the sensor.* Typically, inductive proximity switches are not affected by RFI when a transceiver is 1 ft away from the inductive switch. However, if closer than 1 ft, the switch may operate without a target present.
- *Signal frequency.* The signal frequency may be the determining factor that will cause a particular device to false-operate.
- *Signal intensity.* Radio-frequency transceivers usually are portable devices with a power rating of 5 W maximum.
- *Inductive proximity package.* The sensor package construction may determine how well the device resists RFI.
- *Approach to the sensor.* A transceiver approaching the connecting cable of a switch may affect it at a greater distance than if it was brought closer to the sensing face. As RFI protection varies from device to device and manufacturer to manufacturer, most manufacturers have taken steps to provide the maximum protection against false operation due to RFI.

- *Showering arc.* Showering arc is the term applied to induced line current/voltage spikes. The spike is produced by the electrical arc on an electromechanical switch or contactor closure. The current spike is induced from lines connected to the electromechanical switch to the lines connected to the inductive proximity switch if the lines are adjacent and parallel to one another. The result can be false operation of the inductive proximity switch. The spike intensity is determined by the level of induced voltage and the duration of the spike. Avoiding running cables for control devices in the same wiring channel as those for the contactor or similar leads may eliminate spikes. Most electrical code specifications require separation of control device leads from electromechanical switch and contractor leads.

UNDERSTANDING CAPACITIVE PROXIMITY SENSORS

Principles of Operation

A capacitive proximity sensor operates much like an inductive proximity sensor. However, the means of sensing is considerably different. Capacitive sensing is based on dielectric capacitance. *Capacitance* is the property of insulators to store an electric charge. A capacitor consists of two plates separated by an insulator, usually called a *dielectric.* When the switch is closed (Fig. 1.59), a charge is stored on the two plates.

The distance between the plates determines the ability of a capacitor to store a charge and can be calibrated as a function of stored charge to determine discrete ON and OFF switching status.

Figure 1.60 illustrates the principle as it applies to the capacitive sensor. One capacitive plate is part of the switch, the sensor face (the enclosure) is the insulator, and the target is the other plate. Ground is the common path.

The capacitive proximity sensor has the same four basic elements as an inductive proximity sensor:

- Sensor (the dielectric plate)
- Oscillator circuit
- Detector circuit
- Solid-state output circuit

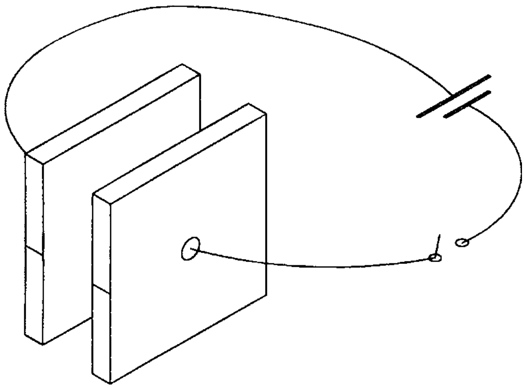


FIGURE 1.59 Capacitive principle.

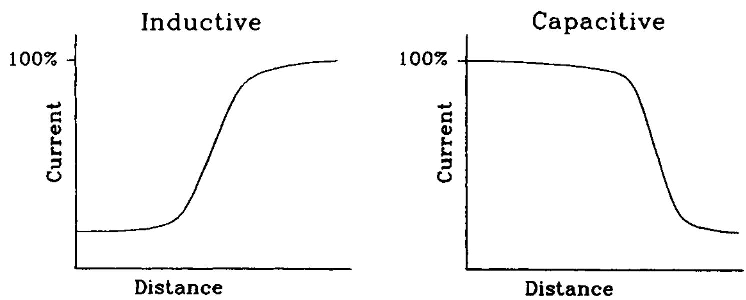


FIGURE 1.60 Capacitive sensor.

The oscillator circuit in a capacitive switch operates like one in an inductive proximity switch. The oscillator circuit includes feedback capacitance from the external target plate and the internal plate. In a capacitive switch, the oscillator starts oscillating when sufficient feedback capacitance is detected. In an inductive proximity switch, the oscillation is damped when the target is present (Fig. 1.61).

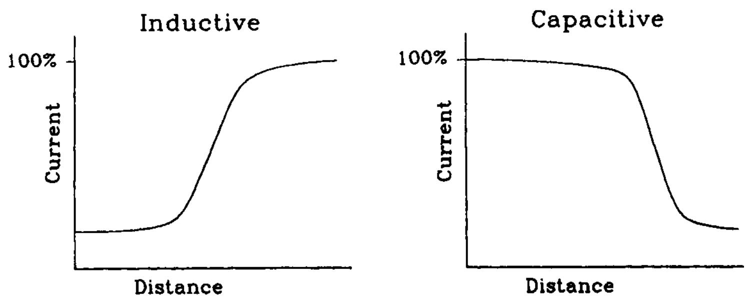


FIGURE 1.61 Oscillator damping of inductive and capacitive sensors.

In both capacitive and inductive switch types, the difference between the operate and release amplitude in the oscillator and corresponding detector circuit is referred to as the *hysteresis* of the sensor. It corresponds to the difference between target detection and release distances from the sensor face.

Features of Capacitive Sensors

The major characteristics of capacitive proximity sensors are:

- They can detect nonmetallic targets.
- They can detect lightweight or small objects that cannot be detected by mechanical limit switches.
- They provide a high switching rate for rapid response in object counting applications.
- They can detect liquid targets through nonmetallic barriers, (glass, plastic, and so on).
- They have a long operational life with a virtually unlimited number of operating cycles.
- The solid-state output provides a bounce-free contact signal.

Capacitive proximity sensors have two major limitations:

- They are affected by moisture and humidity.
- They must have an extended range for effective sensing.

Sensing Range

Capacitive proximity sensors have a greater sensing range than inductive proximity sensors, as illustrated in the following.

Tubular diameter, mm	Inductive extended range, mm	Capacitive extended range, mm
18	8	10
30	15	20
34	—	40

The sensing distance for capacitive proximity sensors is a matter of plate diameter, just as coil size is for inductive proximity sensors. Capacitive sensors basically measure a dielectric gap. Accordingly, it is desirable to be able to compensate for target and application conditions with a sensitivity adjustment for the sensing range. Most capacitive proximity switches are equipped with a sensitivity adjustment potentiometer (Fig. 1.62).

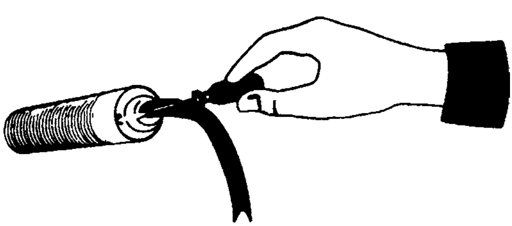


FIGURE 1.62 Sensitivity adjustment.

TABLE 1.5 Target Material Correction

Material	Factor
Mild steel	1.0
Cast iron	1.0
Aluminum and copper	1.0
Stainless steel	1.0
Brass	1.0
Water	1.0
Polyvinylchloride (PVC)	0.5
Glass	0.5
Ceramic	0.4
Wood	≥0.2
Lubrication oil	0.1

Target Material and Size

The sensing range of capacitive sensors, like that of inductive proximity sensors, is determined by the type of material. Table 1.5 lists the sensing-range derating factors that apply to capacitive proximity sensors. Capacitive sensors can be used to detect a target material through a nonmetallic interposing material like glass or plastic. This is beneficial in detecting a liquid through the wall of a plastic tank or through a glass sight tube. The transparent interposing material has no effect on sensing. For all practical purposes, the target size can be determined in the same manner as for inductive proximity sensors.

Surrounding Conditions

Capacitive proximity devices are affected by component tolerances and temperature variations. As with inductive devices, capacitive proximity devices are affected by the following surrounding conditions:

- *Embeddable mounting.* Capacitive sensors are generally treated as nonshielded, nonembeddable devices.
- *Flying chips.* Capacitive devices are more sensitive to metallic and nonmetallic chips.
- *Adjacent sensors.* Allow more space than inductive proximity devices because of the greater sensing range of capacitive devices.
- *Target background.* Relative humidity may cause a capacitive device to operate even when a target is not present. Also, the greater sensing range and ability to sense nonmetallic target materials dictate greater care in applying capacitive devices with target background conditions.
- *Magnetic fields.* Capacitive devices are not usually applied in welding environments.
- *Radio-frequency interference.* Capacitive sensor circuitry can be affected by RFI in the same way an inductive device can.
- *Showering arc.* An induced electrical noise will affect the circuitry of a capacitive device in the same way it does an inductive device.

UNDERSTANDING LIMIT SWITCHES

A limit switch is constructed much like the ordinary light switch used in home and office. It has the same ON/OFF characteristics. The limit switch usually has a pressure-sensitive mechanical arm. When an object applies pressure on the mechanical arm, the switch circuit is energized. An object might have a magnet attached that causes a contact to rise and close when the object passes over the arm.

Limit switches can be either *normally open* (NO) or *normally closed* (NC) and may have multiple poles (Fig. 1.63). A normally open switch has continuity when pressure is applied and a contact is made, while a normally closed switch opens when pressure is applied and a contact is separated. A single-pole switch allows one circuit to be opened or closed upon switch contact, whereas a multiple-pole switch allows multiple circuits to be opened or closed.

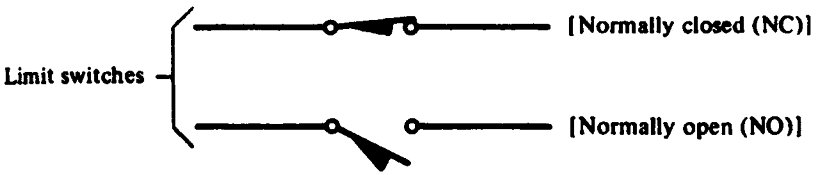


FIGURE 1.63 Normally open/normally closed microswitches.

Limit switches are mechanical devices. They have three potential problems:

- They are subject to mechanical failure.
- Their mean time between failures (MTBF) is low compared to noncontact sensors.
- Their speed of operation is relatively slow; the switching speed of photoelectric microsensors is up to 3000 times faster.

INDUCTIVE AND CAPACITIVE SENSORS IN MANUFACTURING

Inductive and capacitive proximity sensors interface to control circuits through an output circuit, for manufacturing applications. Also, the control circuit type is a determining factor in choosing an output circuit. Control circuits, whether powered by AC, DC, or AC/DC, can be categorized as either *load powered* or *line powered*.

Load-powered devices are similar to limit switches and are connected in series with the controlled load. These devices have two connection points and are often referred to as *two-wire switches*. Operating current is drawn through the load. When the switch is not operated, the switch must draw a minimum operating current, referred to as *residual current*. When the switch is operated or damped (i.e., a target is present), the current required to keep the sensor operating is the minimum holding current (Fig. 1.64). The residual current is not a consideration for low-impedance loads such as relays and motor starters. However, high-impedance loads, most commonly programmable logic controllers, require a residual current of less than 2 mA. Most sensors offer 1.7 mA or less.

In some manufacturing applications, a particular type of PLC will require less than 1.7 mA of residual current. In such applications, a loading resistor is added in parallel to the input to the PLC load. Then, minimum holding current may range up to 20 mA, depending

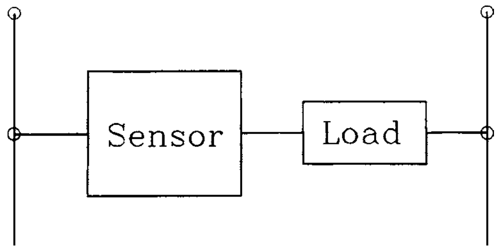


FIGURE 1.64 Load-powered residual current.

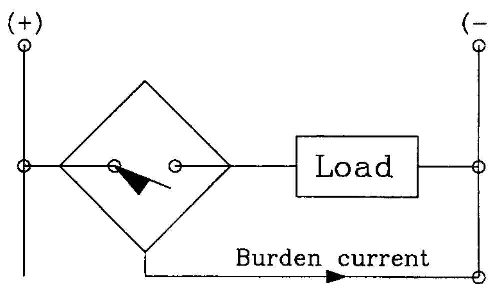


FIGURE 1.65 Line-powered burden current.

on the sensor specification. If the load impedance is too high, there will not be enough load current level to sustain the switch state.

Inductive proximity sensors with a holding current of 4 mA or less can be considered low-holding-current sensors. These devices can be used with PLCs without concern for the minimum holding current.

Line-powered devices derive current, usually called *burden current*, from the line and not through the controller load. These devices are called three-wire switches because they have three connections (Fig. 1.65).

The operating current for a three-wire sensor is burden current, and is typically 20 mA. Since the operating current does not pass through the load, it is not a major concern for the circuit design.

Relays

An output circuit relay is a mechanical switch available in a variety of contact configurations. Relays can handle load currents at high voltages, allowing the sensor to directly interface with motors, large solenoids, and other inductive loads. They can switch either AC or DC loads. Contact life depends on the load current and frequency of operation. Relays are subject to contact wear and resistance buildup. Because of contact bounce, they can produce erratic results with counters and programmable controllers unless the input is filtered. They can add 10 to 25 ms to an inductive or capacitive switch response time because of their mechanical nature (Fig. 1.66).

Relays are familiar to manufacturing personnel. They are often used with inductive or capacitive proximity sensors since they provide multiple contacts. The good and bad features of a relay are summarized next.

Relay advantages	Relay disadvantages
Switches high currents/loads	Slow response time
Multiple contacts	Mechanical wear
Switches AC or DC voltages	Contact bounce
Tolerant of inrush current	Affected by shocks and vibration

Triac Devices

A triac is a solid-state device designed to control AC current (Fig. 1.67). Triac switches turn ON in less than a microsecond when the gate (control leg) is energized, and shut OFF at the zero crossing of the AC power cycle.

Because a triac is a solid-state device, it is not subject to the mechanical limitations of a relay such as mechanical bounce, pitting, corrosion of contacts, and shock and vibration. Switching response time is limited only by the time it takes the 60-Hz AC power to go through one-half cycle (8.33 ms) (Fig. 1.68).

As long as a triac is used within its rated maximum current and voltage specifications, life expectancy is virtually infinite. Triac devices used with inductive or capacitive sensors

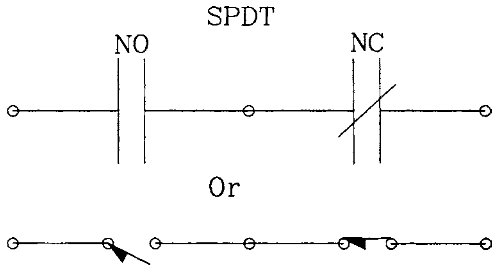


FIGURE 1.66 Relay output.

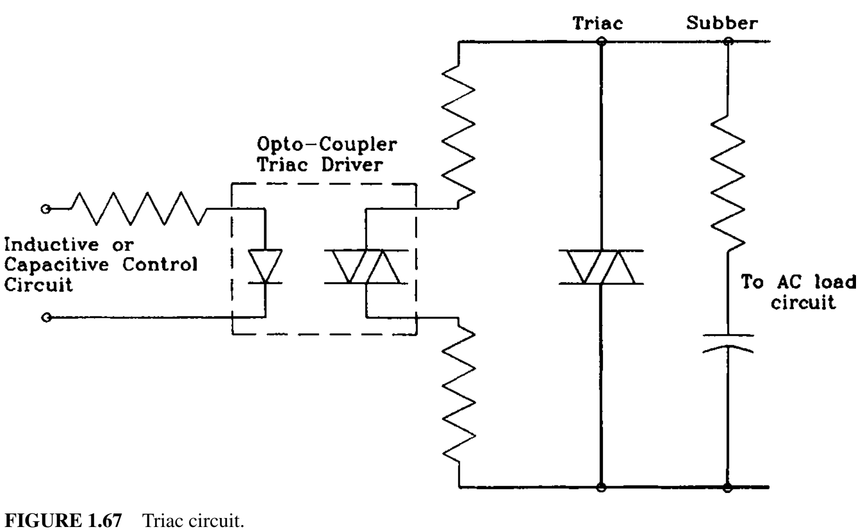


FIGURE 1.67 Triac circuit.

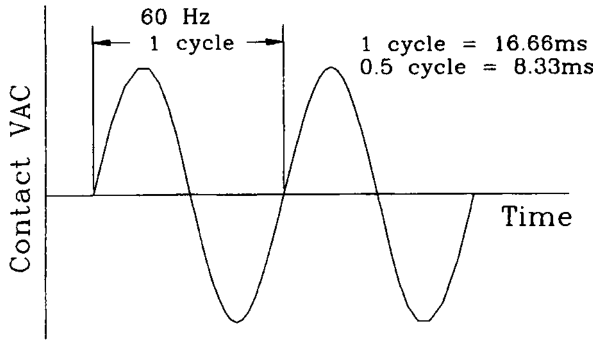


FIGURE 1.68 AC power cycle.

generally are rated at 2-A loads or less. Triac limitations can be summarized as follows: (1) shorting the load will destroy a triac, and (2) directly connected inductive loads or large voltage spikes from other sources can false-trigger a triac.

To reduce the effect of these spikes, a snubber circuit composed of a resistor and capacitor in series is connected across the device. Depending on the maximum switching load, an appropriate snubber network for switch protection is used. The snubber network contributes to the OFF state leakage to the load. The leakage must be considered when loads requiring little current, such as PLCs, are switched. In the ON state, a drop of about 1.7 V rms is common (Fig. 1.69). Good and bad features of triacs are listed next.

Triac advantages	Triac disadvantages
Fast response time (8.33 ms)	Can be falsely triggered by large inductive current
Tolerant of large inrush currents	Snubber contributes to OFF state leakage current
Can be directly interfaced with programmable controllers	Can be destroyed by short circuits
Infinite life when operated within rated voltage/current limits	

Transistor DC Switches

Transistors are solid-state DC switching devices. They are most commonly used with low-voltage DC-powered inductive and capacitive sensors as the output switch. Two types are employed, depending on the function (Fig. 1.70).

In an NPN transistor, the current source provides a contact closure to the DC positive rail. The NPN current sink provides a contact to the DC common. The transistor can be thought of as a single-pole switch that must be operated within its voltage and maximum current ratings (Fig. 1.71).

Any short circuit on the load will immediately destroy a transistor that is not short-circuit protected. Switching inductive loads creates voltage spikes that exceed many times the maximum rating of the transistor. Peak voltage clamps such as zener diodes or transorbs are utilized to protect the output device. Transistor outputs are typically rated to switch loads of 250 mA at 30 V DC maximum (Fig. 1.72).

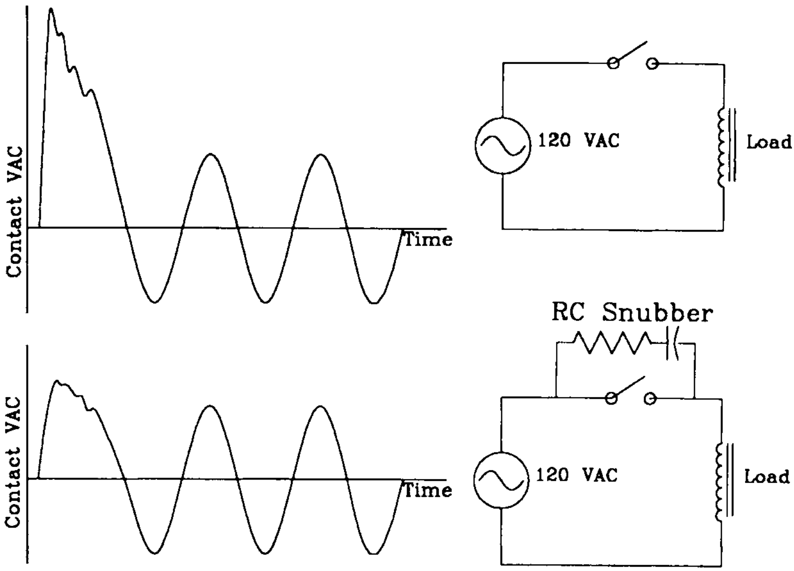


FIGURE 1.69 Snubber circuit.

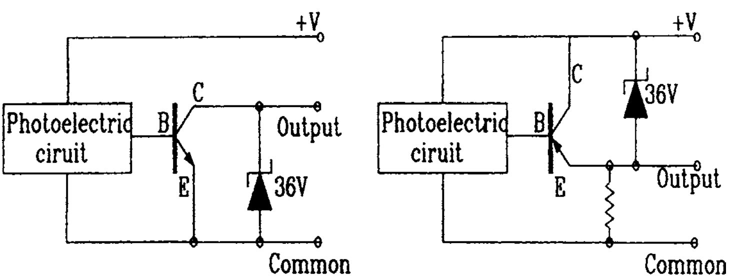


FIGURE 1.70 DC circuit logic.

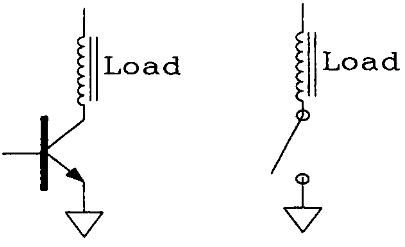


FIGURE 1.71 Transistor switch.

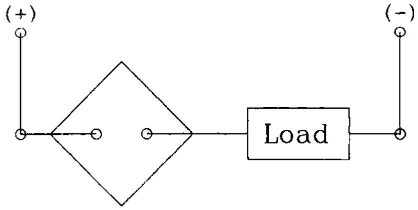


FIGURE 1.72 Voltage clamp.

Advantage and disadvantage of transistors use in switching circuits are shown below.

Transistor advantages	Transistor disadvantages
Virtually instantaneous response	Low current handling capacity
Low OFF state leakage and voltage drop	Cannot handle inrush current unless clamped
Infinite life when operated within rated current/voltage	Can be destroyed by short circuit unless protected
Not affected by shock and vibration	

Output Configuration. Output configurations are categorized as follows:

- Single output—normally open (NO)
- Single output—normally closed (NC)
- Programmable output—NO or NC
- Complementary output—NO and NC

The functions of normally open and normally closed output are defined in Table 1.6.

Inductive and Capacitive Control/Output Circuits

A single output sensor has either an NO or an NC configuration and cannot be changed to the other configuration (Fig. 1.73).

A programmable output sensor has one output, NO or NC, depending on how the output is wired when installed. These sensors are exclusively two-wire AC or DC (Fig. 1.74).

A complementary output sensor has two outputs, one NO and one NC. Both outputs change state simultaneously when the target enters or leaves the sensing field. These sensors are exclusively three-wire AC or DC (Fig. 1.75).

The choice of control circuit and output logic plays an important part in determining the reliability of data collection. The choice of control circuit and output logic depends on the following parameters:

- *AC or DC control voltage.* Use of AC control may seem to require the use of an AC-configured sensor. However, interface circuitry can allow for DC sensors even if the main control voltage source is AC.
- *Control circuit current requirements.* Usually control circuits operating in the 200- to 300-mA range can use either AC or DC sensors. Circuits with 0.5-A and higher current will dictate the type of sensor to be used.

TABLE 1.6 Output Logic

Output configuration	Target state	Oscillator state	Output
NO	Absent	Undamped	Nonconducting (OFF)
	Present	Damped	Conducting (ON)
NC	Absent	Undamped	Conducting (ON)
	Present	Damped	Nonconducting (OFF)

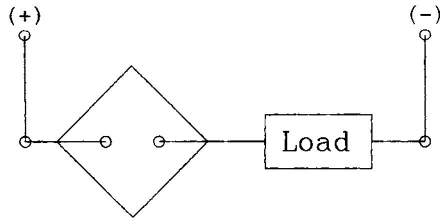


FIGURE 1.73 Single output.

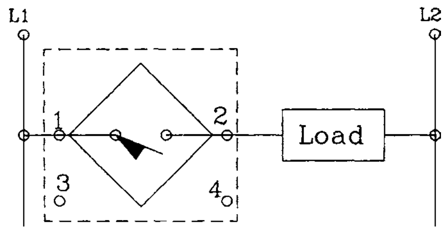


FIGURE 1.74 Programmable output.

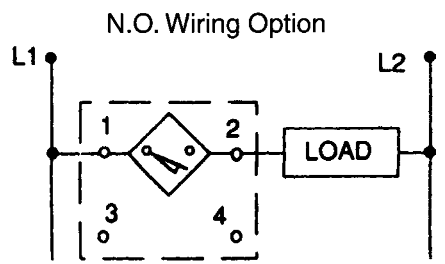


FIGURE 1.75 Complementary output.

- *Application output requirements.* NO output is the most commonly used output type. Controlled circuit configurations may dictate use of NC or complementary-type configured sensors.
- *Switching speed requirements.* AC circuits are limited in their operations per second. DC circuits may be required for applications involving counting or high speed.
- *Connecting logic device.* The device to which the sensor is connected—such as a programmable controller, relay, solenoid, or timer/counter—is usually the most important factor in sensor circuit and output configuration.

Accessories for Sensor Circuits

Sensor circuits and their output configurations must have various types of indicators and protection devices, such as:

- Light-emitting diode (LED) indicators
- Short-circuit protectors

- Reverse-polarity protectors—DC three-wire
- Wire terminators—color-coded wire
- Pin connector type and pin-out designator

LED Indicators. LED indicators provide diagnostic information on the status of sensors (e.g., operated or not operated), that is vital in computer-integrated manufacturing. Two LEDs also indicate the status of complementary-type sensor switches, power ON/OFF status, and short-circuit condition.

Short-Circuit Protection. Short-circuit protection is intended to protect the switch circuit from excessive current caused by wiring short circuits, line power spikes from high inrush sources, or lightning strikes. This option involves special circuitry that either limits the current through the output device or turns the switch OFF. The turn-off–type switch remains inoperative until the short circuit has been cleared—with power disconnected. Then, power is reapplied to the sensor. A second LED is usually furnished with this type of device to indicate the shorted condition.

Reverse-Polarity Protection. Reverse-polarity protection is special circuitry that prevents damage in a three-wire DC sourcing (PNP) or sinking (NPN) device when it is connected to control circuitry incorrectly. Although reverse polarity is relatively common, not all switches are equipped with this option.

Wire Termination. Wire terminals are common on limit-switch enclosure-type sensors. The terminal designations are numbered and correspond to the device wiring diagram (Fig. 1.76). Cable/wire stub terminations are most common on tubular sensors. Color-coded conductors are essential for correct wiring. Most sensor wires are color-coded to comply with industry wire color-code standards.

Pin Connectors. Pin-connector-terminal sensors feature a male pin connector receptacle on the switch or at the end of the wire/cable stub. The female receptacle is at the end of the matching cable cord. Most industry-standard pin connectors are either the mini type—approximately 18 mm in diameter—or the micro type—approximately 12 mm in diameter (Fig. 1.77).

Inductive and Capacitive Switching Logic

The outputs of two or more inductive or capacitive proximity sensors can be wired together in series or parallel to perform logic functions. The ON, or activated, switch function can

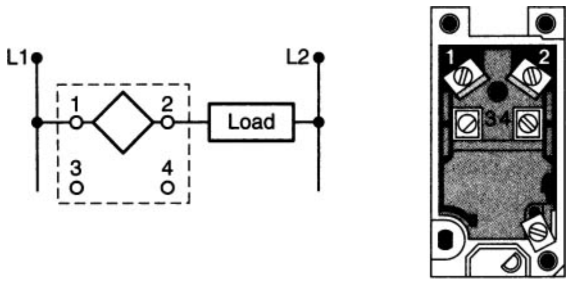


FIGURE 1.76 Wire terminal.

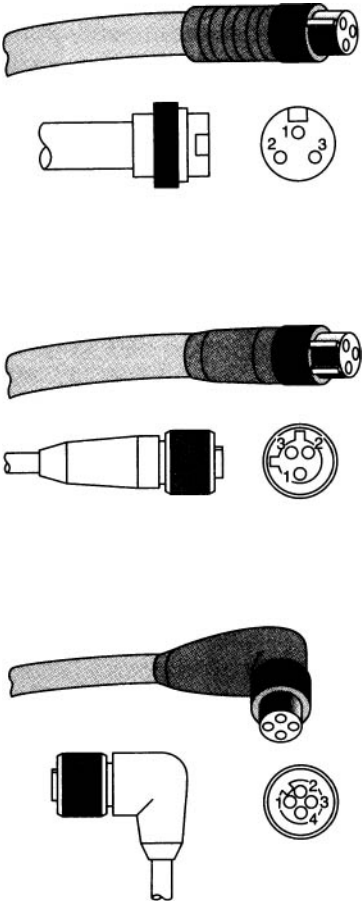


FIGURE 1.77 Pin connector.

be either a normally open or a normally closed output function, depending on the desired control logic.

Although sensors are the most effective means for the data acquisition role in manufacturing, care must be exercised when sensors are integrated with various production operations. The following factors will affect the performance of the switch logic circuit:

- Excessive leakage current in parallel-connected load-powered devices
- Excessive voltage drop in series-connected devices
- Inductive feedback with line-powered sensors with parallel connections

Parallel-Connection Logic—OR Function. The binary OR logic in Table 1.7 indicates that the circuit output is ON (1) if one or more of the sensors in parallel connection is ON.

It is important to note that, in two-wire devices, the OFF state residual current is additive (Fig. 1.78). If the circuit is affected by the total leakage applied, a shunt (loading) resistor may have to be applied. This is a problem in switching to a programmable controller or other high-impedance device.

Example. $I_a + I_b + I_c = I_t$

$$1.7 + 1.7 + 1.7 = 5.1 \text{ mA}$$

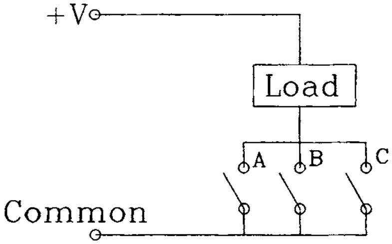
Three-wire 10 to 30 V can also be connected in parallel for a logic OR circuit configuration. Figure 1.79 shows a current *sourcing* (PNP) parallel connection.

Figure 1.80 shows a current *sinking* (NPN) parallel connection. It may be necessary to utilize blocking diodes to prevent inductive feedback (or reverse polarity) when one of the sensors in parallel is damped while the other is undamped. Figure 1.81 demonstrates the use of blocking diodes in this type of parallel connection.

TABLE 1.7 Binary Logic Chart—Parallel Connection

A	B	C	OUT
0	0	0	0
0	0	1	1
0	1	0	1
1	0	0	1

0 = OFF, 1 = ON



C	B	A	Out
0	0	0	0
0	0	1	1
0	1	0	1
1	0	0	1

FIGURE 1.78 Parallel sensor arrangement.

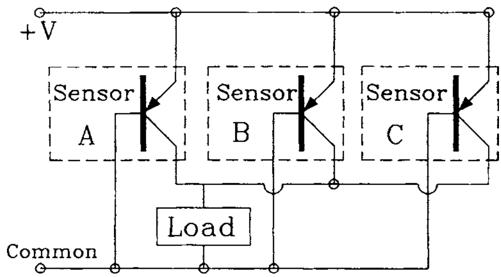


FIGURE 1.79 Sourcing (PNP) parallel sensor arrangement.

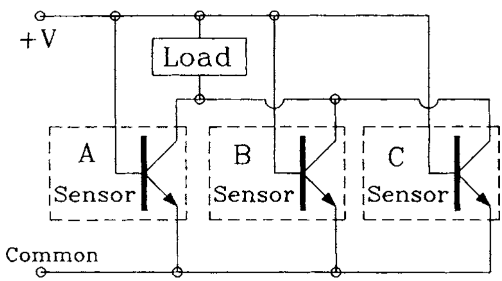


FIGURE 1.80 Sinking (NPN) parallel sensor arrangement.

Series-Connection Logic—AND Function. Figure 1.82 shows AND function logic indicating that the series-connected devices must be ON (1) in order for the series-connected circuit to be ON.

The voltage drop across each device in series will reduce the available voltage the load will receive. Sensors, as a general rule, have a 7- to 9-V drop per device. The minimum operating voltage of the circuit and the sum of the voltage drop per sensor will determine the number of devices in a series-connected circuit. Figure 1.83 shows a typical two-wire AC series-connected circuit.

Series connection is generally applied to two-wire devices, most commonly two-wire AC. 10- to 30-V DC two-wire connections are not usually practical for series connection because of the voltage drop per device and minimum operating voltage. Three-wire devices

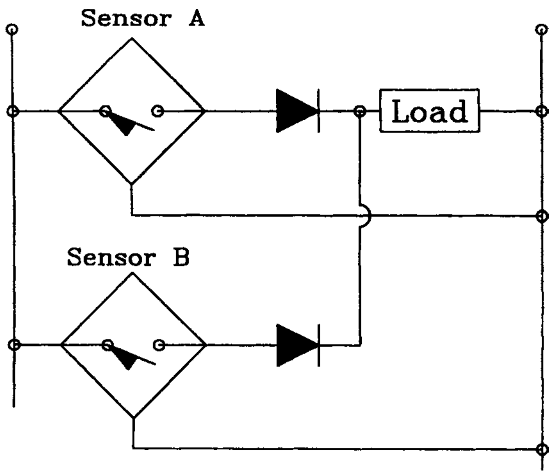


FIGURE 1.81 Blocking diodes.

B	A	Out
0	0	0
0	1	0
1	0	0
1	1	1

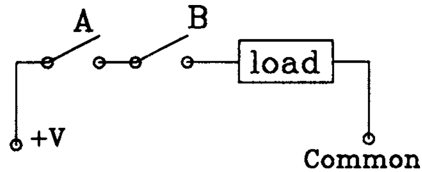


FIGURE 1.82 Series AND logic.

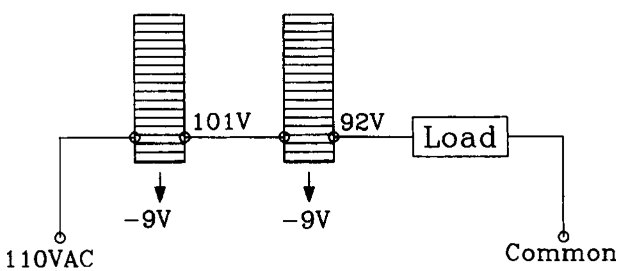


FIGURE 1.83 Series connected, load powered.

are generally not used for series connection. However, the following characteristics should be considered for three-wire series-connected circuits (Fig. 1.84):

- Each sensor must carry the load current and the burden current for all the downstream sensors (Fig. 1.84).
- When conducting, each sensor will have a voltage drop in series with the load, reducing the available voltage to the load. As with two-wire devices, this and the minimum operating voltage will limit the number of devices wired in series.

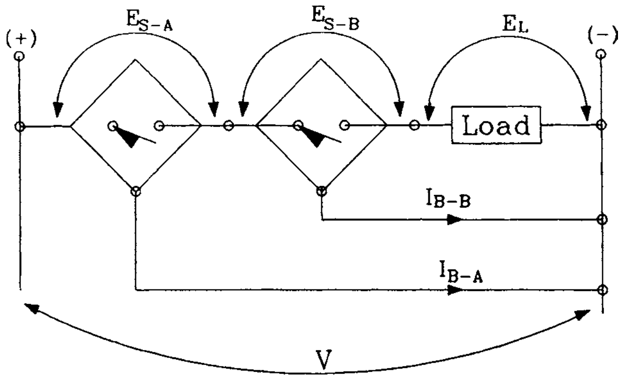


FIGURE 1.84 Series connected, line powered.

- When upstream sensors are not conducting, the downstream sensors are disconnected from their power source and are incapable of responding to a target until the upstream sensors are activated (damped). Time before availability will be increased due to the response in series.

Series and parallel connections that perform logic functions with a connection to a PLC are not common practice. Utilizing sensors this way involves the preceding considerations. It is usually easier to connect directly to the PLC inputs and perform the desired logic function through the PLC program.

Inductive and Capacitive Sensor Response Time—Speed of Operation

When a sensor receives initial power on system power-up, the sensor cannot operate. The sensor operates only after a delay called *time delay before availability* (Fig. 1.85).

In AC sensors, this delay is typically 35 ms. It can be as high as 100 ms in AC circuits with very low residual current and high noise immunity. In DC sensors, the time delay is typically 30 ms.

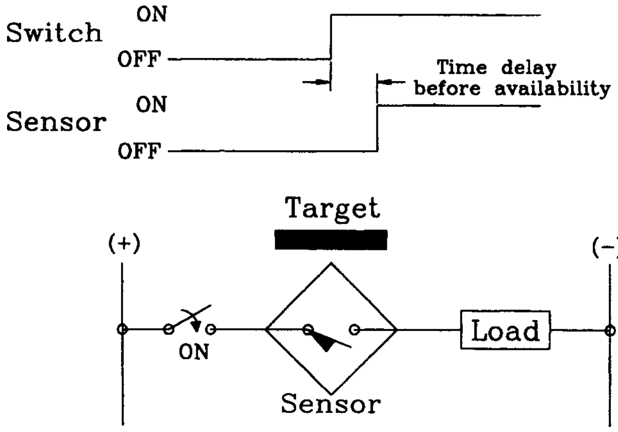


FIGURE 1.85 Time delay prior to availability.

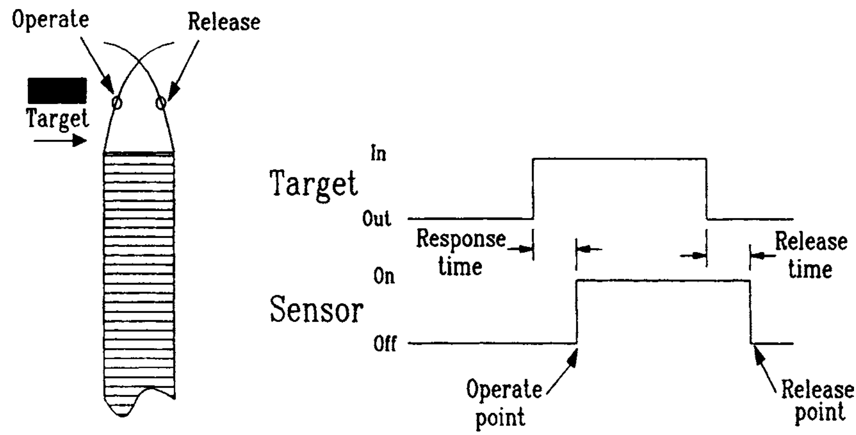


FIGURE 1.86 Response/release times.

Response and Release Time. A target entering the sensing field of either an inductive or a capacitive sensor will cause the detector circuit to change state and initiate an output. This process takes a certain amount of time, called *response time* (Fig. 1.86).

Response time for an AC sensor is typically less than 10 ms. DC devices respond in microseconds. Similarly, when a target leaves the sensing field, there is a slight delay before the switch restores to the OFF state. This is the *release time.* Release time for an AC device is typically one cycle (16.66 ms). The DC device release time is typically 3 ms.

High-Speed Operation. Mechanical devices such as limit switches and relays do not operate at speeds suitable for high-speed counting or other fast-operating-circuit needs. Solid-state devices, however, can operate at speeds of 10, 15, or more operations per second. DC devices can operate at speeds of 500, 1000, or more operations per second.

In order to properly achieve high-speed operation, some basic principles must be applied.

Maximum Target Length. A response delay occurs when a sensor has a target entering the sensing field, as previously stated. There is a similar delay for the respective load to operate. The time from when the sensor conducts and the load operates is the *load response time*. Together, these delays make up the *system response time* T_o.

Similarly, there are delays when the target reaches the release point in the sensing field caused by the sensor release time and the corresponding *load release time*. In order to ensure that the sensor will operate reliably and repeatedly, the target must stay in the field long enough to allow the load to respond. This is called *dwell time*. Figure 1.87 illustrates the time functions for reliable repeatable sensor operation. Figure 1.88 illustrates the dwell range.

Target Duty Cycle. Response (turn-on) times for the sensor and the controlled load may be considerably different from the release (turn-off) times for the same devices. Conditions for the target duty cycles are illustrated by Fig. 1.89. Note that the target is not out of the sensing field long enough to allow the sensor to turn off the controlled load. The application must be arranged so that both sensor and load turn ON and OFF reliably and repeatedly.

Timing Functions. When an inductive control is operating a logic function, an output is generated for the length of time an object is detected (Fig. 1.90).

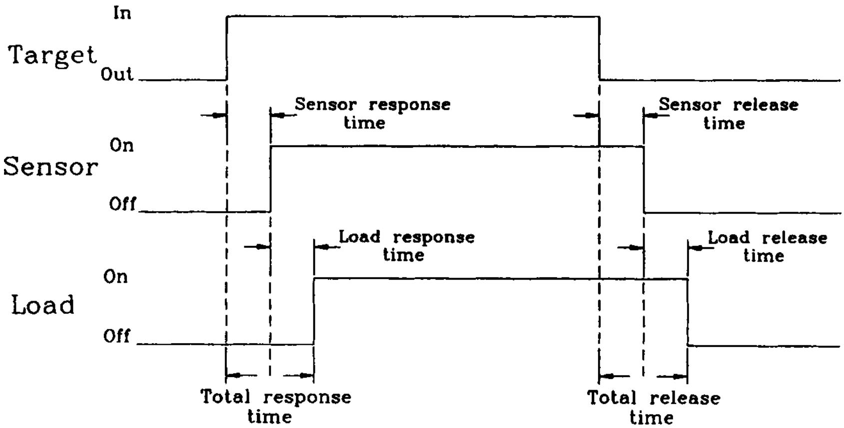


FIGURE 1.87 Maximum target length.

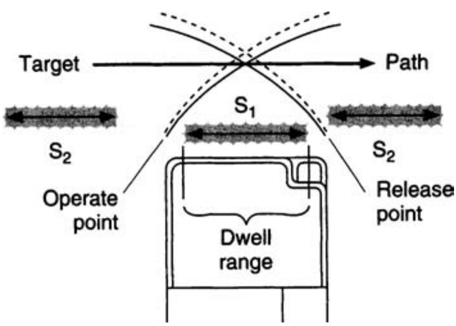


FIGURE 1.88 Dwell range.

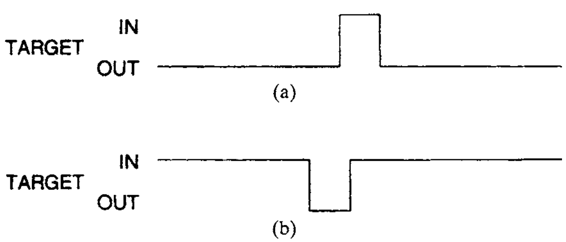


FIGURE 1.89 Target duty cycle. (*a*) Critical response time (turn-on). (*b*) Critical release time (turn-off).

ON Delay Logic. ON delay logic allows the output signal to turn on only after the object has been detected for a predetermined period of time. The output will turn off immediately after the object is no longer detected. This logic is useful if a sensor must avoid false interruption from a small object. ON delay is useful in bin fill or jam detection, since it will not false-trigger in the normal flow of objects going past (Fig. 1.91).

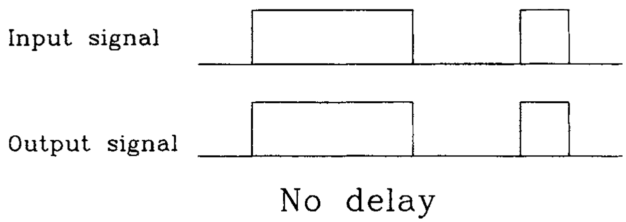


FIGURE 1.90 No delay.

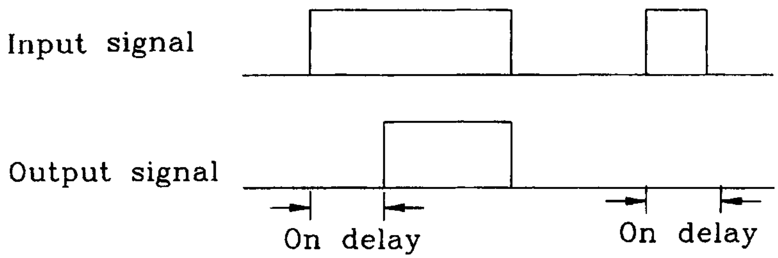


FIGURE 1.91 ON delay.

OFF Delay Logic. OFF delay logic holds the output on for a predetermined period of time after an object is no longer detected. The output is turned on as soon as the object is detected. OFF delay ensures that the output will not drop out despite a short period of signal loss. If an object is once again detected before the output times out, the signal will remain ON. OFF delay logic is useful in applications susceptible to periodic signal loss (Fig. 1.92).

ON/OFF Delay Logic. ON/OFF delay logic combines ON and OFF delay so that the output will be generated only after the object has been detected for a predetermined period of time, and will drop out only after the object is no longer detected for a predetermined period of time. Combining ON and OFF delay smoothes the output of the inductive proximity control (Fig. 1.93).

One-Shot Delay. One-shot logic generates an output of predetermined length no matter how long an object is detected. A standard one-shot must time out before it can be retriggered. One-shot logic is useful in applications that require an output of specified length (Fig. 1.94).

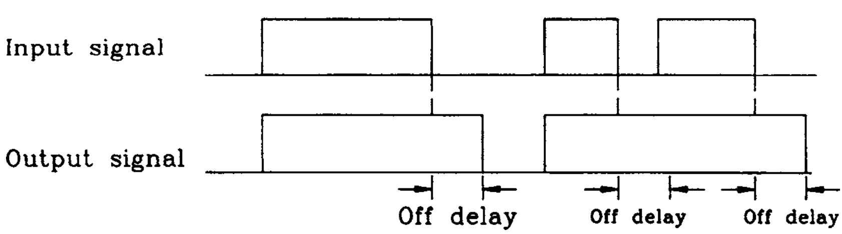


FIGURE 1.92 OFF delay.

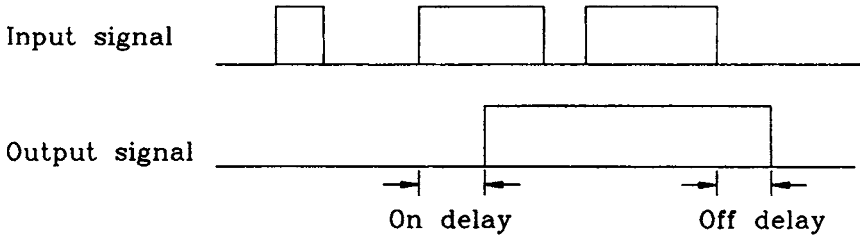


FIGURE 1.93 ON/OFF delay.

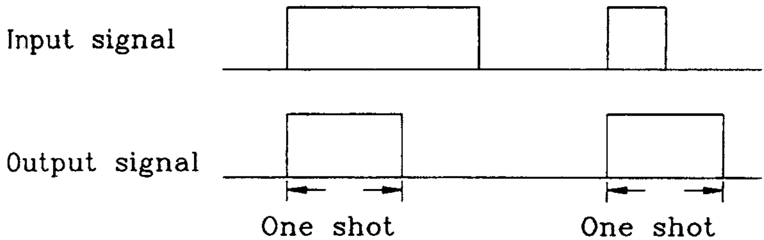


FIGURE 1.94 One-shot.

UNDERSTANDING MICROWAVE-SENSING APPLICATIONS

Microwave sensors are valuable tools in the industrial environment for measuring motion, velocity, direction of movement, and range. They are rugged devices capable of operating in hostile environments. They are intrinsically safe since they have no moving parts and require low power. They will not harm operators and function effectively in explosive environments. They can successfully measure large military and industrial objects over large distances and can provide a great deal of information about the objects, as observed during the Persian Gulf War in 1991.

Microwave technology has long been an effective method of measuring the parameters of motion and presence. Applications range from simple intrusion alarms that merely indicate an object has entered its field of view to complex military radar systems that define the existence, location, and direction of motion.

Microwave sensing technology can be classified into five categories:

- *Motion sensing.* Sensing a moving object in a defined domain—for example, detecting an intruder in a prohibited area.
- *Presence sensing.* Sensing that an object exists in a defined domain at a given time. This concept is vital in industrial control systems where the arrival of an object may not be noticed.
- *Velocity sensing.* Sensing the linear speed of an object in a specified direction. This concept is used by police to detect speeding cars.
- *Direction-of-motion sensing.* Determining whether a target is moving away from or toward the microwave sensor device. This concept is particularly important for manufacturers of automated guided vehicle systems for obstacle avoidance. It is also used to detect whether objects or personnel are approaching or departing from automatic doors.

- *Range sensing.* Measuring the distance from the sensor to an object of interest. Applications include sensing the level of oil or chemical solutions in tanks and containers.

Characteristics of Microwave Sensors

Microwave sensor general characteristics important in industrial and commercial applications are:

- *No contact.* Microwave sensors operate without actually contacting the object. This is particularly important if the object is in a hostile environment or sensitive to wear. They can monitor the speed of power-plant generator shafts, continuously monitoring acceleration and deceleration in order to maintain a constant rotational speed. Microwave sensors can effectively penetrate nonmetallic surfaces, such as fiberglass tanks, to detect liquid levels. They can also detect objects in packaged cartons.
- *Rugged.* Microwave sensors have no moving parts and have proven their reliability in extensive military use. They are packaged in sealed industrial enclosures to endure the rigors of the production environment.
- *Environmental reliability.* Microwave sensors operate reliably in harsh inhospitable environments. They can operate from −55°C to +125°C in dusty, dirty, gusty, polluted, and poisonous areas.
- *Intrinsically safe.* Industrial microwave sensors can be operated in an explosive atmosphere because they do not generate sparks due to friction or electrostatic discharge. Microwave energy is so low that it presents no concern about hazard in industrial applications.
- *Long range.* Microwave sensors are capable of detecting objects at distances of 25 to 45,000 mm or greater, depending on the target size, microwave power available, and the antenna design.
- *Size of microwave sensors.* Microwave sensors are larger than inductive, capacitive, and limit switch sensors. However, use of higher microwave frequencies and advances in microwave circuit development allow the overall package to be significantly smaller and less costly.
- *Target size.* Microwave sensors are better suited to detect large objects than smaller ones such as a single grain of sand.

Principles of Operation

Microwave sensors consist of three major parts: (1) transmission source, (2) focusing antenna, and (3) signal processing receiver.

Usually the transmitter and receiver are combined together in one module, which is called a *transceiver*. A typical module of this type is used by intrusion alarm manufacturers for an indoor alarm system. The transceiver contains a Gunn diode mounted in a small precession cavity that, upon the application of power, oscillates at microwave frequencies. A special cavity design will cause this oscillation to occur at 10.525 GHz, which is one of the few frequencies that the U.S. Federal Communications Commission (FCC) has set aside for motion detectors. Some of this energy is coupled through an iris into an adjoining waveguide. Power output is in the 10- to 20-mW range. The DC input power for this stage (8 V at 150 mA) should be well regulated, since the oscillator is voltage-sensitive. The sensitivity of the system can be significantly reduced by noise (interference).

At the end of the waveguide assembly, a flange is fastened to the antenna. The antenna focuses the microwave energy into a beam, the characteristics of which are determined by the application. Antennas are specified by beam width or gain. The higher the gain, the longer the range and the narrower the beam. An intrusion alarm protecting a certain domain would require a wide-beam antenna to cover the area, while a traffic control microwave sensor would require a narrow-beam high-gain antenna to focus down the road.

Regardless of the antenna selection, when the beam of microwave energy strikes an object, some of the microwave energy is reflected back to the module. The amount of energy will depend on the composition and shape of the target. Metallic surfaces will reflect a great deal, while Styrofoam and plastic will be virtually transparent. A large target area will also reflect more than a small one.

The reflected power measured at the receiver decreases by the fourth power of the distance to the target. This relationship must be taken into consideration when choosing the transmitted power, antenna gain, and signal processing circuitry for a specific application.

When the reflected energy returns to the transceiver, the mixer diode will combine it with a portion of the transmitted signal. If the target is moving toward or away from the module, the phase relationships of these two signals will change and the signal out of the mixer will be an audio frequency proportional to the speed of the target. This is called the *Doppler frequency*. This is of primary concern in measuring velocity and direction of motion. If the target is moving across in front of the module, there will be no Doppler frequency, but there *will* be sufficient change in the mixer output to allow the signal processing circuitry to detect it as unqualified motion in the field.

The signal from the mixer will be in the microvolt to millivolt range, so amplification will be needed to provide a useful level. This amplification should also include 60-Hz and 120-Hz notch filters to eliminate interference from power lines and fluorescent light fixtures, respectively. The remaining bandwidth should be tailored to the application.

Besides amplification, a comparator and output circuitry relays are added to suit the application (Fig. 1.95).

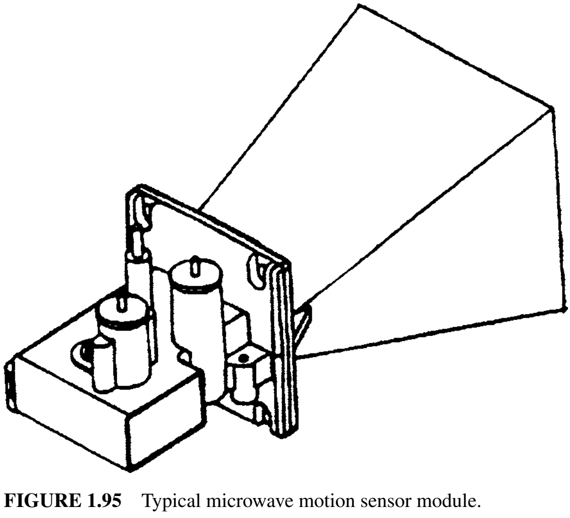


FIGURE 1.95 Typical microwave motion sensor module.

Detecting Motion with Microwave Sensors

The presence of an object in the microwave field disturbs the radiated field. There may be a Doppler frequency associated with the disturbance. The signal from the mixer to the signal processing circuitry may vary with a large amplitude and long duration so it can be detected. The amplitude gain and the delay period are of specific importance in tailoring the device for particular applications, such as motion detection. These sensors are primarily used in intrusion alarm applications where it is only necessary to detect the movement rather than derive further information about the intruder. The sensitivity would be set to the minimum necessary level to detect a person-sized object moving in the domain to be protected to prevent pets or other nonhostile moving objects from causing a false alarm. In addition, some response delay would be introduced for the same reason, requiring continuous movement for some short period of time.

Other applications include parts counting on conveyer belts; serial object counting in general; mold ejection monitoring, particularly in hostile environments; obstacle avoidance in automated guided vehicle systems; fill indication in tanks; and invisible protection screens. In general, this type of sensor is useful where the objects to be sensed are moving in the field of interest (Fig. 1.96).

Other devices that compete for the same applications are ultrasonic, photoelectric, and infrared sensors.

In the intrusion alarm manufacturing industry, microwave sensors have advantages such as a longer range and insensitivity to certain environmental conditions. Ultrasonic sensors are sensitive to drafts and high-frequency ambient noise caused by bells and steam escaping from radiators. Infrared sensors are sensitive to thermal gradients caused by lights turning on and off. The effectiveness of infrared sensors is severely reduced at high ambient temperatures. However, utilizing dual technologies is recommended to minimize false alarms—combining microwave technology with infrared technology, for example. It is necessary for the intruder to be sensed by both technologies before an alarm is given. In other applications, microwave sensors can show advantages over photoelectric sensors in the areas of longer range, increased area of coverage, operation in hostile environments, and in applications where it is necessary to see through one medium (such as a cardboard box or the side of a nonmetallic tank) to sense the object on the other side.

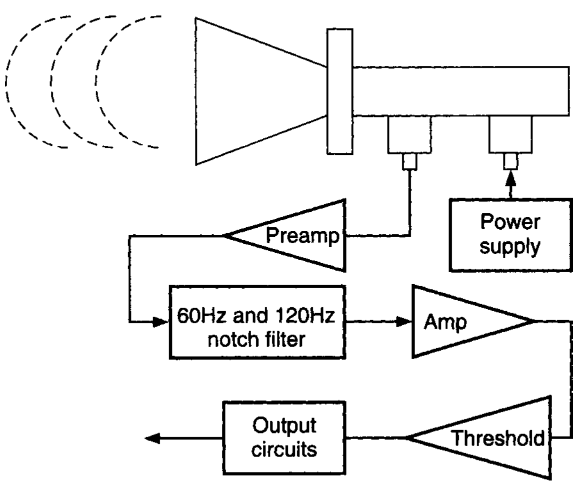


FIGURE 1.96 Microwave motion sensor.

If the target is moving toward or away from the transceiver there will be an audio-frequency (Doppler) signal out of the mixer diode that is proportional to the velocity of the target. The frequency of this signal is given by the formula:

$$F_d = 2V\,(F_t/c) \tag{1.2}$$

where F_d = Doppler frequency
V = velocity of the target
F_t = transmitted microwave frequency
c = speed of light

If the transmitted frequency is 10.525 GHz (the motion detector frequency), this equation simplifies to:

$$F_d = 31.366 \text{ Hz} \times V \quad \text{in miles/hour}$$

or

$$F_d = 19.490 \text{ kHz} \times V \quad \text{in kilometers/hour} \tag{1.3}$$

or

$$F_d = 84.313 \text{ kHz} \times V \quad \text{in furlongs/fortnight}$$

This assumes that the target is traveling directly at or away from the transceiver. If there is an angle involved, then the equation becomes:

$$F_d = 2V(F_t/c)\cos\Theta \tag{1.4}$$

where Θ is the angle between the transceiver and the line of movement of the target. Evidently, as the target is moving across the face of the transceiver, $\cos\Theta = 0$, and the frequency is 0. If the angle is kept below 18°, however, the measured frequency will be within 5 percent of the center frequency (Fig. 1.97).

Signal processing for this module must include amplification, a comparison network to shape the signal into logic levels, and a timing and counting circuit to either drive a display device or compare the frequency to certain limits. If more than one moving object is in the microwave field, it may be necessary to discriminate on the basis of amplitude or frequency bandwidth, limiting it to exclude unwanted frequencies. Velocities near 3 km/h and 6 km/h are also difficult to measure with this system since the corresponding Doppler frequencies are 60 and 120 Hz, which are prime interference frequencies from power lines and fluorescent fixtures. Extra shielding or isolation will be necessary in this case. False alarm rates may also be reduced by counting a specific number of cycles before triggering an output. This will actually correspond to the target moving a defined distance.

Microwave sensors are well suited for measuring the velocity of objects, which most other sensors cannot do directly. Inductive, photoelectric, and other sensors can measure radial

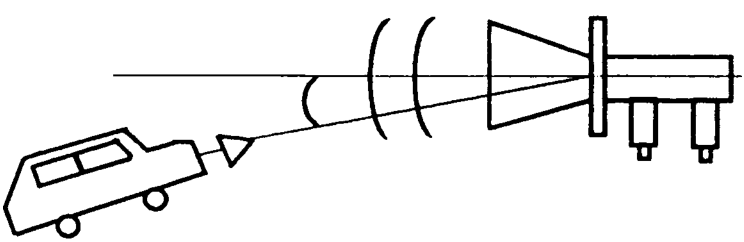


FIGURE 1.97 Angular velocity measurement.

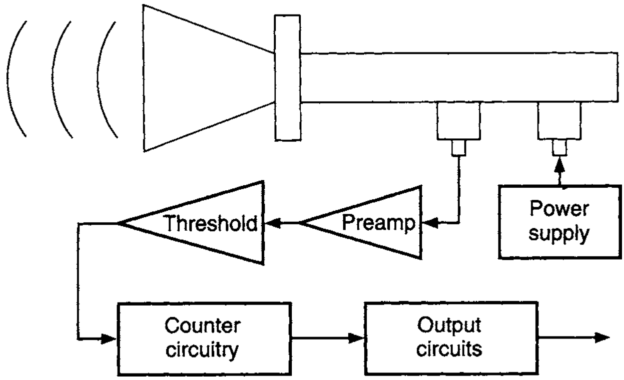


FIGURE 1.98 Velocity sensing module.

velocity. For example, inductive photoelectric sensors measure radial velocity when configured as a tachometer, and if the rotating element is configured as a trailing wheel, then linear velocity can be defined. Photoelectric sensors can also be set up with appropriate signal processing to measure the time that a moving object takes to break two consecutive beams. This restricts the measurement to a specific location. Multiple beams would be needed to measure velocity over a distance, whereas a single microwave sensor could accomplish the same result.

Aside from their use in police radars, microwave sensors can measure the speed of baseball pitches. There are many industrial applications for these sensors as well. Microwave sensors are an excellent means of closed-loop speed control of a relatively high-speed rotating shaft (3600 r/min). Other applications include autonomous-vehicle speed monitoring and independent safety monitoring equipment for heavy and high-speed machine tools. Also, a microwave sensor will detect an overvelocity condition (Fig. 1.98).

A microwave sensor, mounted on a tractor or other farm equipment to measure ground speed, will play an important role in reducing excessive distribution of seeds and fertilizer per acre. The ordinary wheel driven–speedometer is not sufficiently accurate because of wheel slippage. Accurate speed measurement is necessary in these vehicles, so that seeds and fertilizer are spread at a specific rate by the accessory equipment; an over- or underestimate of the speed will result in the wrong density per acre.

Detecting Presence with Microwave Sensors

A static object can be detected in the field of a microwave sensor. The purpose of this detection is to determine that the object is still in the field of interest and has not departed. This is particularly desirous in control systems where the controller is performing other tasks and then accesses the sensor to determine whether there is a sensed object at that particular time. In this situation, presence sensing is especially advantageous since the output can be verified by further interrogations to eliminate false sensing.

To detect the presence of an object, a microwave sensor with a separate transmitter and receiver must be used. A transceiver in this application is not adequate, although the transmitter and the receiver can be mounted in the same enclosure. The receiver must not sense any energy unless the object is present in the field. A means to modulate the transmitter is needed, and the receiver should be narrow-band to amplify and detect the modulated reflection. The sensitivity of the receiver must be adjustable to allow for ambient reflections.

Microwave sensors have been extensively and successfully tested at various fast-food drive-through vending locations. Other types of sensors such as ultrasonic and photoelectric sensors, were also tested, less successfully. They were sensitive to the environment. It was discovered that frost heaving of the ground would eventually cause their buried loop to fail, and the cost of underground excavation to replace the loop was exorbitant.

Another application of the microwave sensor is the door-opening market. The microwave sensor will check, for safety reasons, the area behind a swinging door to detect whether there is an individual or an object in the path. Ultrasonic sensors may perform the same task, yet range and environmental conditions often make a microwave sensor more desirable.

A microwave sensor can check boxes to verify that objects actually have been packed therein. The sensor has the ability to see through the box itself and triggers only if an object is contained in the box. This technology relies on the sensed object being more reflective than the package, a condition that is often met.

Measuring Velocity with Microwave Sensors

Microwave sensors are ideally suited to measuring linear velocity. Police radar is a simple example of a Doppler-frequency-based velocity sensor. This technology can be applied wherever it is necessary to determine velocity in a noncontact manner.

Detecting Direction of Motion with Microwave Sensors

Direction of motion—whether a target is moving toward or away from the microwave sensor—can be determined by through use of the Doppler-frequency concept (Fig. 1.99) by adding an extra mixer diode to the module. A discriminating effect is generated by the additional diode, which is located in the waveguide such that the Doppler outputs from the two mixers differ in phase by one-quarter wavelength, or 90°. These outputs will be separately amplified and converted into logic levels. The resulting signals can then be fed into a digital phase-discrimination circuit to determine the direction of motion. Such circuits are commonly found in motion control applications in conjunction with optical encoders. Figure 1.100 shows the phase relationships of the different directions.

Outputs from this module can vary widely to suit the application. The simplest is two outputs, one for motion and the other for direction (toward or away). These outputs can

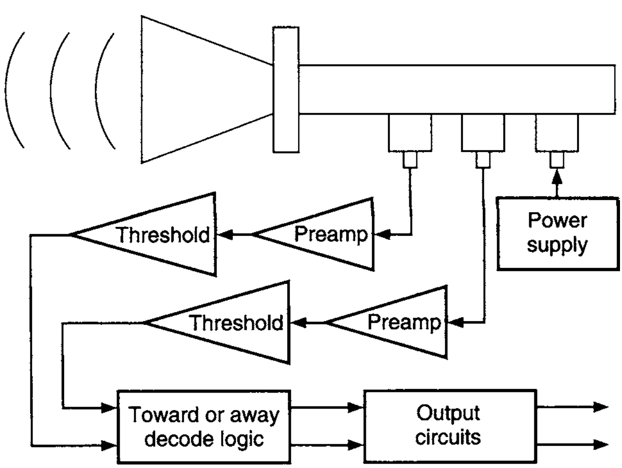


FIGURE 1.99 Direction of motion sensor schematic.

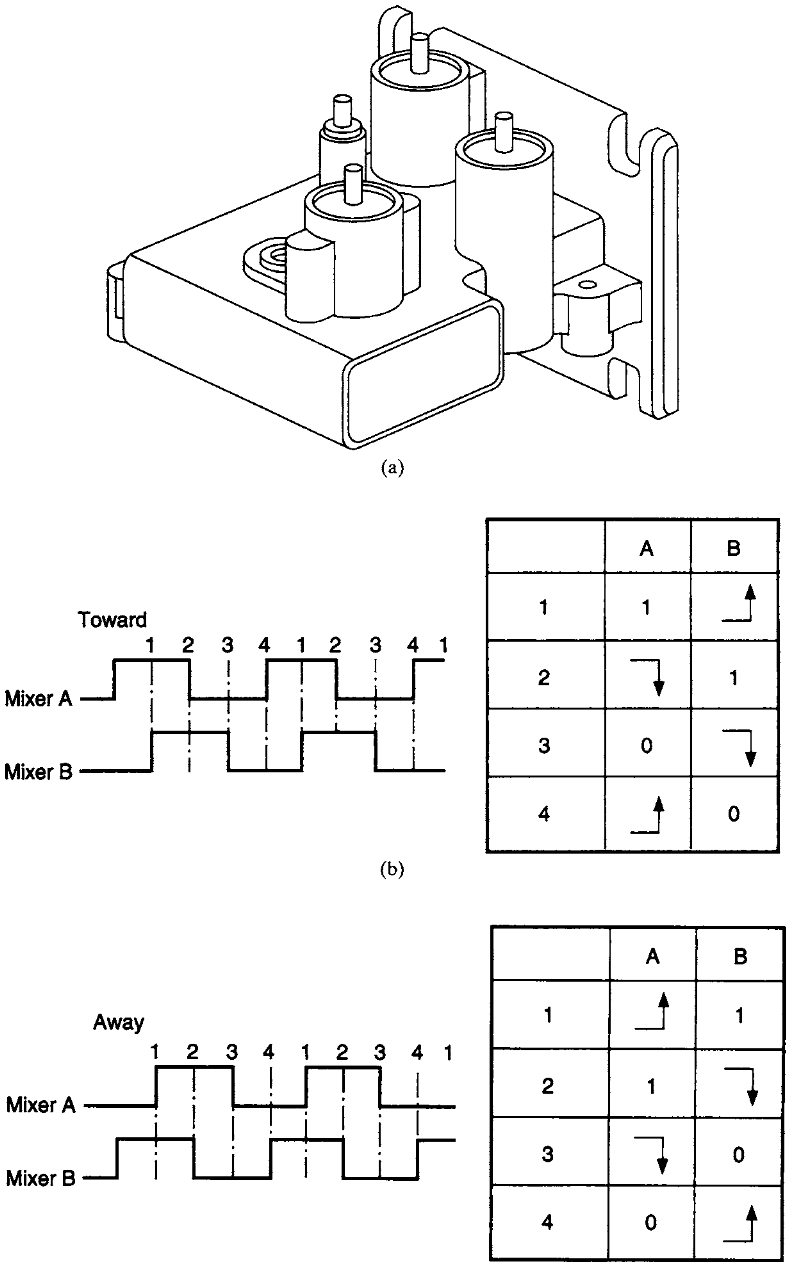


FIGURE 1.100 (*a*) Direction of motion sensor device. (*b*) Motion logic "away." (*c*) Direction logic "toward."

be added to a third, which provides the velocity of the target. The combination of signals could be analyzed to provide a final output when specific amplitude, direction, distance, and velocity criteria are met (Fig. 1.99).

In the door-opening field, using the amplitude, direction, distance, and velocity information reduces the number of false openings. This extends the life of the door mechanism, besides saving heat if the door is an entrance to a heated building.

In this case, the measurements by circuitry indicate the following characteristics:

Characteristic	Measurement
Person-sized object	Amplitude of return
Moving at walking pace	Velocity
Toward or away	Direction
Specific time before opening	Distance

Detecting Range with Microwave Sensors

An early-warning military radar system depends on costly microwave sensors. A small yacht may use a microwave sensor selling for less than $1000 to detect targets at ranges up to 5 miles.

Regardless of their cost, microwave range sensors for commercial, industrial, and military applications employ essentially the same measuring technique. They transmit a narrow pulse of energy and measure the time required for the return from the target. Since microwave energy propagates at the speed of light, the time for the pulse to reach the target and return is 2 ns per foot of range. If the range to the target is 1 mi, the time required is 10.56 μs.

Although the microwave power needed is sufficient to raise the sensor temperature to 500°F, the design of the signal processing circuitry to measure the response is not difficult. However, if the target is very close to the transmitter, then the short response time may pose a real problem. At 3 ft, the time response is 6 ns. For 1-in resolution, the circuitry must be able to resolve 167 ps. This may pose a significant problem.

The alternative method to resolve a target at a short range involves changing the frequency of a continuous oscillator. This method is better suited to industrial applications. An oscillator starting at 10.525 GHz and sweeping at 50 MHz in 10 ms in the 6 ns mentioned earlier will have changed its frequency by:

$$(6 \text{ ns} \times 50 \text{ MHz}/0.01 \text{ s}) = 30 \text{ Hz} \tag{1.5}$$

The returning wave will be still at 10.525 GHz. The output from the mixer diode as the sweep continues will be the 30-Hz difference. If this frequency is averaged over time, it is not difficult to resolve a range to 0.001 in.

The preceding calculation indicates that the frequency is high for faraway objects and low for targets that are close. This leads to two conclusions:

- The closer the object is, the lower the frequency, and therefore the longer the measurement will take.
- The signal processing amplifier should have a gain that increases with frequency.

Problems can arise when the target is moving or there are multiple targets in the area of interest. Movement can be detected by comparing consecutive readings and can be used as a discrimination technique. Multiple targets can be defined by narrow-beam antennas to reduce the width of the area of interest. Gain adjustments are also required to eliminate all but the largest target. Audio-bandwidth filters may be used to divide the range into sectors for greater accuracy.

Other sensor types, such as photoelectric and inductive sensors, may be utilized to measure distances. Inductive sensors are used in tank level measurements. They must be coupled with a moving component that floats on the surface of the substance to be measured. Photoelectric sensors measure position by focusing a beam on a point in space and measuring the reflection on a linear array. This can give very precise measurements over a limited range but is subject to adverse environmental conditions. Photoelectric sensors can focus a camera on a target for a better picture. Ultrasonic sensors may perform the same function, but their range is limited and they can be defeated by hostile environments.

Microwave sensors for measuring range have an impressive array of applications, including measurement of the level of liquid or solid in a tank, sophisticated intrusion alarms, autonomous guided vehicle industrial systems, and noncontact limit switching. In tank level sensing in the chemical industry (Fig. 1.101), the microwave sensor is mounted at the top of the tank and measures the distance from that position to the surface of the contents. Since the electronic circuitries can be isolated from the tank contents by a sealed window, it is intrinsically safe. It has the advantage of being a noncontact system, which means there are no moving parts to break or be cleaned. This allows the microwave sensor to be used on aggressive chemicals, liquids, liquefied gases, highly viscous substances, and solids such as grain and coal.

Microwave Technology Advancement

Advances in technology have opened up important new applications for microwave sensors. The expected governmental permission to utilize higher frequencies and the decreasing size of signal processing circuitry will significantly reduce the cost of the sensors and will enable them to detect even smaller targets at a higher resolution. Microwave integrated circuit technology (MICT), presently developed for the Military Microwave Integrated Circuit (MIMIC) program will overflow into the industrial market, causing increases

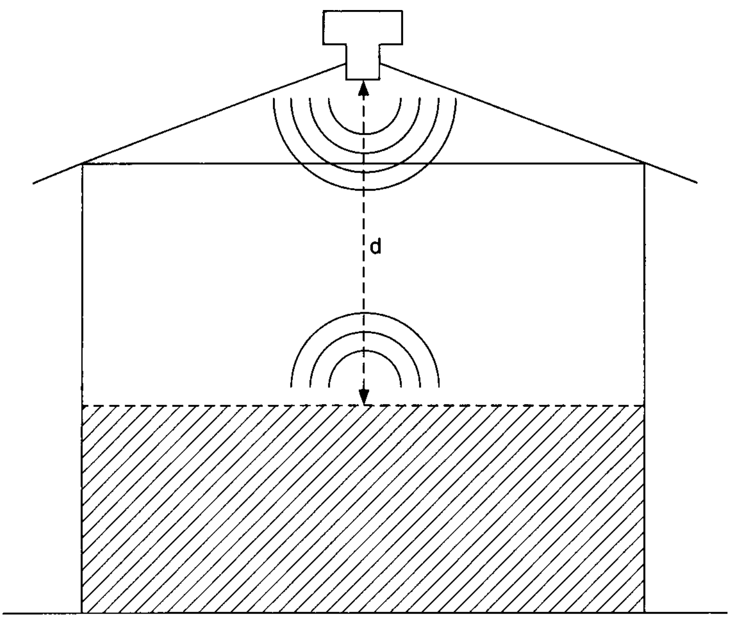


FIGURE 1.101 Tank level sensor.

in performance and in analysis capabilities. Consequently, the utilization of computer-integrated manufacturing technology will be broadened.

THE INFRARED SPECTRUM: UNDERSTANDING AN INFRARED SPECTRUM AND HOW IT ARISES FROM BOND VIBRATIONS WITHIN ORGANIC MOLECULES

How an Infrared Spectrum Is Produced

Visible light is made up of a continuous range of different electromagnetic frequencies, where each frequency can be seen as a different color. Infrared radiation also consists of a continuous range of frequencies that humans can not detect.

If a range of infrared frequencies shines one at a time through a sample of an organic compound, some frequencies are absorbed by the compound. A detector on the other side of the compound would show that some frequencies pass through the compound with almost no loss, but other frequencies are strongly absorbed.

A particular frequency passes through the compound and is measured as the "*percentage transmittance.*"

A percentage transmittance of 100 would mean that all of that frequency passed straight through the compound without any being absorbed. In practice, this is unlikely to occur since there is always some small loss, giving a transmittance of perhaps 95 percent as the best one may achieve.

A transmittance of only 5 percent would mean that nearly all of the particular frequency was absorbed by the compound. A very high absorption of this type indicates the nature of bonds in the compound.

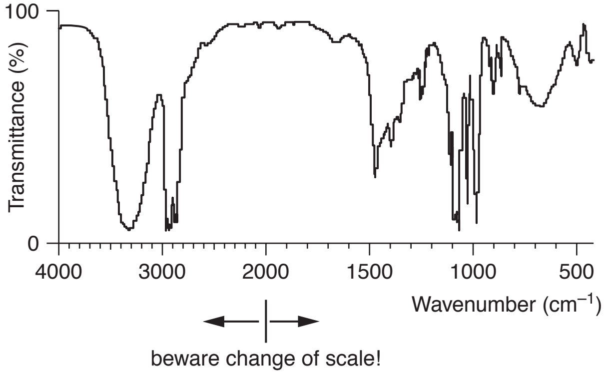
The Infrared Spectrum

A graph is displayed next, showing how the percentage transmittance varies with the frequency of the infrared radiation.

"Wavenumber" is defined as:

$$\text{Wavenumber} = 1/\text{wavelength (cm)} = \text{cm}^{-1}$$

Infra-red spectrum of propan-1-ol, $CH_3CH_2CH_2OH$



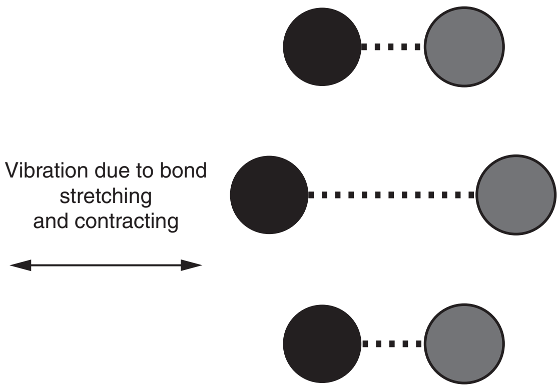
Frequencies to Be Absorbed

Each frequency of light, including infrared, has a certain energy. If a particular frequency is being absorbed as it passes through the compound being investigated, it must mean that its energy is being transferred to the compound.

Energies in infrared radiation correspond to the energies involved in bond vibrations.

Bond Stretching. In covalent bonds, atoms are not joined by rigid links. The two atoms are held together because both nuclei are attracted to the same pair of electrons. The two nuclei can vibrate backwards and forward—toward and away from each other—around an average position.

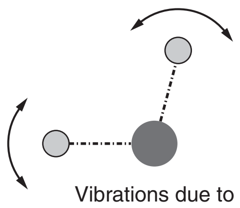
The following diagram shows the stretching that happens in a carbon-oxygen single bond. Other atoms are attached to both the carbon and the oxygen. For example, it could be the carbon-oxygen bond in methanol, CH_3OH.



The energy involved in this vibration depends on things like the length of the bond and the mass of the atoms at either end. Each different bond will vibrate in a different way, involving different amounts of energy.

Bonds are vibrating at all times. However, if the right amount of energy is subjected to the bond, the atom responds to the move and kicks into a higher state of vibration. The amount of energy it needs will vary from bond to bond. Each different bond will absorb a different frequency, and hence the energy of infrared radiation.

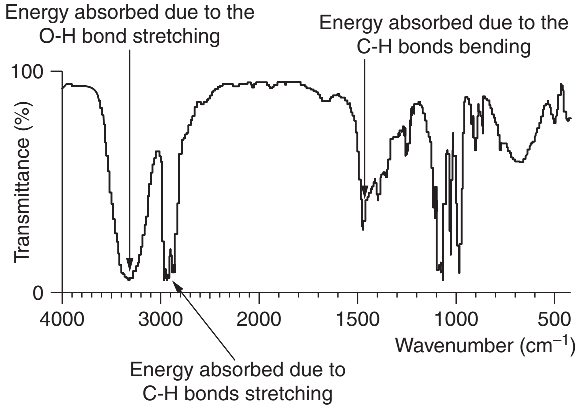
Bond Bending. The diagram shows the bending of the bonds in a water molecule. The bond angle between the two hydrogen-oxygen bonds fluctuates slightly around its average value. Imagine a lab model of a water molecule where the atoms are joined together with springs. These bending vibrations are what one may observe if the model is gently shaken.



Bonds will be vibrating as illustrated at times. If the right amount of energy is subjected to the bond, the bond will respond by jumping to a higher state of vibration. Since the energies involved with the bending will be different for each kind of bond. Each different bond will absorb a different frequency of infrared radiation in order to make this change from one state to a higher one.

The infrared spectrum of propan-1-ol, $CH_3CH_2CH_2OH$:

In the illustrated diagram, three sample absorptions are selected to show how the bond vibrations produce them. The bond stretching and bending produce different troughs in the spectrum.



UNDERSTANDING LASER SENSORS

Two theories about the nature of light have been recognized. The particle theory was the first presented to explain the phenomena that were observed concerning light. According to this theory, light is a particle with mass, producing reflected beams. It was believed that light sources actually generated large quantities of these particles. Through the years, however, many phenomena of light could not be explained by the particle theory, such as reflection of light as it passes through optically transparent materials.

The second theory considered that light was a wave, traveling with characteristics similar to those of water waves. Many, but not all, phenomena of light can be explained by this theory.

A dual theory of light has been proposed, and is presently considered to be the true explanation of light propagation. This theory suggests that light travels in small packets of wave energy called *photons*. Even though photons are *bundles* of wave energy, they have momentum like particles of mass. Thus, light is wave energy traveling with some of the characteristics of a moving particle. The total transmitted as light is the sum of the energies of all the individual photons emitted.

Velocity, frequency, and wavelength are related by the equation:

$$c = f\lambda \tag{1.6}$$

where c = velocity of light, km/s
f = frequency, Hz
λ = wavelength, m

This equation shows that the frequency of a wave is inversely proportional to the wavelength—that is, higher-frequency waves have shorter wavelengths.

Properties of Laser Light

Laser stands for *light amplification by stimulated emission of radiation*. Laser light is monochromatic, whereas standard white light consists of all the colors in the spectrum and is broken into its component colors when it passes through a standard glass prism (Figs. 1.102 and 1.103).

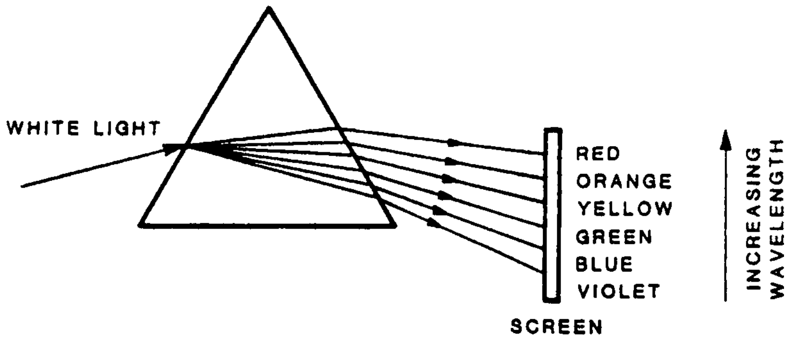


FIGURE 1.102 Standard white light.

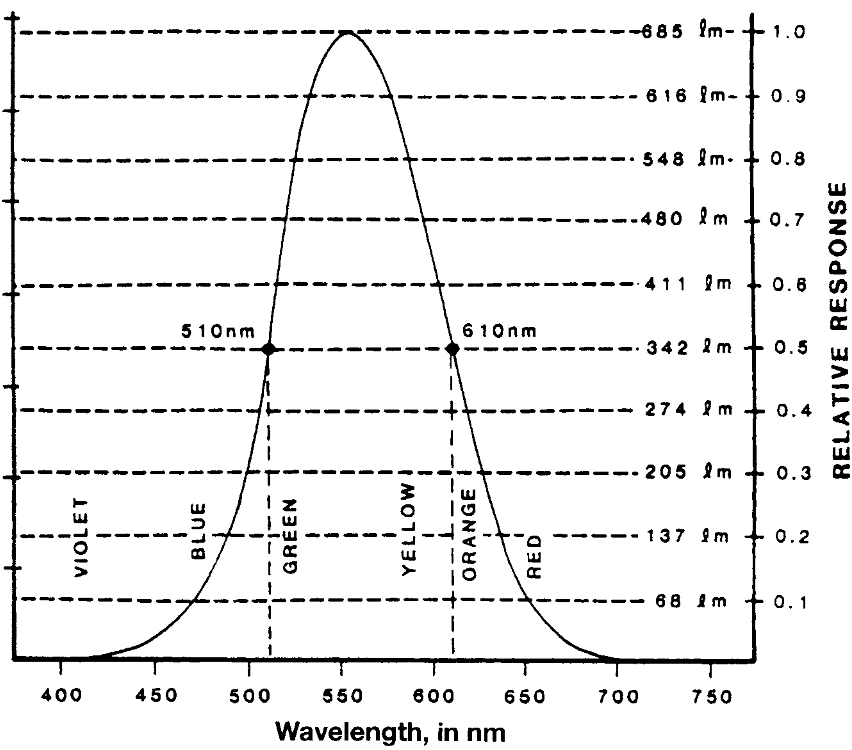


FIGURE 1.103 Spectrum of standard white light.

Essential Laser Components

Laser systems consist of four essential components:

- The active medium
- The excitation mechanism
- The feedback mechanism
- The output coupler

The Active Medium. The active medium is the collection of atoms, ions, or molecules in which stimulated emission occurs. It is in this medium that laser light is produced. The active medium can be a solid, liquid, gas, or semiconductor material. Often, the laser takes its name from that of the active medium. For example, the ruby laser has a crystal of ruby as its active medium while the CO_2 laser has carbon dioxide gas.

The wavelength emitted by a laser is a function of the active medium. This is because the atoms within the active medium have their own characteristic energy levels at which they release photons. It will be shown later that only certain energy levels within the atom can be used to enhance stimulated emission. Therefore, a given active medium can produce a limited number of laser wavelengths, and two different active media cannot produce the same wavelengths. Table 1.8 contains a list of materials commonly used in lasers, along with the corresponding wavelengths these materials produce.

TABLE 1.8 Wavelengths of Laser Materials

Type of active medium	Common material	Wavelength produced, nm
Solid	Ruby	694
	Nd:YAG	1,060
	Nd:glass	1,060
	Erbium	1,612
Liquid	Organic dyes	360–650
Gas	Argon (ionized)	488
	Helium-neon	632.8
	Krypton (ionized)	647
	CO_2	10,600
Semiconductor	Gallium arsenide	850
	Gallium antimonide	1,600
	Indium arsenide	3,200

The active medium is the substance that actually lases. In the helium-neon laser, only the helium lases.

Excitation Mechanism. The excitation mechanism is the device used to put energy into the active medium. The three primary types of excitation mechanisms are optical, electrical, and chemical. All three provide the energy necessary to raise the energy state of the atom, ion, or molecule of the active medium to an excited state. The process of imparting energy to the active medium is called *pumping the laser*.

Optical Excitation. An optical excitation mechanism uses light energy of the proper wavelength to excite the active medium. The light may come from any of several sources, including a flash lamp, a continuous arc lamp, another laser, or even the sun. Although most of these use an electric power supply to produce the light, it is not the electrical energy

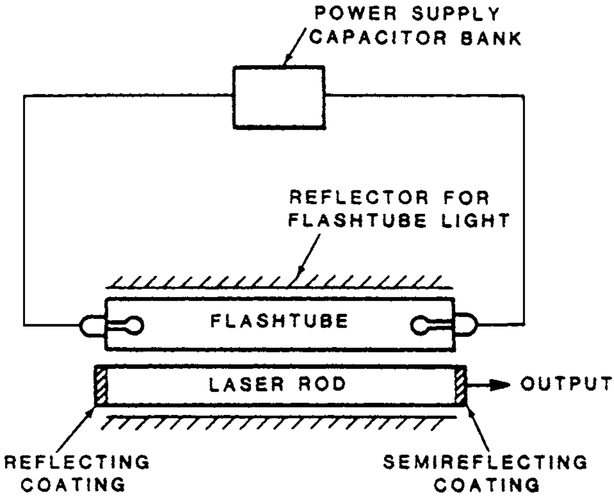


FIGURE 1.104 Solid laser with optical pumping source.

that is used directly to excite the atoms of the active medium but rather the light energy produced by the excitation mechanism.

Optical excitation is generally used with active media that do not conduct electricity—as with solid lasers like the ruby. Fig. 1.104 is a schematic drawing of a solid laser with an optical pumping source.

The sun is considered a possible optical pumping source for lasers in space. The optical energy from the sun could be focused by curved mirrors onto the laser's active medium. Since the size and weight of an electric power supply is of concern in space travel, solar pumping of lasers is an interesting alternative.

Electrical Excitation. Electrical excitation is most commonly used when the active medium will support an electric current. This is usually the case with gases and semiconductor materials.

When a high voltage is applied to a gas, current-carrying electrons, or ions move through the active medium. As they collide with the atoms, ions, or molecules of the active medium, their energy is transferred and excitation occurs. The atoms, ions, and electrons within the active medium are called *plasma*.

Figure 1.105 is a schematic drawing of a gas laser system with electrical excitation. The gas mixture is held in a gas plasma tube and the power supply is connected to the ends of the plasma tube. When the power supply is turned on, electron movement within the tube is from the negative to the positive terminal.

Chemical Excitation. Chemical excitation is used in a number of lasers. When certain chemicals are mixed, energy is released as chemical bonds are made or broken. This energy can be used as a pumping source. It is most commonly used in hydrogen-fluoride lasers, which are extremely high-powered devices used primarily in military weapons and research. These lasers are attractive for military applications because of the large power-to-weight ratio.

Feedback Mechanism. Mirrors at each end of the active medium are used as a *feedback mechanism*. The mirrors reflect the light produced in the active medium back into the medium along its longitudinal axis. When the mirrors are aligned parallel to each other, they form a resonant cavity for the light waves produced within the laser. They reflect the light waves back and forth through the active medium.

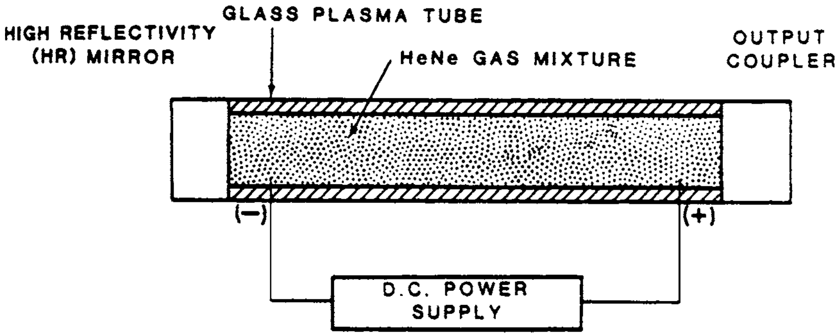


FIGURE 1.105 Gas laser with electrical excitation.

In order to keep stimulated emission at a maximum, light must be kept within the amplifying medium for the greatest possible distance. In effect, mirrors increase the distance traveled by the light through the active medium. The path that the light takes through the active medium is determined by the shape of the mirrors. Figure 1.106 shows some of the possible mirror combinations. Curved mirrors are often used to alter the direction in which the reflected light moves.

Output Coupler. The feedback mechanism keeps the light inside the laser cavity. In order to produce an output beam, a portion of the light in the cavity must be allowed to escape. However, this escape must be controlled. This is most commonly accomplished by using a partially reflective mirror in the feedback mechanism. The amount of reflectance varies with the type of laser. A high-power laser may reflect as little as 35 percent, with the remaining 65 percent being transmitted through the mirror to become the output laser beam. A low-power laser may require an output mirror reflectivity as high as 98 percent, leaving only 2 percent to be transmitted. The output mirror that is designed to transmit a given percentage of the laser light in the cavity between the feedback mirrors is called the *output coupler*.

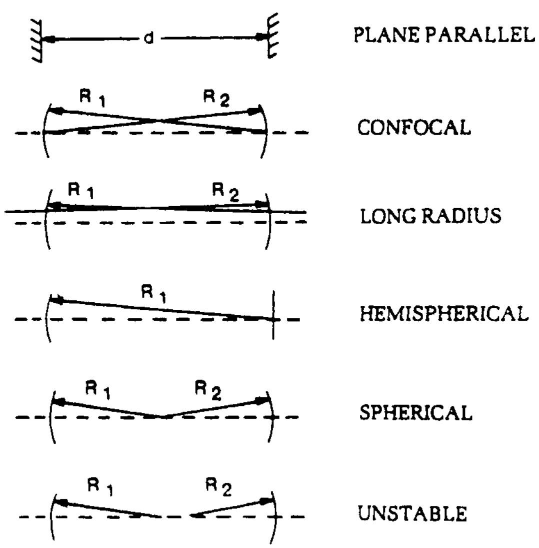


FIGURE 1.106 Mirror combinations for feedback mechanism.

Semiconductor Displacement Laser Sensors

Semiconductor displacement laser sensors, consisting of a light-metering element and a position-sensitive detector (PSD), detect targets by using triangulation. A light-emitting diode or semiconductor laser is used as the light source. A semiconductor laser beam is focused on the target by the lens. The target reflects the beam, which is then focused on the PSD, forming a beam spot. The beam spot moves on the PSD as the target moves. The displacement of the workpiece can then be determined by detecting the movement of the beam spot.

Industrial Applications of Semiconductor Displacement Lasers. The laser beam emitted from laser diode in the transmitter is converged into a parallel beam by the lens unit. The laser beam is then directed through the slit on the receiver and focused on the light-receiving element. As the target moves through the parallel laser beam, the change in the size of the shadow is translated into the change in received light quantity (voltage). The resulting voltage is used as a comparator to generate an analog output voltage.

Industrial Applications of Laser Sensors*

Electrical and electronics industries:

- *Warpage and pitch of IC leads.* The visible beam spot facilitates the positioning of the sensor head for small workpieces. Warpage and pitch can be measured by scanning IC leads with the sensor head (Fig. 1.107).
- *Measurement of lead pitch of electronic components.* The sensor performs precise noncontact measurement of pitch using a laser beam (Fig. 1.108).
- *Measurement of disk head movement.* The laser sensor is connected to a computer in order to compare the pulse input to the disk head drive unit with actual movement. The measurement is done on-line, thus increasing productivity (Fig. 1.109).

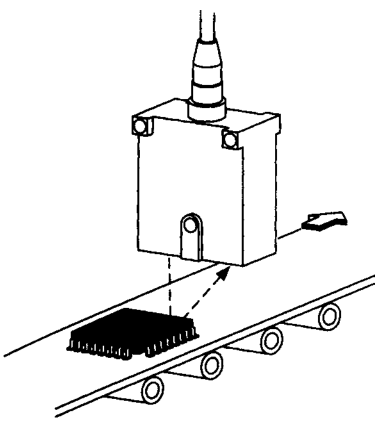


FIGURE 1.107 Warpage and pitch of IC lead.

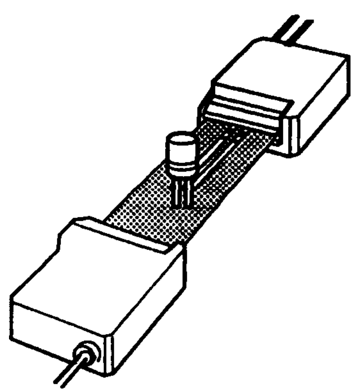


FIGURE 1.108 Measurement of lead pitch of electronic components.

*A few nonlaser optical sensors are included, as indicated.

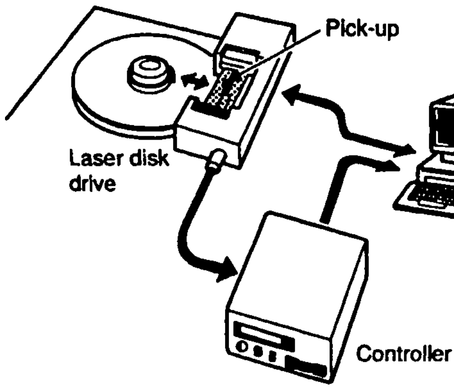


FIGURE 1.109 Measurement of disk head movement.

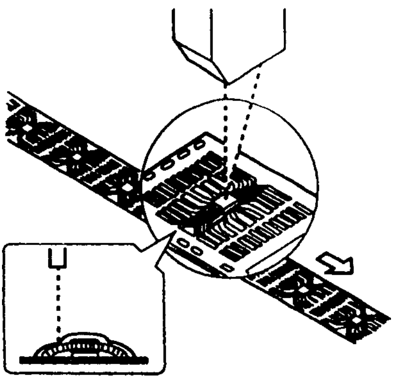


FIGURE 1.110 Detection of presence/absence of resin coating.

- *Detection of presence/absence of resin coating.* The laser displacement sensor determines whether a resin coating was formed after wire bonding (Fig. 1.110).
- *Detection of double-fed or mispositioned resistors prior to taping.* Through-beam-type sensor heads are positioned above and below the resistors traveling on a transfer line. A variation on the line changes the quantity of light in the laser beam, thus signaling a defect (Fig. 1.111).
- *Detection of defective shrink wrapping of videocassette.* Defective film may wrap or tear during shrink wrapping. The laser sensor detects defective wrapping by detecting a change in the light quantity on the surface of the videocassette (Fig. 1.112).

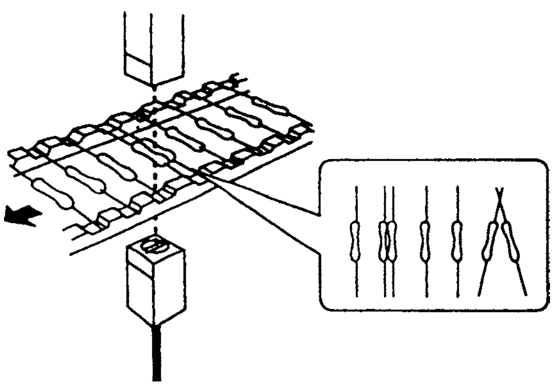


FIGURE 1.111 Detection of double-fed or mispositioned resistors.

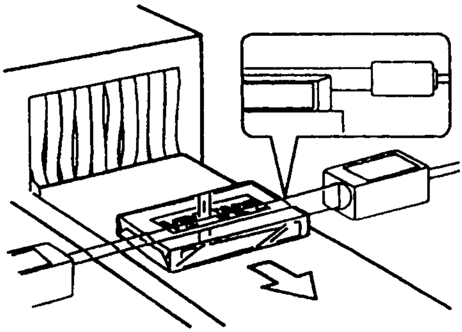


FIGURE 1.112 Detection of defective shrink wrapping of videocassette.

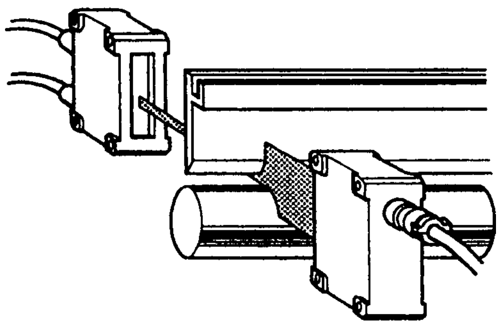


FIGURE 1.113 Measurement of gap between roller and doctor blade.

- *Measurement of gap between roller and doctor blade.* Measures the gap between the roller and the doctor blade in submicrometer units. The sensor's automatic measurement operation eliminates reading errors (Fig. 1.113).
- *Measurement of surface run-out of laser disk.* The surface run-out of a laser disk is measured at a precision of 0.5 µm. The sensor head enables measurement on a mirror-surface object (Fig. 1.114).

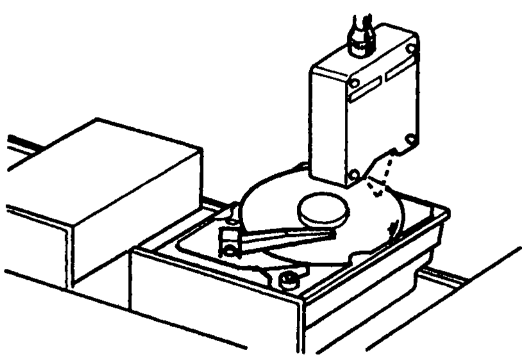


FIGURE 1.114 Measurement of surface run-out of laser disk.

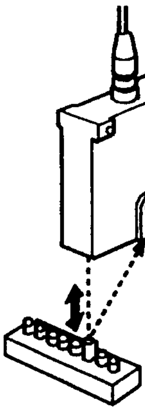


FIGURE 1.115 Displacement of printer impact pins.

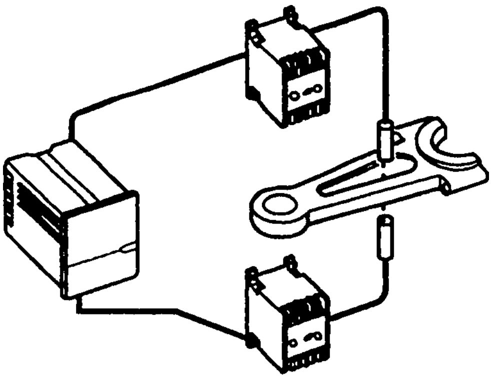


FIGURE 1.116 Measurement of thickness of connecting rod.

- *Displacement of printer impact pins.* The visible beam spot facilitates positioning of the head of a pin-shaped workpiece, enabling measurement of the vertical displacement of impact pins (Fig. 1.115).

Automotive manufacturing industries:

- *Measurement of thickness of connecting rod.* Measures the thickness of the connecting rod by processing the analog inputs in the digital meter relay (Fig. 1.116).
- *Measurement of depth of valve recesses in piston head.* Measures the depth of the valve recesses in the piston head so that chamber capacity can be measured. Iron jigs are mounted in front of the sensor head, and the sensor measures the distance the jigs travel when they are pressed onto the piston head (Fig. 1.117).*

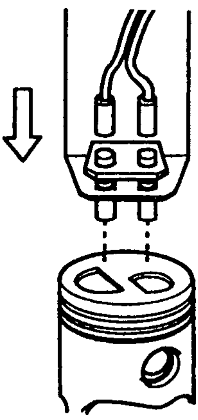


FIGURE 1.117 Measurement of depth of valve recesses in piston head.

- *Measurement of height of radiator fin.* Detects improper radiator fin height by comparing the bottom value of the analog output with a stored pair of tolerances (Fig. 1.118).
- *Measurement of outer diameter of engine valve.* The laser scan micrometer allows on-line measurement of the outer diameter of engine valves simply by positioning a separate sensor head on either side of the conveyer (Fig. 1.119).
- *Positioning of robot arm.* The laser displacement sensor is used to maintain a specific distance between the robot arm and target. The sensor outputs a plus or minus voltage if the distance becomes greater or less, respectively, than the 100-mm reference distance (Fig. 1.120).
- *Detection of damage on microdiameter tool.* Detects a break, chip, or excess swarf from the variation of light quantity received (Fig. 1.121).

*Nonlaser sensor.

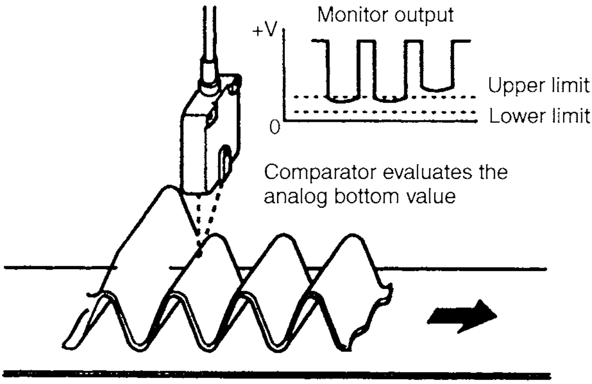


FIGURE 1.118 Measurement of height of radiator fin.

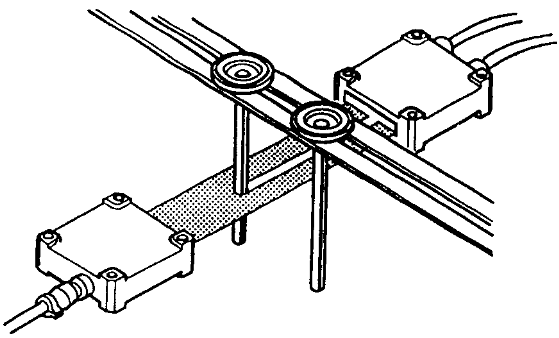


FIGURE 1.119 Measurement of outer diameter of engine valve.

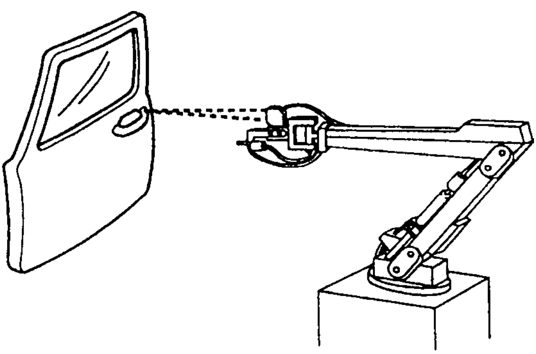


FIGURE 1.120 Positioning of robot arm.

Metal/steel/nonferrous industries:

- *Detection of misfeeding in high-speed press.* The noncontact laser sensor, timed by a cam in the press, confirms the material feed by monitoring the pilot holes and then outputs the result to an external digital meter relay (Fig. 1.122).

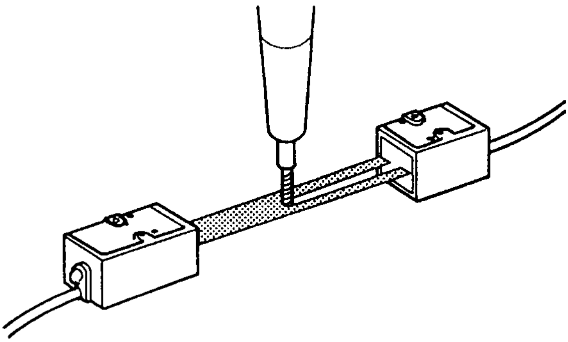


FIGURE 1.121 Detection of damage on microdiameter tool.

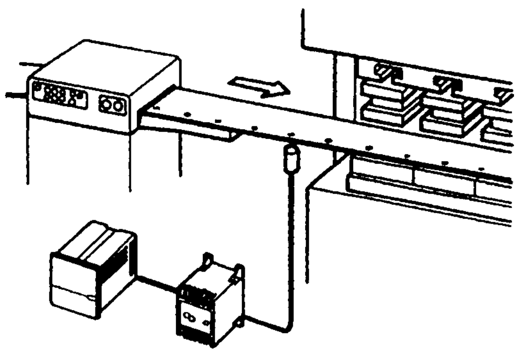


FIGURE 1.122 Detection of misfeeding in high-speed press.

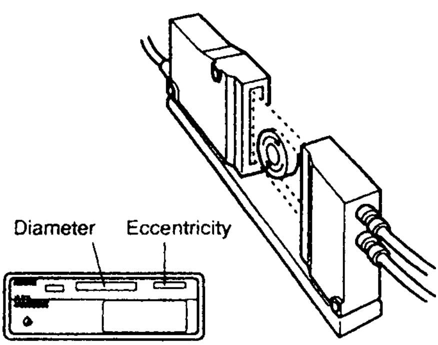


FIGURE 1.123 Simultaneous measurement of outer diameter and eccentricity of ferrite core.

- *Simultaneous measurement of outer diameter and eccentricity of ferrite core.* Simultaneously measures the outer diameter and eccentricity of a ferrite core with a single sensor system. The two measured values can then be simultaneously displayed on a single controller unit (Fig. 1.123).
- *Confirmation of roller centering.* The analog output of the inductive displacement sensor is displayed as a digital value, thus allowing a numerical reading of the shaft position (Fig. 1.124).
- *Measurement of the height and inner diameter of sintered metal ring.* Determines the height and inner diameter of the metal ring by measuring the interrupted areas of the parallel laser beam (Fig. 1.125).
- *Measurement of outer diameter of wire in two axes.* Simultaneously measures in *x* axis to determine the average value of the outer diameter, thereby increasing dimensional stability (Fig. 1.126).

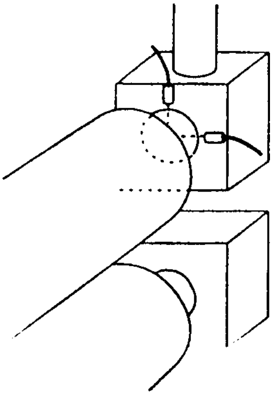


FIGURE 1.124 Confirmation of roller centering.

- *Measurement of outer diameter after centerless grinding.* The scanning head allows continuous noncontact measurement of a metal shaft immediately after the grinding process (Fig. 1.127).

Food processing and packaging:

- *Detection of material caught during heat sealing.* Detects material caught in the distance between rollers (Fig. 1.128).
- *Detection of missing or doubled packing ring in cap.* Detects a missing or doubled rubber ring in caps by using the comparator to evaluate sensor signals (Fig. 1.129).
- *Detection of incorrectly positioned small objects.* The transmitter and the receiver are installed to allow a parallel light beam to scan slightly above the tablet sheets. When a single tablet projects from a line of tablets, the optical axis is interrupted and the light quantity changes (Fig. 1.130).

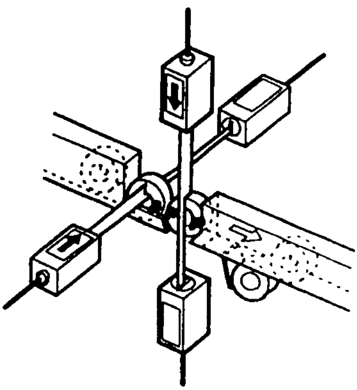


FIGURE 1.125 Measurement of the height and inner diameter of sintered metal ring.

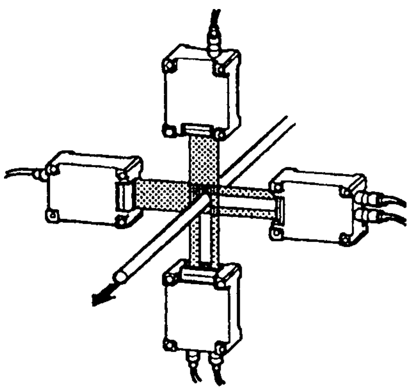


FIGURE 1.126 Measurement of outer diameter of wire in two axes.

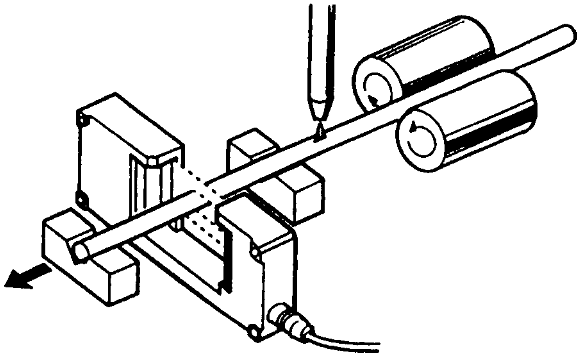


FIGURE 1.127 Measurement of outer diameter after centerless grinding.

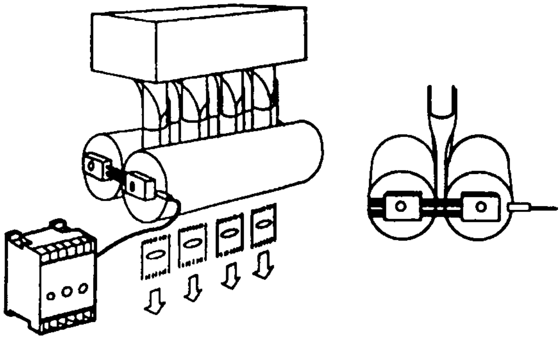


FIGURE 1.128 Detection of material caught during heat sealing.

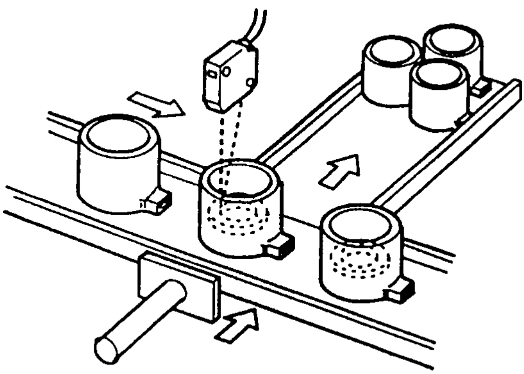


FIGURE 1.129 Detection of missing or doubled packing ring in cap.

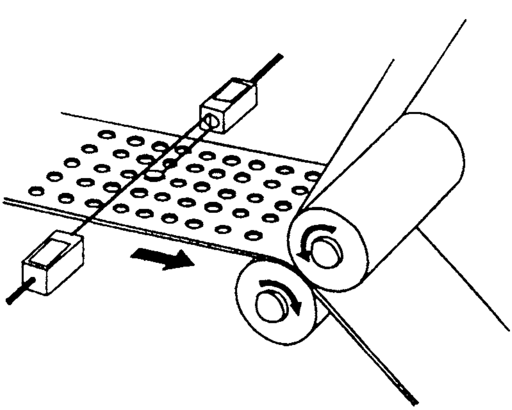


FIGURE 1.130 Detection of incorrectly positioned small objects.

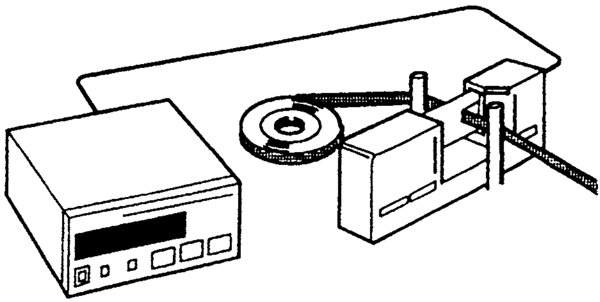


FIGURE 1.131 Measurement of tape width.

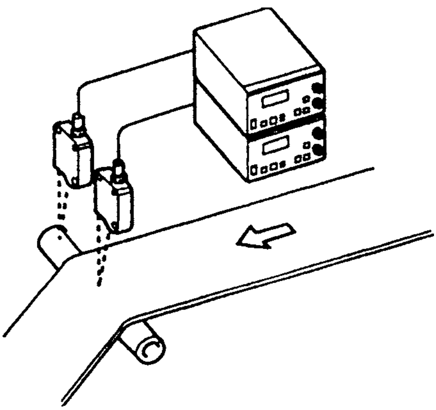


FIGURE 1.132 Measurement of sheet thickness.

- *Measurement of tape width.* Measures the width of running tape to the submicrometer level; 100 percent inspection improves product quality (Fig. 1.131).
- *Measurement of sheet thickness.* Use of two controllers enables thickness measurement by determining the distance in input values. Thus, thickness measurement is not affected by roller eccentricity (Fig. 1.132).

Automatic machinery:

- *Detection of surface run-out caused by clamped error.* Improper clamping due to trapped chips will change the rotational speed of the workpiece. The multifractional digital meter relay sensor calculates the surface run-out of the rotating workpiece, compares it with the stored value, and outputs a detection signal (Fig. 1.133).
- *Detection of residual resin in injection molding machine.* When the sensor heads are such that the optical axis covers the surface of the die, any residual resin will interface with this axis (Fig. 1.134).

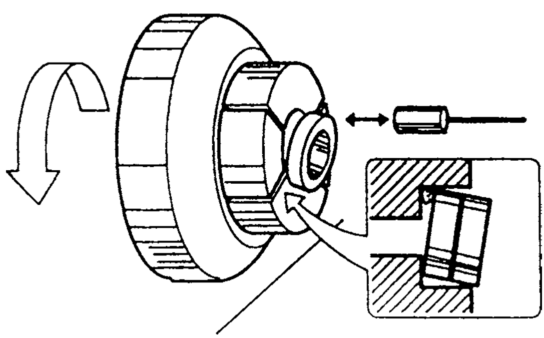


FIGURE 1.133 Detection of surface run-out caused by clamped error.

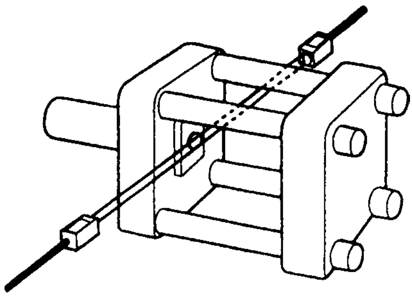


FIGURE 1.134 Detection of residual resin.

- *Measurement of travel of camera lens.* A separate sensor can be installed without interfering with the camera body, thus assuring a highly reliable reading of lens travel (Fig. 1.135).
- *Measurement of rubber sheet thickness.* With the segment function that allows the selection of measuring points, the thickness of a rubber sheet (i.e., the distance between the rollers) can be easily measured (Fig. 1.136).
- *Measurement of stroke of precision table.* Detects even minute strokes at a resolution of 0.05 μm. In addition, the AUTO ZERO function allows indication of relative movement (Fig. 1.137).
- *Measurement of plasterboard thickness.* A sensor head is placed above and below the plasterboard, and its analog outputs are fed into a digital meter relay. The meter relay indicates the absolute thickness value (Fig. 1.138).

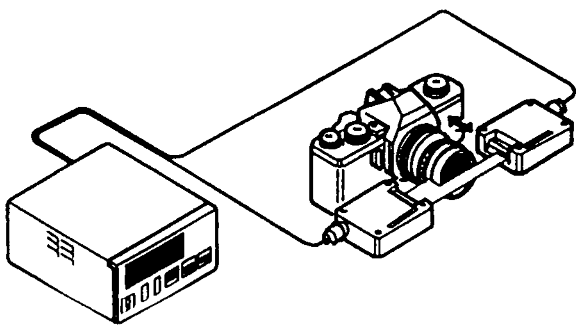


FIGURE 1.135 Measurement of travel of camera lens.

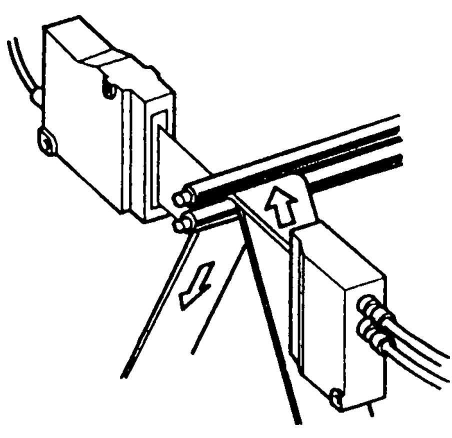


FIGURE 1.136 Measurement of rubber sheet thickness.

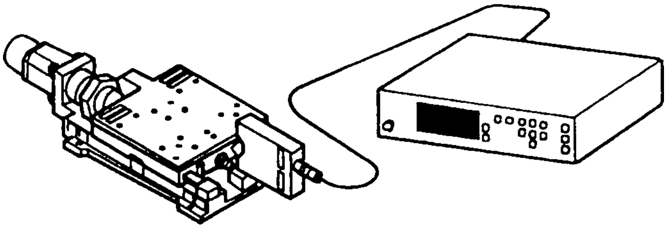


FIGURE 1.137 Measurement of stroke of precision table.

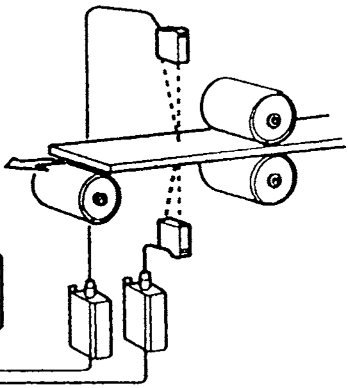


FIGURE 1.138 Measurement of plasterboard thickness.

REFERENCES

1. Chappel, A. (ed.), *Optoelectronics: Theory and Practice*, McGraw-Hill, New York, 1978.
2. Doebelin, E. O., *Measurement Systems: Application and Design*, 4th ed., McGraw-Hill, New York, 1990.
3. Holliday, D., and R. Resnick, *Physics*, Wiley, New York, 1975.
4. International Organization for Standardization, "Statistical Interpretation of Data: Comparison of Two Means in the Case of Paird Observations," ISO 3301–1975.
5. Lion, K. L., *Elements of Electrical and Electronic Instrumentation*, McGraw-Hill, New York, 1975.
6. Neubert, H. K. P., *Instrument Transducers*, 2d ed., Clarendon Press, Oxford, 1975.
7. Ogata, K., *Modern Control Engineering*, 2d ed., Prentice-Hall, Englewood Cliffs, N.J., 1990.
8. Rock, I., *Lightness Constancy, Perception*, W. H. Freeman, New York, 1984.
9. Seippel, R. G., *Optoelectronics*, Reston Publishing Co., Reston, Va., 1981.
10. Shortley, G., and D. Williams, *Quantum Property of Radiation*, Prentice-Hall, Englewood Cliffs, N.J., 1971.
11. Todd, C. D. (Bourns Inc.), *The Potentiometer Handbook*, McGraw-Hill, New York, 1975.
12. White, R. M., "A Sensor Classification Scheme," *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, March, 1987.

CHAPTER 2
FIBER OPTICS IN SENSORS AND CONTROL SYSTEMS

INTRODUCTION

Accurate position sensing is crucial to automated motion control systems in manufacturing. The most common components used for position sensing are photoelectric sensors, inductive proximity sensors, and limit switches. They offer a variety of options for manufacturing implementation, from which highly accurate and reliable systems can be created. Each option has its features, strengths, and weaknesses that manufacturing personnel should understand for proper application. Three types of sensors are used in manufacturing applications:

- *Photoelectric sensors.* Long-distance detection
- *Inductive proximity sensors.* Noncontact metal detection
- *Limit switches.* Detection with traditional reliability

PHOTOELECTRIC SENSORS—LONG-DISTANCE DETECTION

A photoelectric sensor is a switch that is turned on and off by the presence or absence of receiving light (Fig. 2.1). The basic components of a photoelectric sensor are a power supply, a light source, a photodetector, and an output device. The key is the photodetector, which is made of silicon, a semiconductor material that conducts current in the presence of light. This property is used to control a variety of output devices vital for manufacturing operation and control, such as mechanical relays, triacs, and transistors, which in turn control machinery.

Early industrial photoelectric controls used focused light from incandescent bulbs to activate a cadmium sulfide photocell (Fig. 2.2). Since they were not modulated, ambient light such as that from arc welders, sunlight, or fluorescent light fixtures could easily false-trigger these devices. Also, the delicate filament in the incandescent bulbs had a relatively short life span, and did not hold up well under high vibration and the kind of shock loads normally found in an industrial environment. Switching speed was also limited by the slow response of the photocell to light/dark changes (Fig. 2.1).

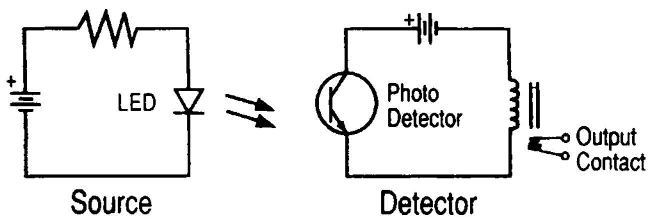


FIGURE 2.1 Photoelectric sensor.

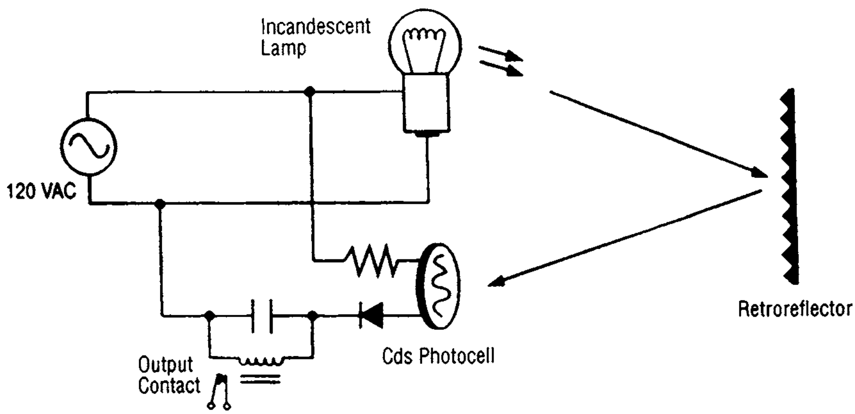


FIGURE 2.2 Early photoelectric control.

Light-Emitting Diodes

Photoelectric sensors use an effective light source, light-emitting diodes (LEDs), which were developed in the early 1960s. LEDs are solid-state devices that emit light when current is applied (Fig. 2.3). This is the exact opposite of the photodetector, which emits current when light is received.

LEDs have several advantages over incandescent bulbs and other light sources. LEDs can be turned on and off very rapidly, are extremely small, consume little power, and last as long as 100,000 continuous hours. Also, since LEDs are solid-state devices, they are much more immune to vibration than incandescent bulbs.

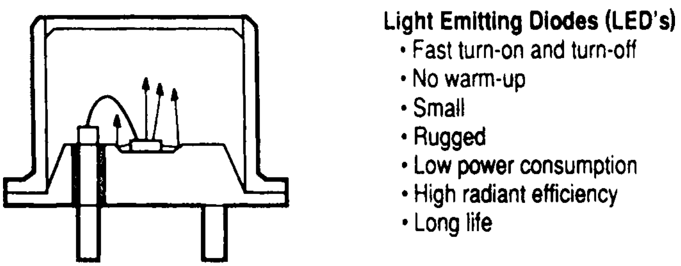


FIGURE 2.3 Light-emitting diode.

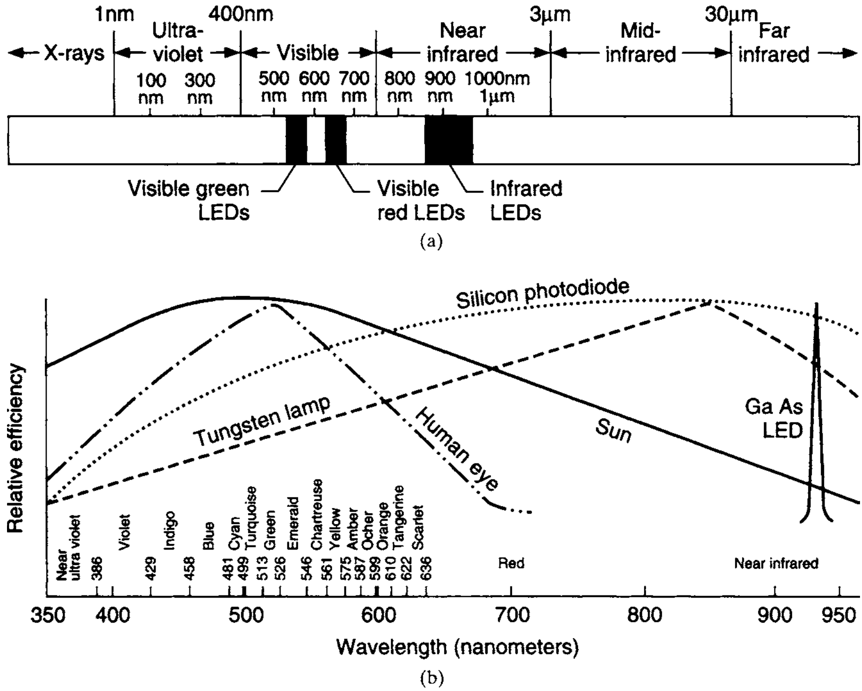


FIGURE 2.4 (*a*) LED emission wavelengths. (*b*) Infrared gallium arsenide LED emission.

LEDs emit light energy over a narrow wavelength (Fig. 2.4*a*). Infrared (ir) gallium arsenide LEDs emit energy only at 940 nm (Fig. 2.4*b*). As this wavelength is at the peak of a silicon photodiode's response, maximum energy transfer between source and detector is achieved.

A silicon photodetector's sensitivity to light energy also peaks in the infrared light spectrum. This contributes to the high efficiency and long range possible when silicon photodetectors are used in conjunction with gallium arsenide LEDs.

In recent years, visible LEDs have been introduced as light sources in photoelectric controls. Because the beam is visible to the naked eye, the principle advantage of visible LEDs is ease of alignment. Visible beam photoelectric controls usually have lower optical performance than ir LEDs.

The modes of detection for optical sensors are (1) through-beam and (2) reflection (diffuse reflection and reflex detection).

Through-Beam Sensors

Through-beam sensors have separate source and detector elements aligned opposite each other, with the beam of light crossing the path that an object must cross (Fig. 2.5). The effective beam area is that of the column of light that travels straight between the lenses (Fig. 2.6).

Because the light from the source is transmitted directly to the photodetector, through-beam sensors offer the following benefits:

- Longest range for sensing
- Highest possible signal strength

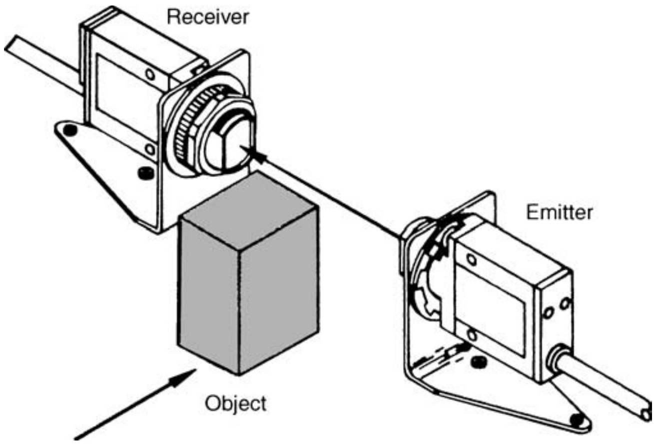


FIGURE 2.5 Through-beam sensor.

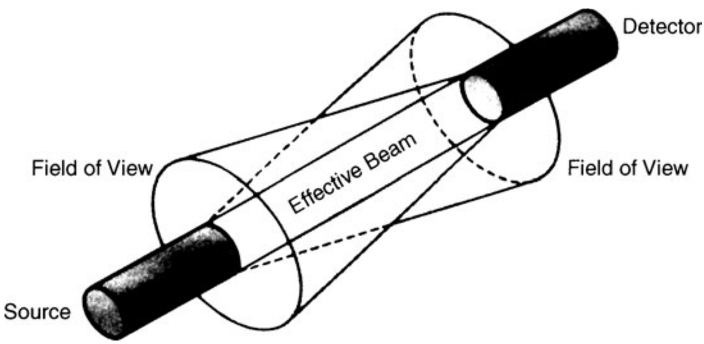


FIGURE 2.6 Effective beam area.

- Greatest light/dark contrast ratio
- Best trip point repeatability

The limitations of through-beam sensors are as follows:

- They require wiring of the two components across the detection zone.
- It may be difficult to align the source and the detector.
- If the object to be detected is smaller than the effective beam diameter, an aperture over the lens may be required (Fig. 2.7).

Reflex Photoelectric Controls

Reflex photoelectric controls (Fig. 2.8) position the source and detector parallel to each other on the same side of the target. The light is directed to a retroreflector and returns to the detector. The switching and output occur when an object breaks the beam.

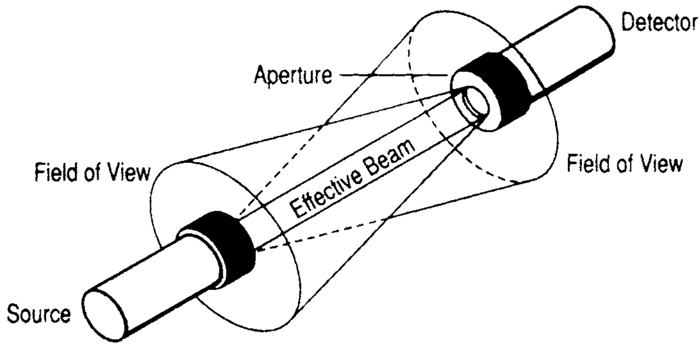


FIGURE 2.7 Sensor with aperture over lens for detecting small objects.

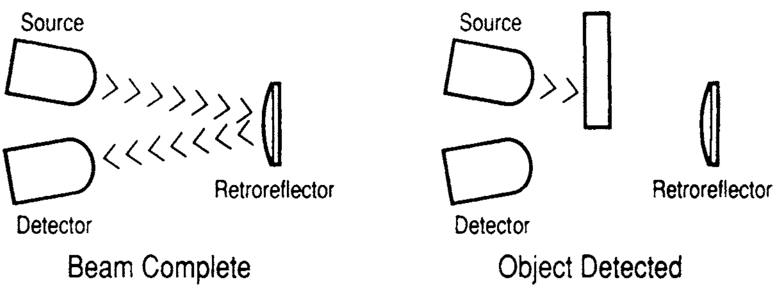


FIGURE 2.8 Reflex photoelectric controls.

Since the light travels in two directions (hence twice the distance), reflex controls will not sense as far as through-beam sensors. However, reflex controls offer a powerful sensing system that is easy to mount and does not require that electrical wire be run on both sides of the sensing area. The main limitation of these sensors is that a shiny surface on the target object can trigger false detection.

Polarized Reflex Detection

Polarized reflection controls use a polarizing filter over the source and detector that conditions the light such that the photoelectric control sees only light returned from the reflector (Fig. 2.9). A polarized reflex sensor is used in applications where shiny surfaces such as metal or shrink-wrapped boxes may false-trigger the control.

Polarized reflex sensing is achieved by combining some unique properties of polarizers and retroreflectors. These properties are (1) polarizers pass light that is aligned along only one plane and (2) corner-cube reflectors depolarize light as it travels through the face of the retroreflector (Fig. 2.10).

Light from the source is aligned by a polarizer. When this light reflects off the retroreflector, it is depolarized. The returning light passes through another polarizing filter in front of the detector. The detector's polarizer is oriented at 90° to the source's polarizer. Only the light that has been rotated by the corner cube retroreflector can pass through the detector's polarizer. Light that bounces off other shiny objects, and has not been rotated 90°, cannot pass through the detector's polarizer, and will not trigger the control.

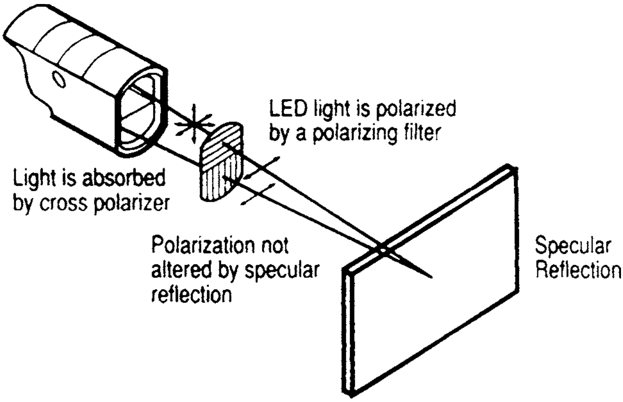


FIGURE 2.9 Polarization reflection controls.

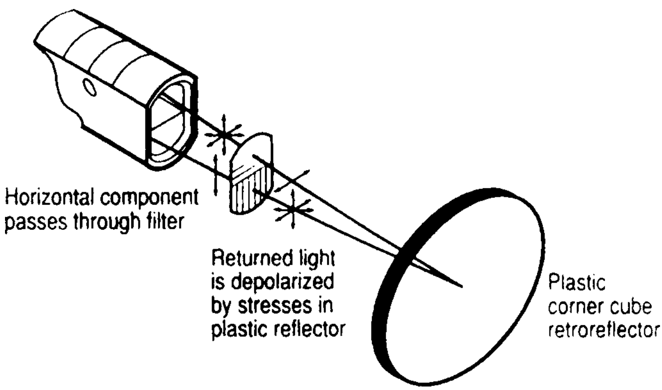


FIGURE 2.10 Corner-cube reflector.

Polarized reflex sensors will not work with reflective tape containing glass beads. Also, shiny objects wrapped with clear plastic shrink-wrap will potentially false-trigger a polarized reflex control, since under certain conditions these act as a corner-cube reflector.

The polarized reflex detection sensor has the following advantages:

- It is not confused by the first surface reflections from target objects.
- It has a high dark/light contrast ratio.
- It is easily installed and aligned. One side of the sensing zone only need be wired.

It also has certain limitations:

- The operating range is half that of a nonpolarized sensor because much of the signal is lost in the polarizing filters.
- The sensor can be fooled by shiny objects wrapped with shrink-wrap material.

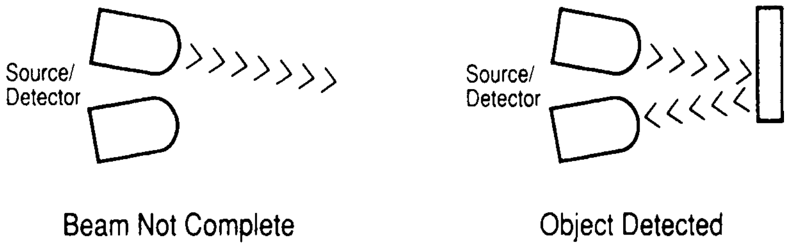


FIGURE 2.11 Proximity detection.

Proximity (Diffuse-Reflection) Detection

Proximity detection is similar to reflex detection, because the light source and detector elements are mounted on the same side (Fig. 2.11). In this application, the sensors detect light that is bounced off the target object, rather than the breaking of the beam. The detection zone is controlled by the type, texture, and composition of the target object's surface.

Focused proximity sensors are a special type of proximity sensor where the source and the detector are focused to a point in front of the sensor (Fig. 2.12). Focused proximity sensors can detect extremely small objects, or look into holes or cavities in special applications. Background objects will not false-trigger a focused proximity sensor since they are "cross-eyed" and cannot see past a certain point.

Advantages of the focused proximity sensor are:

- Installation and alignment are simple. The control circuit can be wired through only one side of the sensing zone.
- It can detect differences in surface reflectivity.

It also has certain limitations:

- It has a limited sensing range.
- The light/dark contrast and sensing range depend on the target object's surface reflectivity.

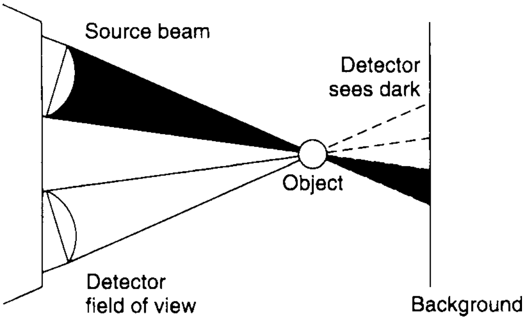


FIGURE 2.12 Focused proximity sensor.

Automated Guided Vehicle System

The wide sensing field of diffuse reflective photoelectric sensors makes them ideally suited for use as obstacle detection sensors. Sensing distance and field width are adjustable to form the obstacle detection zone vital for manufacturing operations.

Two outputs, *near* and *far*, provide switching at two separate sensing distances that are set by corresponding potentiometers. The sensing distance of the far output is adjustable up to 3-m maximum. The sensing distance of the near output is adjustable from 30 to 80 percent of the far output. Indicators include a red LED that glows with the near output ON and a yellow LED that glows with the far output ON.

An ideal application for this family of sensors is the automated guided vehicle (AGV), which requires both slow-down and stop controls to avoid collisions when obstacles enter its path (Fig. 2.13).

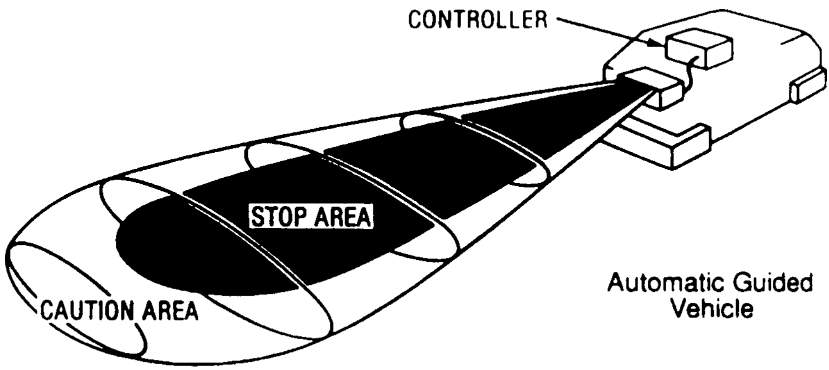


FIGURE 2.13 Application of diffuse reflective photoelectric sensor in automated guided vehicle system.

A modulated infrared light source provides immunity to random operation caused by ambient light. Additionally, unwanted sensor operation caused by adjacent sensor interference (crosstalk) is also eliminated through the use of multiple-position modulated frequency adjustments.

FIBER OPTICS

Fiber optics have greatly expanded the applications of photoelectric sensors. Fiber optics uses bundles of thin plastic or glass fibers that operate on a principle discovered in 1854 by John Tyndahl. When Tyndahl shined a beam of light through a stream of water, instead of emerging straight from the stream of water as might be expected, the light tended to bend with the water as it arced towards the floor. Tyndahl discovered that the light was transmitted along the stream of water. The light rays inside the water bounced off the internal walls of the water and were thereby contained (Fig. 2.14). This principle has come to be known as *total internal reflection*.

Industry has since discovered that the principle of total internal reflection also applies to small-diameter glass and plastic fibers, and this has lead to rapid growth of applications throughout the industry. Because optical fibers are small in diameter and flexible, they can bend and twist into confined places. Also, because they contain no electronics, they can operate in much higher temperatures—as high as 400°F—and in areas of high vibration.

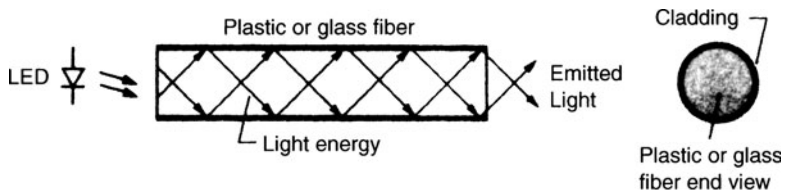


FIGURE 2.14 Total internal reflection.

They are limited by sensing distances, which typically are 80 mm in the proximity mode and 400 mm in the through-beam mode. Also, because of their small sensing area, optical fibers can be fooled by a small drop of water or dirt over the sensing area.

Fiber optics is used to transmit data in the communication field and to transmit images or light in medicine and industry. Photoelectric controls use fiber optics to bend the light from the LED source and return it to the detector so sensors can be placed in locations where common photoelectric sensors could not be applied.

Fiber optics used with photoelectric controls consists of a large number of individual glass or plastic fibers that are sheathed in suitable material for protection. The optical fibers used with photoelectric controls are usually covered by either PVC or stainless-steel jackets. Both protect the fibers from excessive flexing and the environment (Fig. 2.15).

Optical fibers are transparent fibers of glass or plastic used for conducting and guiding light energy. They are used in photoelectric sensors as "light pipes" to conduct sensing light into and out of a sensing area.

Glass optical-fiber assemblies consist of a bundle of 0.05-mm-diameter discrete glass optical fibers housed within a flexible sheath. Glass optical fibers are also able to withstand hostile sensing environments. Plastic optical-fiber assemblies consist of one or two acrylic monofilaments in a flexible sheath.

Two basic styles of fiber-optic assemblies are available: (1) individual fiber optics (Fig. 2.16) and (2) bifurcated fiber optics (Fig. 2.17).

Individual fiber-optic assemblies guide light from an emitter to a sensing location, or to a receiver from a sensing location. Bifurcated fibers use half their fiber area to transmit light and the other half to receive light.

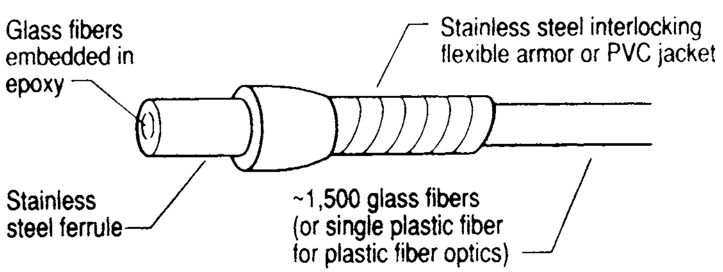


FIGURE 2.15 Jacketed glass fibers.

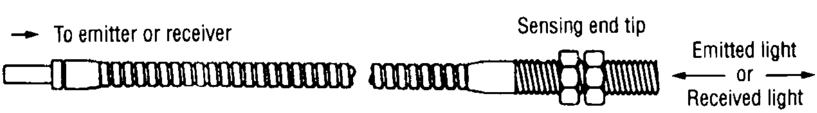


FIGURE 2.16 Individual fiber-optic assembly.

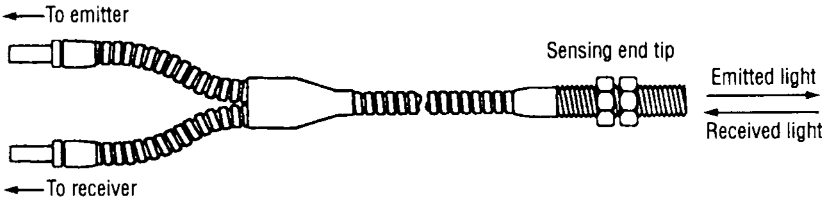


FIGURE 2.17 Bifurcated fiber-optic assembly.

Individual Fiber Optics

A fiber-optic assembly having one control end and one sensing end is used for piping photoelectric light from an emitter to the sensing location or from the sensing location back to a receiver. It is usually used in pairs in the opposed sensing mode, but can also be used side by side in the diffuse proximity mode or angled for the specular reflection or mechanical convergent mode.

Bifurcated Fiber Optics

A bifurcated fiber-optic assembly is branched to combine emitted light with received light in the same assembly. Bifurcated fibers are used for diffused (divergent) proximity sensing, or they may be equipped with a lens for use in the retroreflective mode.

Three types of sensing modes are used in the positioning of a sensor so the maximum amount of emitted energy reaches the receiver sensing element:

- Opposed sensing mode (Fig. 2.18)
- Retroreflective sensing mode (Fig. 2.19)
- Proximity (diffused) sensing mode (Fig. 2.20)

Opposed sensing is the most efficient photoelectric sensing mode and offers the highest level of optical energy to overcome lens contamination, sensor misalignment, and long scanning ranges. It is also often referred to as *direct scanning* and sometimes called the *beam break* mode.

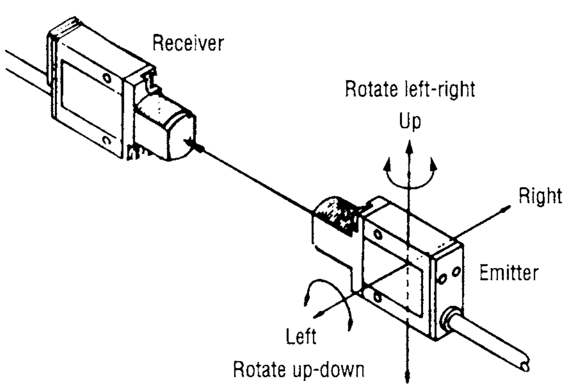


FIGURE 2.18 Opposed sensing mode. For alignment, move emitter or receiver up-down, left-right, and rotate.

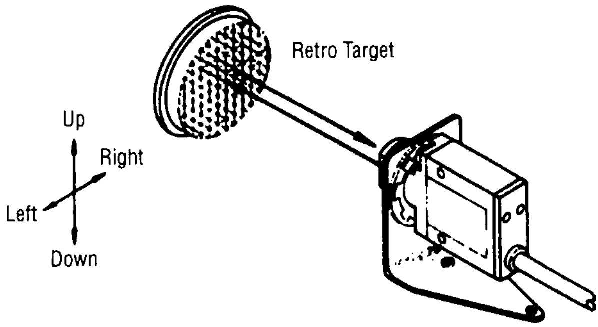


FIGURE 2.19 Retroreflective sensing mode. For alignment, move target up-down, left-right.

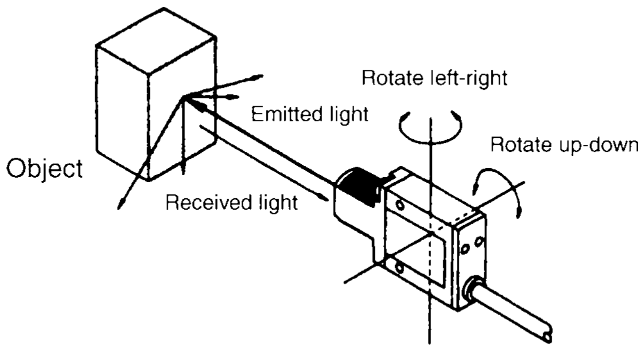


FIGURE 2.20 Proximity sensing mode. For alignment, rotate up-down, left-right.

The addition of fiber optics to photoelectric sensing has greatly expanded the application of photoelectric devices. Because they are small in diameter and flexible, optical fibers can bend and twist into tiny places formerly inaccessible to bulky electronic devices.

Optical fibers operate in the same sensing modes as standard photoelectric controls—through-beam, proximity, and reflex. The sizes and shapes of sensing tips have been developed to accommodate many applications.

Optical fibers have a few drawbacks:

- They have a limited sensing distance. Typical sensing distance in the proximity mode is 80 mm; 380 mm for the through-beam mode.
- They are typically more expensive than other photoelectric sensing controls.
- They are easily fooled by a small drop of water or dirt over the sensing surface.

Optical fibers' advantages include:

- Sensing in confined places
- The ability to bend around corners
- No electronics at the sensing point

- Operation at high temperatures (glass)
- Total immunity from electrical noise and interference
- Easily cut to desired lengths (plastic)

OPTICAL FIBER PARAMETERS

The most important parameters affecting optical-fiber performance are:

- Excess gain
- Background suppression
- Contrast
- Polarization

Excess Gain

Excess gain is the measure of energy available between the source and the detector to overcome signal loss due to dirt or contamination. Excess gain is the single most important consideration in choosing a photoelectric control in manufacturing. It is the extra punch that the sensor has available within its detecting region.

By definition, excess gain is the ratio of the amount of light the detector sees to the minimum amount of light required to trip the sensor. This ratio is depicted graphically for all photoelectric sensors. In Fig. 2.21, excess gain is plotted along the vertical logarithmic

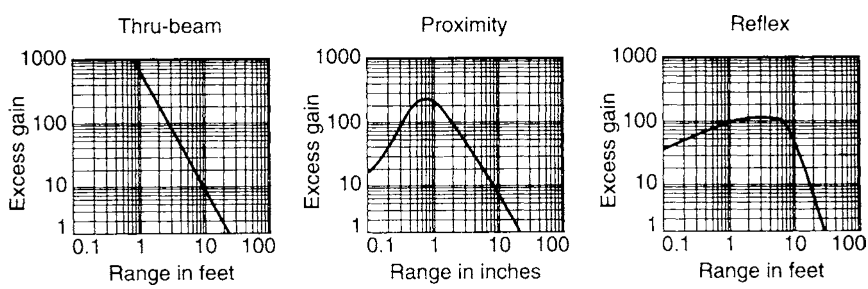


FIGURE 2.21 Excess gain curves.

axis, starting at 1 (the minimum amount of light required to trigger the detector). Every point above 1 represents the amount of light above that required to trigger the photoelectric control—the excess gain.

Often, the standard of comparison for choosing between different photoelectric sensors is range. Actually, more important to most applications is the excess gain. For a typical application, the higher the excess gain within the sensing region, the more likely the application will work. It is the extra margin that will determine whether the photoelectric control will continue to operate despite the buildup of dirt on the lens or the presence of contamination in the air.

Example. An application requires detecting boxes on a conveyer in a filthy industrial environment (Fig. 2.22). The boxes will pass about 2 to 5 mm from the sensor as they move

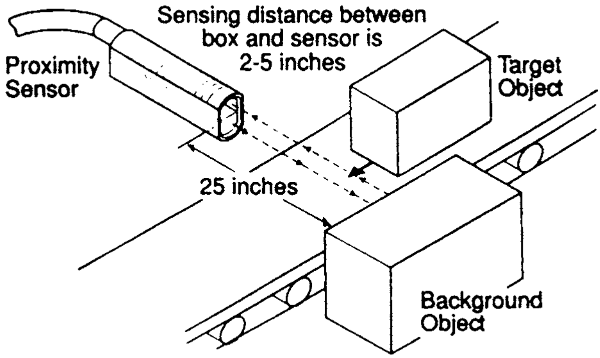


FIGURE 2.22 Box detection.

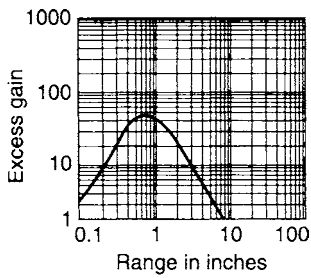


FIGURE 2.23 Excess gain curve for sensor 1.

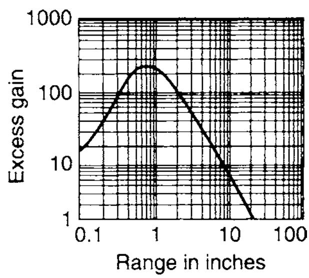


FIGURE 2.24 Excess gain curve for sensor 2.

along the conveyer at the sensing location. Given a choice between the two proximity sensors whose excess gain curves appear in Figs. 2.23 and 2.24, which photoelectric control should be selected for this application?

If the decision were based solely on specified range, the unit described in Fig. 2.23 would be selected. However, if units were installed in this application, it might fail after a short time in operation. Over time, contaminants from the environment would settle on the lens, decreasing the amount of light the sensor sees. Eventually, enough lens contamination would accumulate that the photoelectric control would not have enough excess gain to overcome the signal loss created by the coating, and the application would fail.

A better choice for this application would be the unit represented in Fig. 2.24. It delivers much more excess gain in the operating region required for this application and will therefore work much more successfully than the other unit.

Background Suppression

Background suppression enables a diffuse photoelectric sensor to have high excess gain to a predetermined limit and insufficient excess gain beyond that range, where it might pick up objects in motion and yield a false detection. By using triangular ranging, sensor developers have created a sensor that emits light that reflects on the detector from two different target positions. The signal received from the more distant target is subtracted from that of the closer target, providing high excess gain for the closer target.

Contrast

Contrast measures the ability of a photoelectric control to detect an object; it is the ratio of the excess gain under illumination to the excess gain in the dark. All other things being equal, the sensor that provides the greatest contrast ratio should be selected. For reliable operation, a ratio of 10:1 is recommended.

Polarization

Polarization is used in reflection sensors in applications where shiny surfaces, such as metal or shrink-wrapped boxes, may trigger the control falsely. The polarizer passes light along only one plane (Fig. 2.25), and the corner-cube reflectors depolarize the light as it passes through the plastic face of the retroreflector (Fig. 2.10). Only light that has been rotated by the corner-cube retroreflector can pass through the polarizer, whereas light that bounces off other shiny objects cannot.

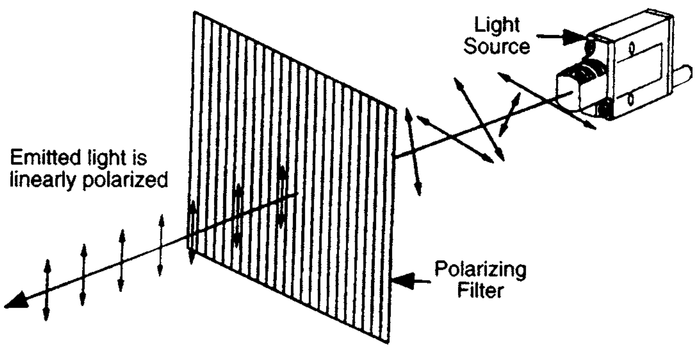


FIGURE 2.25 Polarization.

Like regular reflex photoelectric sensors, polarized sensors have a high light/dark contrast ratio and are simple to install and align. However, the polarizers do limit the sensor's operating range because light is lost passing through them.

INDUCTIVE PROXIMITY SENSORS—NONCONTACT METAL DETECTION

Inductive proximity sensors are another common choice for position sensing. An inductive proximity sensor consists of four basic elements:

- The sensor, which comprises a coil and ferrous core
- An oscillator circuit
- A detector circuit
- A solid-state output

In the circuitry in Fig. 2.26, the oscillator generates an electromagnetic field that radiates from the sensor's face. This field is centered around the axis and detected by the ferrite

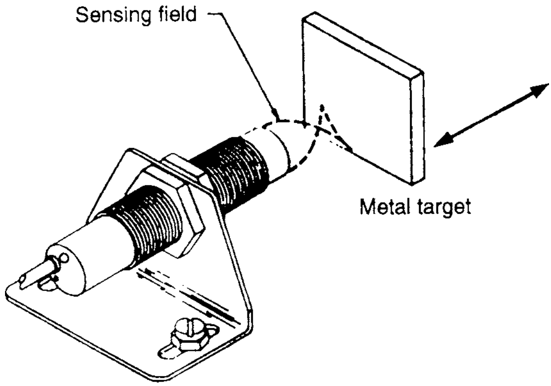


FIGURE 2.26 Inductive proximity sensor.

core to the front of the sensor assembly. When a metal object enters the electromagnetic field, eddy currents are induced in the surface of the target. This loads in the oscillator circuit, which reduces its amplitude.

The detector circuit detects the change in the oscillator amplitude and, depending on its programming, switches ON and OFF at a specific oscillator amplitude. The sensing circuit returns to its normal state when the target leaves the sensing area and the oscillator circuit regenerates.

The nominal sensing range of inductive proximity sensors is a function of the diameter of the sensor and the power available to generate the electromagnetic field. This is subject to a manufacturing tolerance of ±10 percent, as well as a temperature drift tolerance of ±10 percent. The target size, shape, and material will have an effect on the sensing range. Smaller targets will reduce the sensing range, as will targets that are not flat or are made of nonferrous material.

Basically two types of inductive proximity sensors are used: (1) shielded and (2) nonshielded. The shielded version has a metal cover around the ferrite core and coil assembly. This focuses the electromagnetic field to the front of the sensor and allows it to be imbedded in metal without influencing the sensing range. The nonshielded sensor can sense on the side as well as in front of a sensor. It requires a nonmetallic area around the sensor to operate correctly.

Inductive proximity sensors have several benefits:

- High repeatability. Visibility of the environment is not an issue since inductive proximity sensors can sense only electromagnetic fields. Therefore, environments from dirt to sunlight pose no problem for inductive proximity sensors. Also, because they are noncontact sensors, nothing wears.
- Shock and vibration resistance.
- Versatile connections. They can connect two or three wires with an AC, AC/DC, or DC power supply and up to four-wire DC connections.
- Wide operating temperature range. They operate between −20 to +70°C, ±10 percent.
- Very fast response, particularly in the DC models. These sensors can detect presence, send a signal, and reset in 50 μs (2000 times per second) in DC models.

Inductive proximity sensors are generally limited by their sensing distances and the material they sense. The effective range is limited to about 25 mm for most models and can be extended to only about 100 mm with the large models.

LIMIT SWITCHES—TRADITIONAL RELIABILITY

Limit switches are mechanical position-sensing devices that offer simplicity, robustness, and repeatability to processes. Mechanical limit switches are the oldest and simplest of all presence- or position-sensing devices: contact is made and a switch is engaged. This simplicity contributes generally to the cost advantage of limit switches. Yet, they can provide the control capabilities and versatility demanded in today's error-free manufacturing environment. The key to their versatility is the various forms they can take in the switch, actuating head, and lever operator. Two-step, dual-pole limit switches exist that can detect and count two products of different sizes and can provide direct power control to segregate or process the items differently. The lever operator will rotate 10° to activate one set of contacts and 20° to activate another set. Because of the high amperage they can handle, limit switches can control up to ten contacts from the movement of a single lever.

They are easy to maintain because the operator can hear the operation of the switch and can align it easily to fit the application. They are also robust. Limit switches are capable of handling an inrush current ten times that of their steady-state current rating. They have rugged enclosures and have prewiring that uses suitable strain-relief bushings to enable the limit switch to retain cables with 500 to 600 pounds of force on them. Limit switches can also handle direct medium-power switching for varying power factors and inrush stresses. For example, they can control a multihorsepower motor without any interposing starter, relay, or contactor.

Reliability is another benefit. Published claims for repeat accuracy for standard limit switches vary from within 0.03 mm to within 0.001 mm over the temperature range of –4 to +200°F. Limit switches dissipate energy spikes and rarely break down under normal mode surges. They will not be affected by electromagnetic interferences (EMIs); there are no premature responses in the face of EMI. However, because they are mechanical devices, limit switches face physical limitations that can shorten their service life even though they are capable of several million operations. Also, heavy sludge, chips, or coolant can interfere with their operation.

FACTORS AFFECTING THE SELECTION OF POSITION SENSORS

In selecting a position sensor, several key factors should be considered:

- *Cost.* Both initial purchase price and life-cycle cost must be considered.
- *Sensing distance.* Photoelectric sensors are often the best selection when sensing distances are longer than 25 mm. Photoelectric sensors can have sensing ranges as long as 300,000 mm for outdoor or extremely dirty applications, down to 25 mm for extremely small parts or for ignoring background. Inductive proximity sensors and limit switches, on the other hand, have short sensing distances. The inductive proximity sensors are limited by the distance of the electromagnetic field—less than 25 mm for most models—and limit switches can sense only as far as the lever operator reaches.
- *Type of material.* Inductive proximity sensors can sense only ferrous and nonferrous materials, whereas photoelectric and limit switches can detect the presence of any solid material. Photoelectric sensors, however, may require a polarizer if the target's surface is shiny.
- *Speed.* Electronic devices using DC power are the fastest—as fast as 2000 cycles per second for inductive proximity models. The fastest-acting limit switches can sense and reset in 4 ms or about 300 times per second.

- *Environment.* Proximity sensors can best handle dirty, gritty environments, but they can be fooled by metal chips and other metallic debris. Photoelectric sensors will also be fooled or left inoperable if they are fogged or blinded by debris.
- *Types of voltages, connections, and requirements of the device's housing.* All three types can accommodate varying requirements, but the proper selection must be made in light of the power supplies, wiring schemes, and environments.
- *Third-party certification.* The Underwriters Laboratories (UL), National Electrical Manufacturers Association (NEMA), International Electrotechnical Commission (IEC), Factory Mutual, Canadian Standards Association (CSA), and other organizations impose requirements for safety, often based on the type of application. The certification will ensure the device has been tested and approved for certain uses.
- *Intangibles.* These can include the availability of application support and service, the supplier's reputation, local availability, and quality testing statements from the manufacturer.

WAVELENGTHS OF COMMONLY USED LIGHT-EMITTING DIODES

An LED is a semiconductor that emits a small amount of light when current flows through it in the forward direction. In most photoelectric sensors, LEDs are used both as emitters for sensing beams and as visual indicators of alignment or output status for a manufacturing process. Most sensor manufacturers use visible red, visible green, or infrared (invisible) LEDs (Fig. 2.4*b*). This simple device plays a significant part in industrial automation. It provides instantaneous information regarding an object during the manufacturing operation. LEDs, together with fiber optics, allow a controller to direct a multitude of tasks, simultaneously or sequentially.

SENSOR ALIGNMENT TECHNIQUES

A sensor should be positioned so the maximum amount of emitted energy reaches the receiver element in one of three different modes:

- Opposed sensing mode
- Retroreflective sensing mode
- Proximity (diffuse) sensing mode

Opposed Sensing Mode

In this photoelectric sensing mode, the emitter and receiver are positioned opposite each other so the light from the emitter shines directly at the receiver. An object then breaks the light beam that is established between the two. Opposed sensing is always the most reliable mode.

Retroreflective Sensing Mode

Retroreflective sensing is also called the *reflex* mode or simply the *retro* mode. A retroreflective photoelectric sensor contains both emitter and receiver. A light beam is established between the sensor and a special retroreflective target. As in opposed sensing, an object is sensed when it interrupts this beam.

Retro is the most popular mode for conveyer applications where the objects are large (boxes, cartons, etc.), where the sensing environment is relatively clean, and where scanning ranges are typically a few meters in length. Retro is also used for code-reading applications. Automatic storage and retrieval systems and automatic conveyer routing systems use retroreflective code plates to identify locations and/or products.

Proximity (Diffuse) Sensing Mode

In the proximity (diffuse) sensing mode, light from the emitter strikes a surface of an object at some arbitrary angle and is diffused from the surface at all angles. The object is detected when the receiver captures some small percentage of the diffused light. Also called the *direct reflection* mode or simply the photoelectric *proximity* mode, this method provides direct sensing of an object by its presence in front of a sensor. A variation is the ultrasonic proximity sensor, in which an object is sensed when its surface reflects a sound wave back to an acoustic sensor.

Divergent Sensing Mode

This is a variation of the diffuse photoelectric sensing mode in which the emitted beam and the receiver's field of view are both very wide. Divergent mode sensors (Fig. 2.27) have

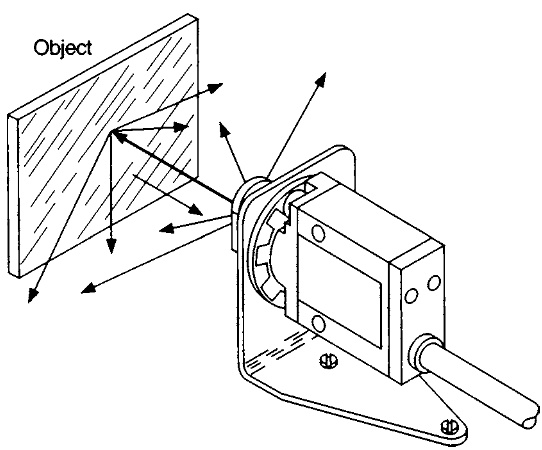


FIGURE 2.27 Divergent sensing mode.

loose alignment requirements, but have a shorter sensing range than diffuse mode sensors of the same basic design. Divergent sensors are particularly useful for sensing transparent or translucent materials or for sensing objects with irregular surfaces (e.g., webs that flutter). They are also used effectively to sense objects with very small profiles, such as small-diameter thread or wire, at close range.

All unlensed bifurcated optical fibers are divergent. The divergent mode is sometimes called the *wide-beam diffuse* (or *proximity*) mode.

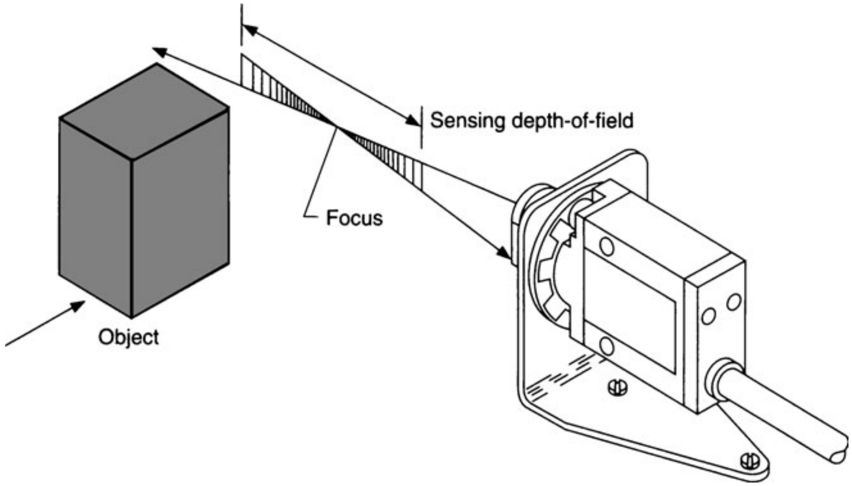


FIGURE 2.28 Convergent sensing mode.

Convergent Sensing Mode

This is a special variation of diffuse mode photoelectric proximity sensing that uses additional optics to create a small, intense, and well-defined image at a fixed distance from the front surface of the sensor lens (Fig. 2.28). Convergent beam sensing is the first choice for photoelectric sensing of transparent materials that remain within a sensor's depth of field. It is also called the *fixed-focus proximity* mode.

Mechanical Convergence

In mechanical convergence (Fig. 2.29), an emitter and a receiver are simply angled toward a common point ahead of the sensor. Although less precise than the optical convergent-beam sensing mode, this approach to reflective sensing uses light more

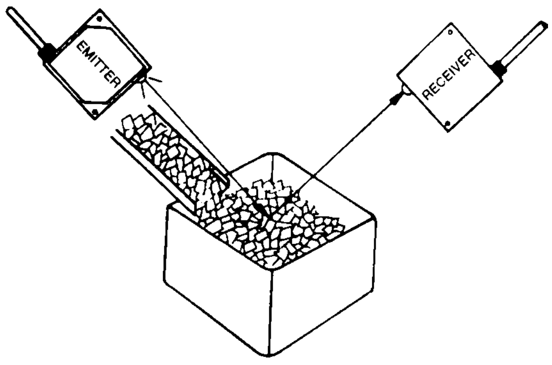


FIGURE 2.29 Mechanical convergence.

efficiently than diffuse sensing and gives a greater depth of field than true optical convergence.

Mechanical convergence may be customized for an application by mounting the emitter and the receiver to converge at the desired distance. Depth of field is controlled by adjusting the angle between the emitter and the receiver.

FIBER OPTICS IN INDUSTRIAL COMMUNICATION AND CONTROL

The application of fiber optics to industrial information transfer is a natural extension of the current commercial uses of this technology in high-data-rate communications. While the primary advantage of fiber optics in traditional application areas has been extremely reliable communication at high rates, exceeding 1 Gbit/s over distances exceeding 100 km, other intrinsic features of the technology are more important than data rate and distance capability in industrial uses.

The physical mechanism of light propagating through a glass fiber has significant advantages that enable sensors to carry data and plant communications successfully and in a timely manner—a fundamental condition that must be constantly maintained in a computer-integrated manufacturing environment. These advantages include:

- The light signal is completely undisturbed by electrical noise. This means that the fiber-optic cables can be laid wherever convenient without special shielding. Fiber-optic cables and sensors are unaffected by electrical noise when placed near arc welders, rotating machinery, electrical generators, and so on, whereas in similar wired applications, even the best conventional shielding methods are often inadequate.
- Fiber-optic communication is devoid of any electrical arcing or sparking, and thus can be used successfully in hazardous areas without danger of causing an explosion.
- The use of a fiber link provides total electrical isolation between terminal points on the link. Over long plant distances, this can avoid troublesome voltage or ground differentials and ground loops.
- A fiber-optic system can be flexibly configured to provide additional utility in existing hardware.

PRINCIPLES OF FIBER OPTICS IN COMMUNICATIONS

An optical fiber (Fig. 2.30) is a thin strand composed of two layers: an inner core and an outer cladding. The core is usually constructed of glass, and the cladding structure, of glass or plastic. Each layer has a different index of refraction, the core being higher. The difference between the index of refraction of the core material, n_1, and that of the surrounding cladding material, n_2, causes rays of light injected into the core to continuously reflect back into the core as they propagate down the fiber.

The light-gathering capability of the fiber is expressed in terms of its numerical aperture (NA), the sine of the half angle of the acceptance cone for that fiber. Simply stated, the larger the NA, the easier it is for the fiber to accept light (Fig. 2.31).

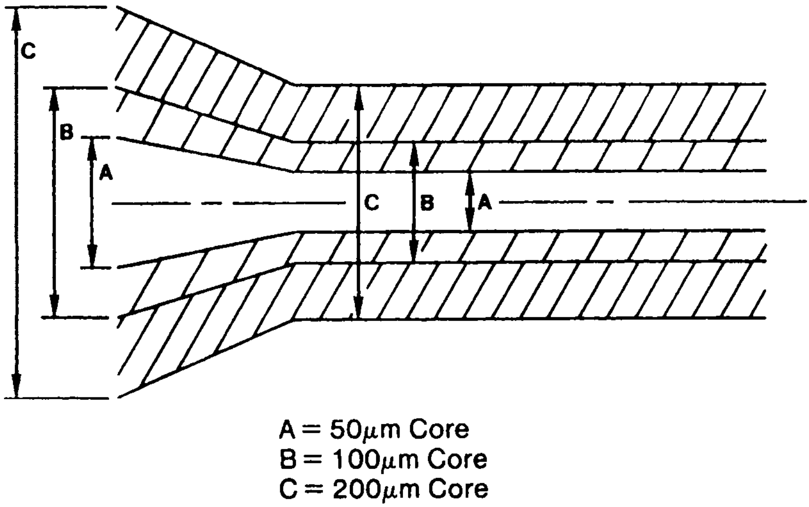


FIGURE 2.30 Structure of optical fiber.

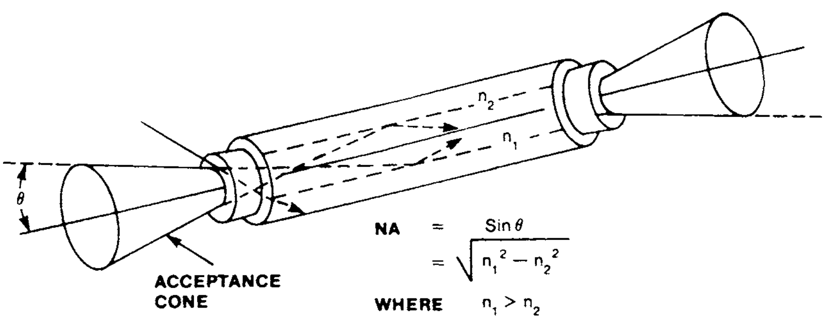


FIGURE 2.31 Numerical aperture (NA) of an optical fiber.

FIBER-OPTIC INFORMATION LINK

A fiber-optic communication system (Fig. 2.32) consists of:

- A light source (LED or laser diode) pulsed by interface circuitry and capable of handling data rates and voltage levels of a given magnitude.
- A detector (photodiode) that converts light signals to electrical signals and feeds interface circuitry to re-create the original electrical signal.
- Fiber-optic cables between the light source and the detector (called the *transmitter* and the *receiver*, respectively).

It is usual practice for two-way communication to build a transmitter and receiver into one module and use a duplex cable to communicate to an identical module at the end of the

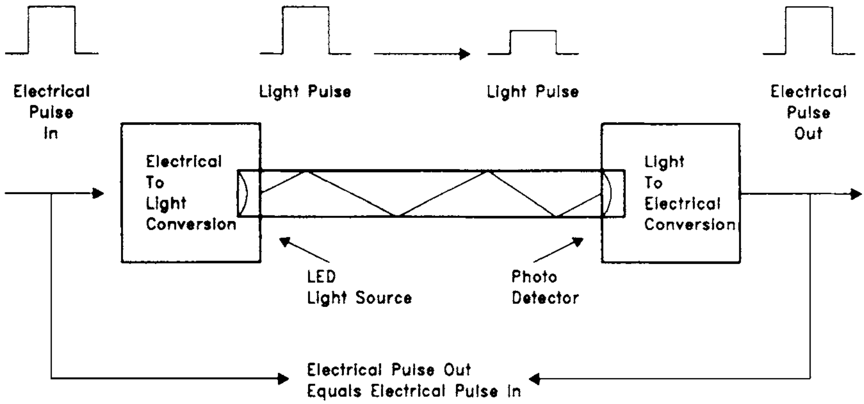


FIGURE 2.32 Fiber-optic communication system.

link. The 820- to 850-nm range is the most frequent for low-data-rate communication, but other wavelengths (1300 and 1550 nm) are used in long-distance systems. Fiber selection must always take transmission wavelength into consideration.

CONFIGURATIONS OF FIBER OPTICS

The selection of optical fiber plays a significant part in sensor performance and information flow. Increased fiber diameter results in higher delivered power, but supports a lower bandwidth. A fiber with a 200-µm core diameter (glass core, silicone plastic cladding), by virtue of its larger diameter and acceptance angle, can transmit five to seven times more light than a 100-µm core fiber, or up to 30 times more light than a 40-µm fiber, the historical telecommunication industry standard. Computer data communications systems commonly employ 62.5-µm fiber because of its ability to support very high data rates (up to 100 Mbaud) while offering increased power and ease of handling. Factory bandwidth requirements for most links are typically one or more orders of magnitude less. Therefore, fiber core size may be increased to 200 µm to gain the benefits of enhanced power and handling ease, the decrease in bandwidth being of no consequence (Fig. 2.31).

Optical Power Budget

An optical power budget examines the available optical power and how it is used and dissipated in a fiber-optic system. It is important to employ the highest possible optical power budget for maximum power margin over the detector requirement. A budget involves four factors:

- Types of light source
- Optical fiber acceptance cone
- Receiver sensitivity
- Splice, coupling, and connector losses

Laser-diode sources are generally not economically feasible or necessary in industrial systems. Light-emitting diodes are recommended for industrial applications. Such systems are frequently specified with transmitted power at 50 μW or greater if 200-μm-core fiber is used.

Successful communication in industry and commercial applications is determined by the amount of light energy required at the receiver, specified as *receiver sensitivity*. The higher the sensitivity, the less light required from the fiber. High-quality systems require power only in the hundreds of nanowatts to low microwatts range.

Splice losses must be low so that as little light as possible is removed from the optical-fiber system. Splice technology to repair broken cable is readily available, permitting repairs in several locations within a system in a short time (minutes) and causing negligible losses. Couplers and taps are generally formed through the process of glass fusion and operate on the principle of splitting from one fiber to several fibers. New active electronic couplers replenish light as well as distribute the optical signal.

An example of an optical power budget follows:

1. The optical power injected into a 200-μm-core fiber is 200 μW (the same light source would inject approximately 40 μW into a 100-μm-core fiber).
2. The receiver sensitivity is 2 μW.
3. The receiver budget (dB) is calculated as:

$$\begin{aligned} \text{dB} &= 10 \log [(\text{available light input})/(\text{required light output})] \\ &= 10 \log [(200\ \mu\text{W})/(2\ \mu\text{W})] \\ &= 10 \log 100 \\ &= 20\ \text{dB} \end{aligned}$$

4. Three major sources of loss are estimated as:
   - 2 to 3 dB loss for each end connector
   - 1 to 2 dB loss for each splice
   - 1 dB/150 m loss for fiber of 200 μm diameter

Most manufacturers specify the optical power budget and translate this into a recommended distance.

Digital Links—Pulsed

The one-for-one creation of a light pulse for an electrical pulse is shown in Fig. 2.32. This represents the simplest form of data link. It does not matter what format and signal level the electrical data takes (e.g., whether IEEE RS-232 or RS-422 standard format or CMOS or TTL logic level), as long as the interface circuitry is designed to accept them at its input or reproduce them at the output. Voltage conversion may be achieved from one end of the link to the other, if desired, through appropriate interface selection.

The light pulses racing down the fiber are independent of electrical protocol. Several design factors are relevant to these and other types of data links as follows:

- *Minimum output power.* The amount of light, typically measured in microwatts, provided to a specific fiber size from the data link's light source
- *Fiber size.* Determined from the data link's light source
- *Receiver sensitivity.* The amount of light, typically measured in microwatts or nanowatts, required to activate the data link's light detector
- *Data rate.* The maximum rate at which data can be accurately transmitted

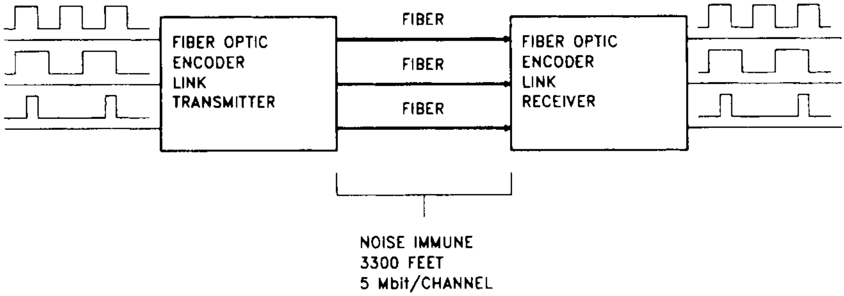


FIGURE 2.33 Three-channel optical link.

- *Bit error rate* (*BER*). The frequency with which a light pulse is erroneously interpreted (for example, 10^{-9} BER means no more than one of 10^9 pulses will be incorrect)
- *Pulse-width distortion.* The time-based disparity between input and output pulse widths

The simple pulse link is also the basic building block for more complex links. Figure 2.33 provides an example of a 5-Mbit three-channel link used to transmit encoder signals from a servomotor to a remote destination.

Digital Links—Carrier-Based

A carrier-based digital link is a system in which the frequency of the optical carrier is varied by a technique known as *frequency-shift keying* (*FSK*). Figure 2.34 illustrates the modulation concept; two frequencies are employed to create the logic 0 and 1 states. This scheme is especially useful in systems where electrical "handshaking" (confirmation of reception and acceptance) is employed. Presence of the optical carrier is the equivalent of the handshake signal, with the data signal presented by frequency.

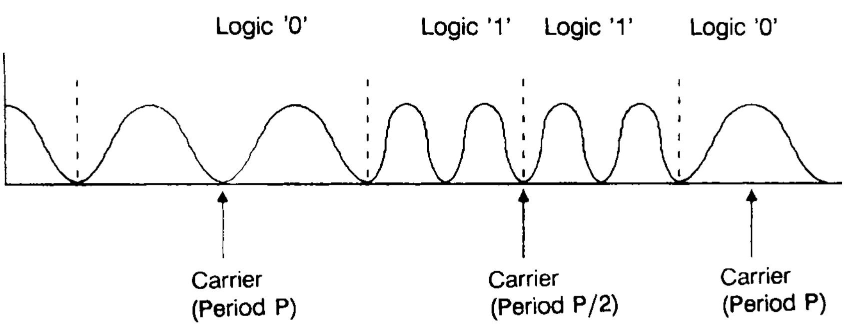


FIGURE 2.34 Modulation concept.

Figure 2.35 illustrates a system where the logic of the fiber-optic line driver recognizes the optical carrier to create a handshake between terminal and processor.

Additionally, since the processor is capable of recognizing only one terminal, the carrier is controlled to deny the handshake to all other terminals once one terminal is actively on line to the processor.

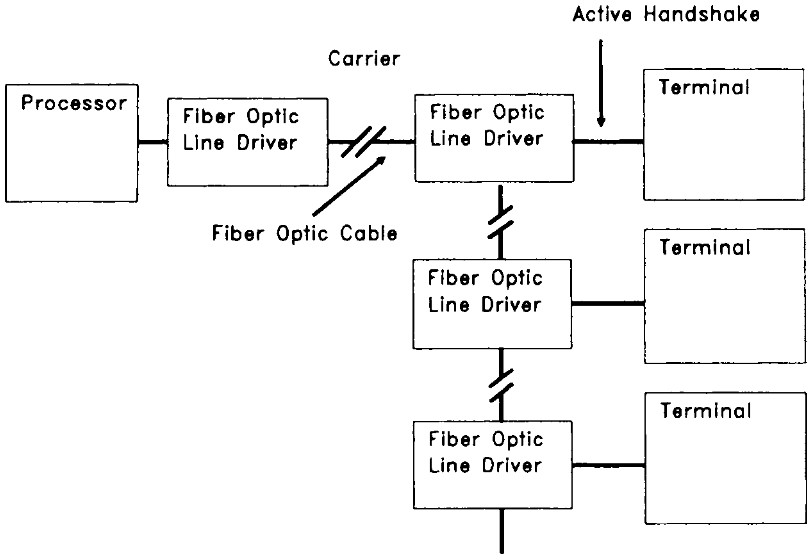


FIGURE 2.35 System employing optical carrier.

Analog Links

It is well-recognized that, in motion control and process measurement and control, transmitting analog information without distortion is important. Analog information can be treated in several ways with fiber optics.

Analog data cannot easily be transmitted through light intensity variation. A number of external factors, such as light source variation, bending losses in cable, and connector expansion with temperature, can affect the amount of raw light energy reaching the detector. It is not practical to compensate for all such factors and deliver accurate analog data. A viable method of transmitting data is to use an unmodulated carrier whose frequency depends on the analog signal level. A more advanced means is to convert the analog data to digital data, where accuracy also is determined by the number of bits used, multiplex the digital bits into one stream, and use the pulsed digital link approach.

Figure 2.36 illustrates a link in which this last approach is used to produce both digital and analog forms of the data at the output.

Video Links

Long-distance video transmission in industrial situations is easily disrupted by radiated noise and lighting. Repeaters and large-diameter coaxial cables are often used for particularly long runs. The use of fiber optics as a substitute for coaxial cable allows propagation of noise-free video over long distances. Either an intensity- or frequency-modulated optical carrier signal is utilized as the transmission means over fiber. With intensity-modulated signals, it is mandatory that some sort of automatic gain control be employed to compensate for light degradation due to varying cable losses, splices, and so on. Figure 2.37 illustrates a typical fiber-optic video link in a machine-vision application.

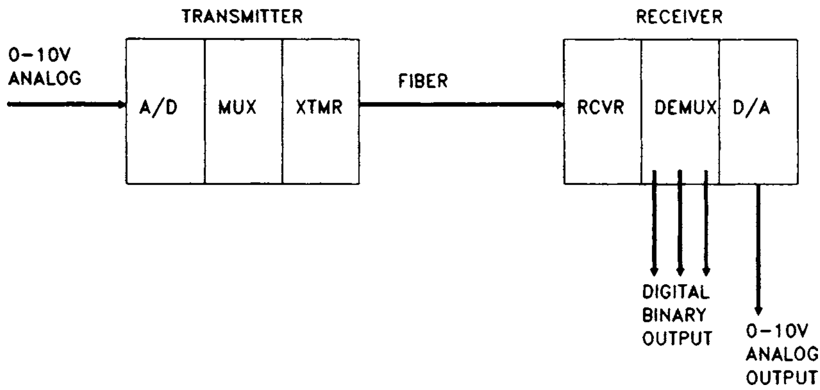


FIGURE 2.36 Analog and digital data transmission.

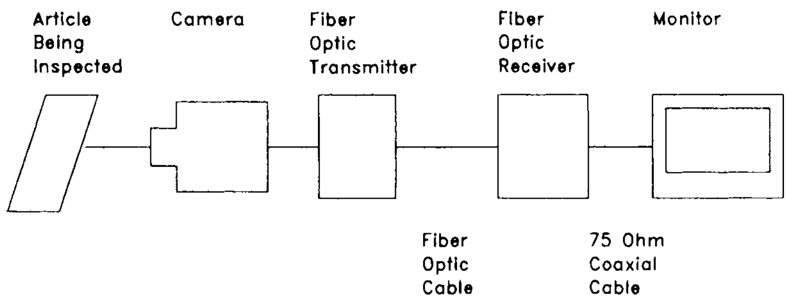


FIGURE 2.37 Fiber-optic video link.

Data Bus Networks

Wiring a system often causes serious problems for designers and communication system integrators regarding the choice of topology. The basic difference between fiber and wiring is that one normally does not splice or tap into fiber as one would with coaxial or twin axial cable to create a drop point.

Daisy Chain Data Bus. The simplest extension of a point-to-point data link is described in Fig. 2.38. It extends continuously from one drop point (node) to the next by using each node as a repeater. The fiber-optic line driver illustrated in Fig. 2.35 is such a system, providing multiple access points from remote terminals to a programmable controller processor. A system with several repeater points is vulnerable to the loss of any repeater, and with it all downstream points, unless some optical bypass scheme is utilized. Figures 2.38 and 2.39 exhibit such a scheme.

Ring Coupler. A preferred choice among several current fiber-optic system designs is the token-passing ring structure.

Signals are passed around the ring, with each node serving to amplify and retransmit. Care must be taken to provide for a node becoming nonoperational. This is usually handled by using some type of bypass switching technique, given that the system provides sufficient optical power to tolerate a bypassed repeater. Another contingency method is to provide for the

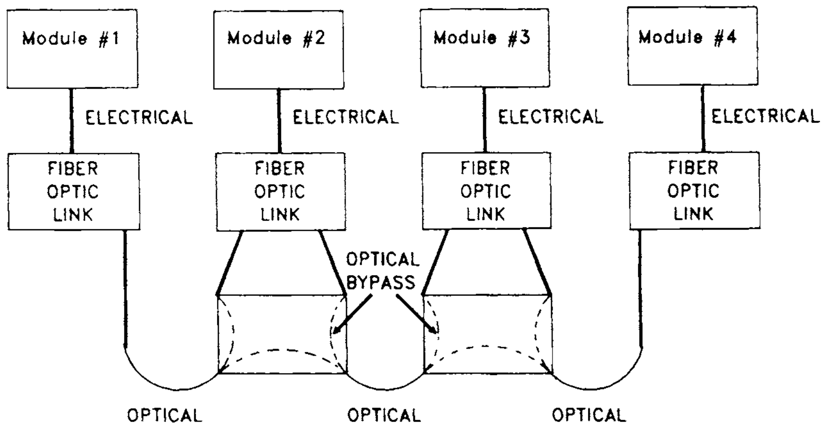


FIGURE 2.38 Point-to-point data link.

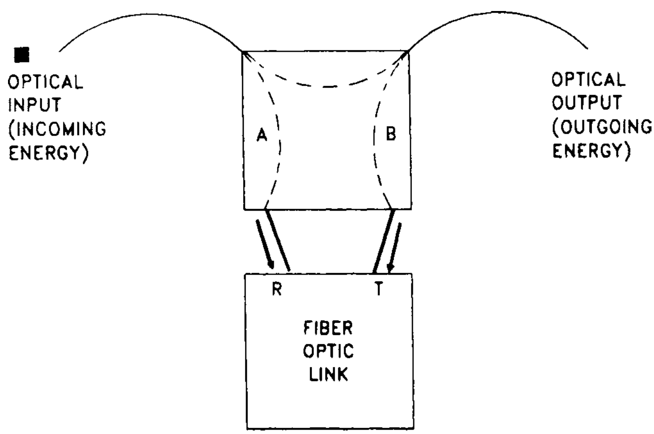


FIGURE 2.39 Daisy chain data bus.

transmitting node to read its own data coming around the ring, and to retransmit in the other direction if necessary, as illustrated in Fig. 2.40. Yet another is to provide for a second pair of fibers paralleling the first, but routed on a physically different path.

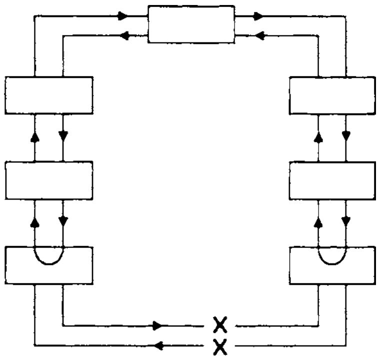


FIGURE 2.40 Ring coupler.

Passive Star Coupler. Certain systems have attempted to utilize a fiber-optic coupling technology offered from the telecommunications and data communications applications areas. When successful, this technique allows tapping into fiber-optic trunk lines, a direct parallel with coaxial or twin axial systems. Light entering into the tap or coupler is split into a given number of output channels. The amount of light in any output channel is determined by the total amount of light input, less system losses, divided by the number of output channels. Additional losses are incurred at the junction of the main data-carrying fibers with the fiber leads from the tap or star. As such, passive couplers are limited to systems with few drops and moderate distances. Also, it is important to minimize termination losses at the coupler caused by the already diminished light output from the coupler. A partial solution is an active in-line repeater, but a superior solution, the *active star coupler*, is described next.

The Active Star Coupler. The basic principle of the active star coupler is that any light signal received as an input is converted to an electrical signal, amplified, and reconverted to optical signals on all other output channels. Figure 2.41 illustrates an eight-port active star coupler, containing eight sets of fiber-optic input/output (I/O) ports. A signal received on the channel 1 input will be transmitted on the channel 2 to 8 output ports. One may visualize the use of the active star coupler as aggregating a number of taps into one box. Should the number of required taps exceed the number of available I/O ports, or should it be desirable to place these tap boxes at several locations in the system, the active star couplers may be jumpered together optically by tying a pair of I/O ports on one coupler to that on another in a hub-and-spoke system.

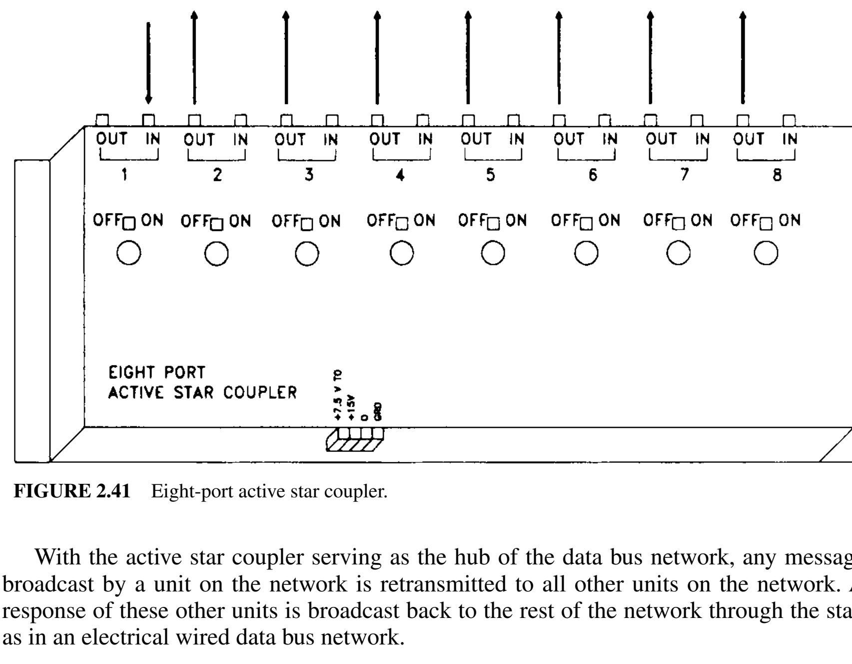


FIGURE 2.41 Eight-port active star coupler.

With the active star coupler serving as the hub of the data bus network, any message broadcast by a unit on the network is retransmitted to all other units on the network. A response of these other units is broadcast back to the rest of the network through the star, as in an electrical wired data bus network.

CONFIGURATIONS OF FIBER OPTICS FOR SENSORS

Fiber-optic sensors for general industrial use have largely been restricted to applications in which their small size has made them convenient replacements for conventional photoelectric sensors. Until recently, fiber-optic sensors have almost exclusively employed standard bundle technology, whereby thin glass fibers are bundled together to form flexible conduits for light.

Recently, however, the advances in fiber optics for data communications have introduced an entirely new dimension into optical sensing technology. Combined with novel but effective transducing technology, they set the stage for a powerful class of fiber-optic sensors.

Fiber-Optic Bundles

A typical fiber-optic sensor probe, often referred to as a *bundle* (Fig. 2.42), is 1.25 to 3.15 mm in diameter and made of individual fiber elements approximately 0.05 mm in diameter. An average bundle will contain up to several thousand fiber elements, each working on the conventional fiber-optic principle of total internal reflection.

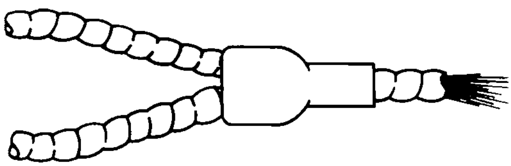


FIGURE 2.42 Fiber-optic sensor probe.

Composite bundles of fibers have an acceptance cone of the light based on the numerical aperture of the individual fiber elements.

$$\begin{aligned} NA &= \sin \Theta \\ &= \sqrt{n_1^2 - n_2^2} \end{aligned} \tag{2.1}$$

where $n_1 > n_2$ and Θ = half the cone angle.

Bundles normally have NA values in excess of 0.5 (an acceptance cone full angle greater than 60°), contrasted with individual fibers for long-distance high-data-rate applications, which have NA values approaching 0.2 (an acceptance cone full angle of approximately 20°).

The ability of fiber-optic bundles to readily accept light, as well as their large total cross-sectional surface area, have made them an acceptable choice for guiding light to a remote target and from the target area back to a detector element. This has been successfully accomplished by using the pipe as an appendage to conventional photoelectric sensors, proven devices conveniently prepackaged with adequate light source and detector elements.

Bundles are most often used in either opposed beam or reflective mode. In the opposed beam mode, one fiber bundle pipes light from the light source and illuminates a second bundle—placed on the same axis at some distance away—which carries light back to the detector. An object passing between the bundles prevents light from reaching the detector.

In the reflective mode, all fibers are usually contained in one probe but divided into two legs at some junction point in an arrangement known as *bifurcate*. One bifurcate leg is then tied to the light source and the other to the detector (Fig. 2.43). Reflection from a target provides a return path to the detector for the light. The target may be fixed so it breaks the beam, or it may be moving so that, when present in the probe's field of view, it reflects the beam.

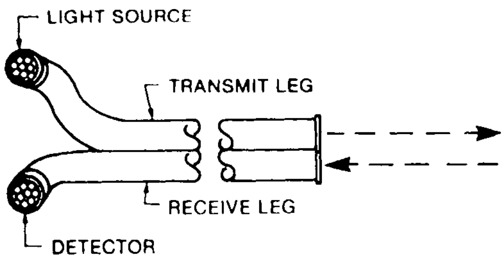


FIGURE 2.43 Reflective mode bifurcate fiber optics.

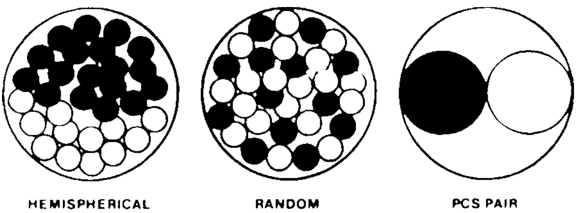


FIGURE 2.44 Bundle construction.

Typical bundle construction in a bifurcate employs one of two arrangements of individual fibers. The sending and receiving fibers are arranged either randomly or hemispherically (Fig. 2.44). As a practical matter, there is little, if any, noticeable impact on the performance of a photoelectric system in any of the key parameters such as sensitivity and scan range.

Application areas for bundle probes include counting, break detection, shaft rotation, and displacement/proximity sensing.

Bundle Design Considerations

Microscopic fiber flaws (Fig. 2.45) such as impurities, bubbles, voids, material absorption centers, and material density variations all diminish the ability of rays of light to propagate down the fiber, causing a net loss of light from one end of the fiber to the other. All these effects combine to produce a characteristic absorption curve, which graphically expresses a

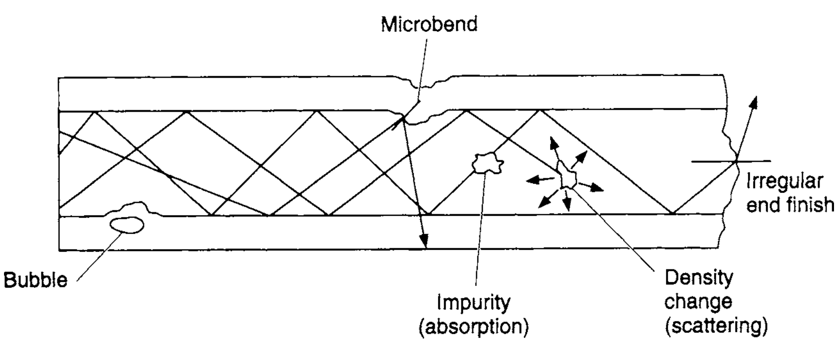


FIGURE 2.45 Microscopic fiber flaws.

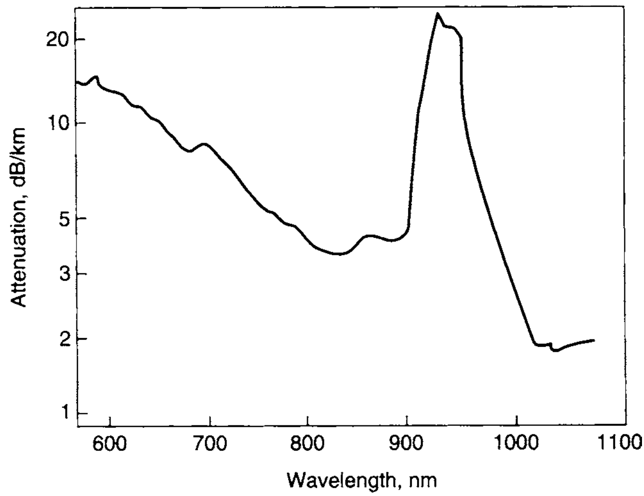


FIGURE 2.46 Characteristic absorption curve.

wavelength-dependent loss relationship for a given fiber (Fig. 2.46). The fiber loss parameter is expressed as attenuation in dB/km as follows:

$$\text{Loss} = -10 \log (p_2/p_1) \tag{2.2}$$

where p_2 = light power output and p_1 = light power input.

Therefore, a 10-dB/km fiber would produce only 10 percent of the input light at a distance of 1 km.

Because of their inexpensive lead silicate glass composition and relatively simple processing techniques, bundles exhibit losses in the 500-dB/km range. This is several orders of magnitude greater than a communications-grade fiber, which has a loss of 10 dB/km. The maximum practical length for a bundle is thus only about 3 m. Further, the absence of coating on individual fibers and their small diameter make them susceptible to breakage, especially in vibratory environments.

Also, because of fiber microflaws, it is especially important to shield fibers from moisture and contaminants. A fiber exposed to water will gradually erode to the point of failure. This is true of any optical fiber, but is especially true of uncoated fibers in bundles.

Fiber Pairs for Remote Sensing

A viable solution to those design problems that exist with fiber bundles is to use large-core glass fibers that have losses on a par with telecommunication fibers. Although its ability to accept light is less than that of a bundle, a 200- or 400-μm core diameter plastic clad silica (PCS) fiber provides the ability to place sensing points hundreds of meters away from corresponding electronics. The fiber and the cable construction in Fig. 2.47 lend themselves particularly well to conduit pulling, vibratory environments, and general physical abuse. These fibers are typically proof-tested for tensile strength to levels in excess of 50,000 lb/in^2. A pair of fibers (Fig. 2.44) is used much like a bundle, where one fiber is used to send light to the sensing point and the other to return light to the detector. The performance limitation of a fiber pair compared to a bundle is reduced scan range; however, lenses may be used to extend the range. A fiber pair may be used

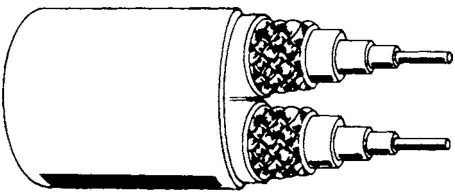


FIGURE 2.47 Pair of fibers.

in one of two configurations: (1) a single continuous probe—that is, an unbroken length of cable from electronics to sensing point, or (2) a fiber-optic extension cord to which a standard probe in either a bundle or a fiber pair is coupled mechanically. This allows the economical replacement, if necessary, of the standard probe, leaving the extension cord intact.

The typical application for a fiber pair is object detection in explosive or highly corrosive environments—for example, ammunition plants. In such cases, electronics must be remote by necessity. Fiber pairs also allow the construction of very small probes for use in such areas as robotics, small object detection, thread break detection, and small target rotation.

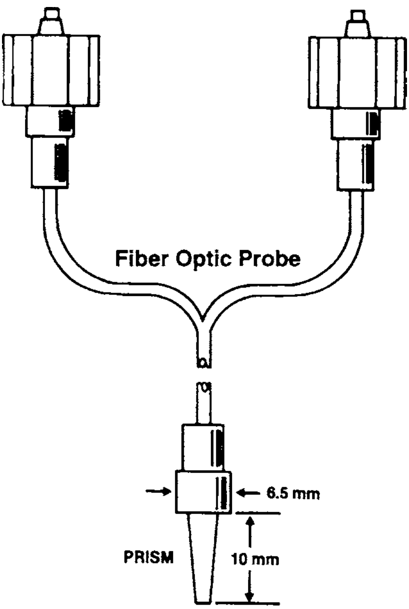


FIGURE 2.48 Prism tip for liquid sensing.

Fiber-Optic Liquid Level Sensing

Another technique for interfacing with fiber-optic probes involves the use of a prism tip for liquid sensing (Fig. 2.48). Light traveling down one leg of the probe is totally internally reflected at the prism-air interface. The index of refraction of air is 1. Air acts as a cladding material around the prism. When the prism contacts the surface of a liquid, light is stripped from the prism, resulting in a loss of energy at the detector. A properly configured system can discriminate between liquid types, such as gasoline and water, by the amount of light lost from the system, a function of the index of refraction of the liquid.

This type of sensor is ideal for set-point use in explosive liquids, in areas where electronics must be remote from the liquid by tens or hundreds of meters, and where foam or liquid turbulence make other level-sensing techniques unusable.

FLEXIBILITY OF FIBER OPTICS

The power of fiber optics is further shown in the flexibility of its system configurations. A master industrial terminal (Fig. 2.49) can access any of a number of remote processors. The flexibility of switching, distance capability, and noise immunity of such a system are its primary advantages.

Figure 2.50 illustrates a passive optical coupler with a two-way fiber-optic link communicating over a single fiber through an on-axis rotary joint. Such a system allows a simple

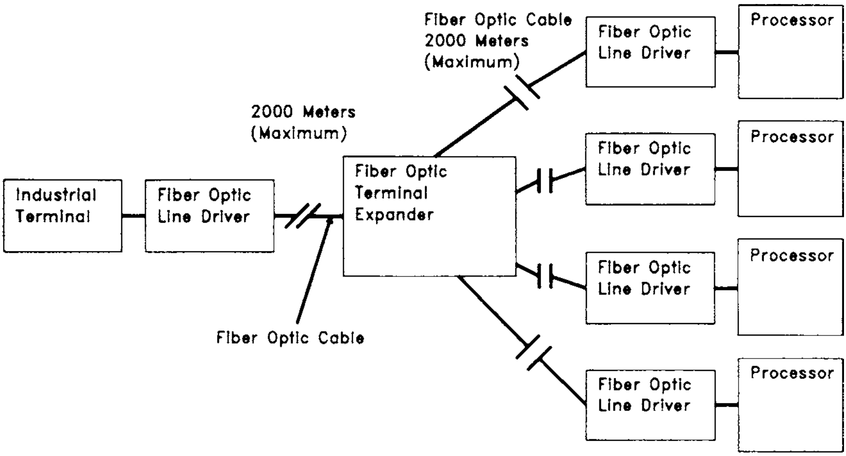


FIGURE 2.49 Master industrial terminal and remote processors.

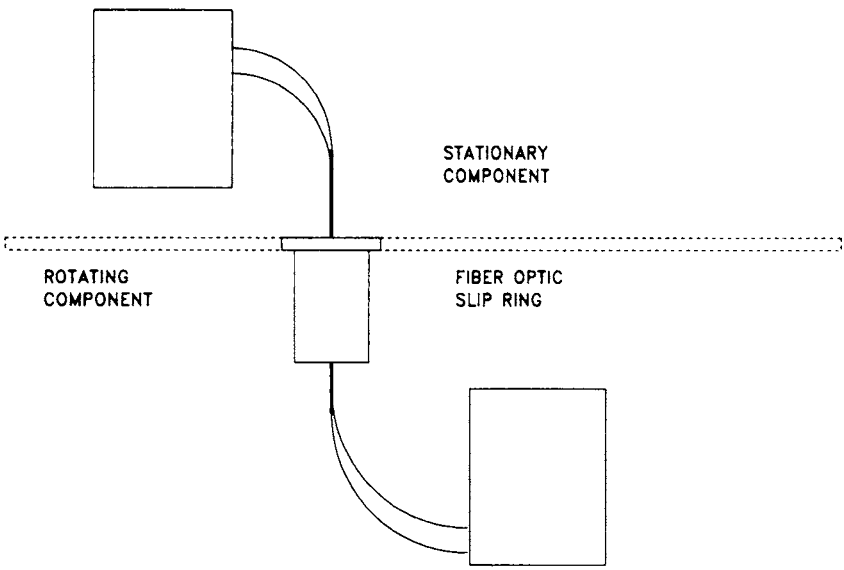


FIGURE 2.50 Passive optical coupler.

uninterrupted communication link through rotary tables or other rotating machinery. This is a true challenge for high-data-rate communication in a wire system.

Fiber-Optic Terminations

Optical fibers are becoming increasingly easier to terminate as rapid advances in termination technology continue to be made. Several manufacturers have connector systems that require no polishing of the fiber end, long a major objection in fiber optics. Products that eliminate epoxy adhesives are also being developed. Field installation times now typically average less than 10 min for large-core fibers (100 and 200 μm) with losses in the 1- to 3-dB range. Further, power budgets for well-designed industrial links normally provide a much greater latitude in making a connection. A 5- to 6-dB-loss connection, while potentially catastrophic in other types of systems, may be quite acceptable in short-haul systems with ample power budgets.

The most popular connector style for industrial communications is the SMA style connector, distinguished by its nose dimensions and configuration, as well as the thread size on the coupling nut. The coupling nut is employed to mechanically join the connector to a mating device on the data link or to a thread splice bushing. Figure 2.51 illustrates an SMA connection to an information link.

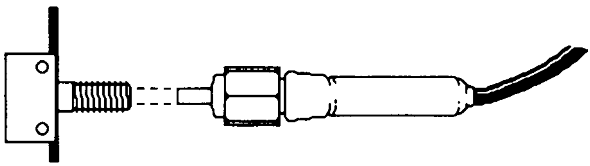


FIGURE 2.51 SMA connection to an information link.

TESTING OF FIBER OPTICS

Optical measurements, perhaps among the most difficult of all physical measurements, are fundamental to the progress and development of fiber-optic technology. Recently, various manufacturers have offered lines of fiber-optic test equipment for use in field and laboratory. Typical field measurement equipment determines the average optical power emitted from the system source, the component and overall system loss, the bit error rate, and the location of breaks in the fiber. Laboratory equipment measures loss through connectors and splicing, characterizes transmitters and receivers, and establishes bit error rate.

Testing of fiber-optic cables or systems is normally done with a calibrated light source and companion power meter. The light source is adjusted to provide a 0-dB reading on the power meter with a short length of jumper cable. The cable assembly under test is then coupled between the jumper and the power meter to provide a reading on the meter, in decibels, that corresponds to the actual loss in the cable assembly.

Alternatively, the power through the cable from the system's transmitter can be read directly and compared with the system's receiver sensitivity specification. In the event of a cable break in a long span, a more sophisticated piece of test equipment, an optical time-domain reflectometer (OTDR), can be employed to determine the exact position of the break.

Test Light Sources

The Photodyne 9XT optical source driver (Fig. 2.52) is a handheld unit for driving LED and laser fiber-optic light sources. The test equipment is designed to take the shock and hard wear of the typical work-crew environment. The unit is powered from two rechargeable nicad batteries or from line voltage. The LED series is suited for measurement applications where moderate dynamic range is required or coherent light should be avoided. The laser modules are used for attenuation measurements requiring extreme dynamic range or where

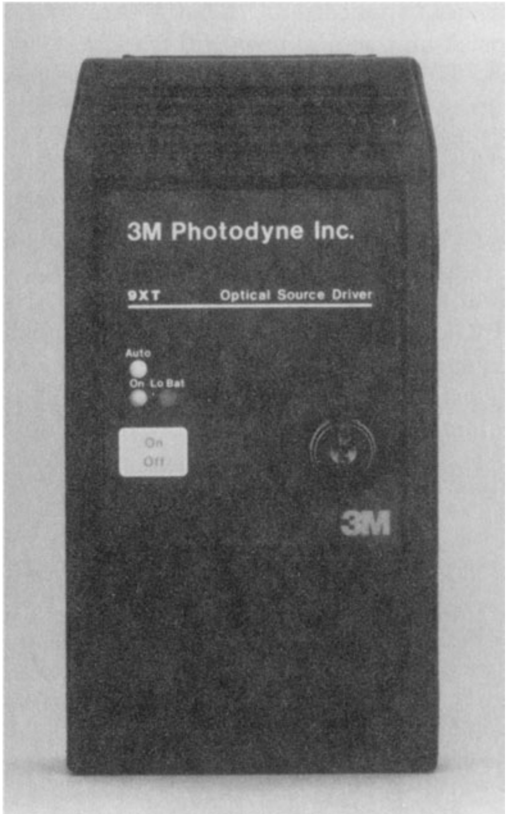


FIGURE 2.52 Photodyne 9XT optical source driver. (*Courtesy 3M Corporation*)

narrow spectral width and coherent light are required. The laser modules are the most powerful source modules available.

Power Meters

The Photodyne 2285XQ fiber-optic power meter (Fig. 2.53) is a low-cost optical power meter for general-purpose average-power measurement, but particularly for fiber-optic applications, in both manual and computer-controlled test setups. A unique and powerful feature of the 2285XSQ is its ratio function. This allows the user to make ratio (*A/B*) measurements by stacking several instruments together via the interface without the need for a controller.

Another very powerful feature is the built-in data logger. With this function, data may be taken at intervals from 1 to 9999 s. This feature is useful in testing optical devices for short- and long-term operation.

At the heart of the instrument is a built-in large-area high-sensitivity indium gallium arsenide (InGaAs) sensor. All industry connectors may be interfaced to the sensor using any of the Photodyne series 2000 connector adapters. The sensor is calibrated for all the fiber-optic windows: 820, 850, 1300, and 1550 nm.

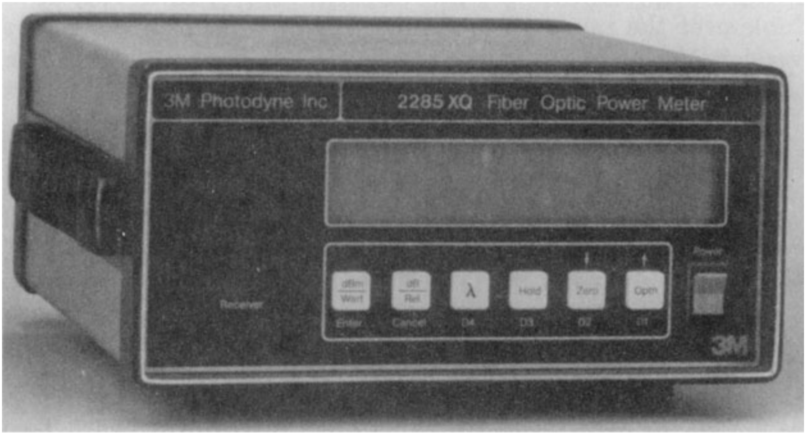


FIGURE 2.53 Photodyne 2285XQ fiber-optic power meter. (*Courtesy 3M Corporation*)

Dual Laser Test Sets

The Photodyne 2260XF and 2260XFA are switchable dual laser test sets, with a transmit and receive section in one unit. They measure and display loss in fiber links at both 1300 and 1550 nm simultaneously in one pass. They are designed for use in installing, maintaining, and troubleshooting single-mode wavelength-division multiplexed (WDM) fiber-optic links operating at 1300- and 1550-nm wavelengths. They may also be used for conventional links operating at either 1300 or 1550 nm.

The essential differences between the two models are that the XF version (Fig. 2.54) has a full complement of measurement units, log and linear (dBm, dB, mW, µW, nW), whereas the XFA version has log units only (dBm, dB). In the XF version, laser power output is adjustable over the range 10 to 20 dBm (100 µW to 21 mW). In the XFA version, it is automatically set to a fixed value of 10 dBm. The XF version allows the user to access wavelengths over the ranges 1250 to 1350 nm and 1500 to 1600 nm in 10-nm steps. The XFA version is fixed at 1300 nm to 1550 nm. To control all functions, the XF version has six keys; the XFA has only two. Although both instruments perform identical tasks equally well, the XF may be seen as the more flexible version and the XFA as the simpler-to-use version.

Test Sets/Talk Sets

The handheld Photodyne 21XTL fiber-optic test set/talk set (Fig. 2.55) is for use in installation and maintenance of fiber cables. This instrument functions as a power meter and test set, as well as a talk set. For maintenance purposes, the user may establish voice communication over the same fiber pair being measured.

The 21XTL as a power meter covers an exceptionally wide dynamic range (–80 to +3 dBm). As an option, the receiver may include a silicon or an enhanced InGaAs photodiode. With the InGaAs version, the user can perform measurements and voice communication at short and long wavelengths. The silicon version achieves a superior dynamic range at short wavelengths only.

The highly stabilized LED ensures repeatable and accurate measurements. Precision optics couples the surface-emitter LED to the fiber core. With this technique, the fiber end will not wear out or scratch. The transmitter is interchangeable, providing complete flexibility of wavelengths and connector types.

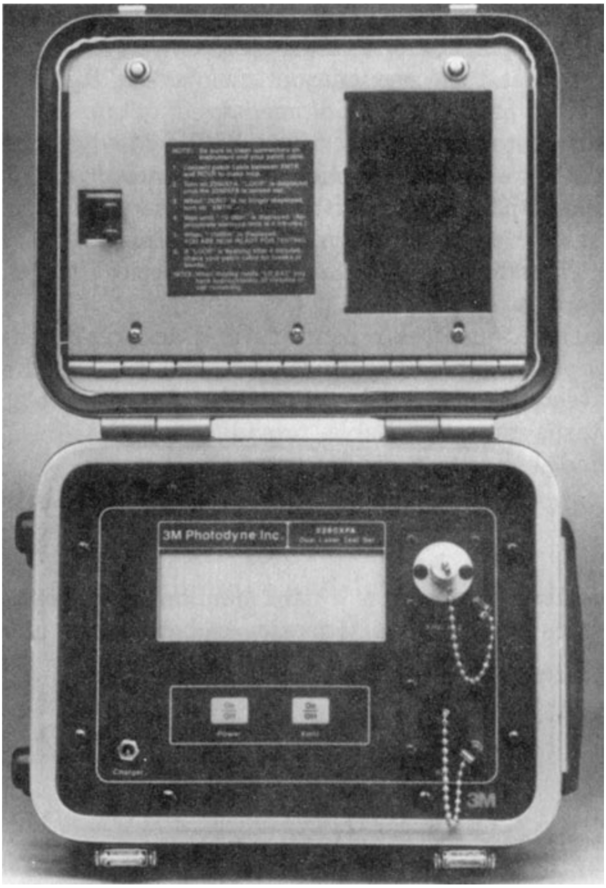


FIGURE 2.54 Photodyne 2260XF switchable dual laser test set. (*Courtesy 3M Corporation*)

Attenuators

The 19XT optical attenuator (Fig. 2.56) is a handheld automatic optical attenuator. It provides continuous attenuation of both short- and long-wave optical signals in single-mode and multimode applications. The calibrated wavelengths are 850 nm/1300 nm multimode and/or 1300 nm/1550 nm single mode. An attenuation range of 70 dB is offered with 0.1-dB resolution and an accuracy of ±0.2 dB typical (±0.5 dB maximum). Unique features allow scanning between two preset attenuation values and include the insertion loss in the reading.

The 19XT allows simple front panel operation or external control of attenuation through analog input and output connections.

Fault Finders

The Photodyne 5200 series optical fault finders (Fig. 2.57) offer flexible alternatives for localizing faults or trouble areas on any fiber-optic network. The 5200 series fault finders

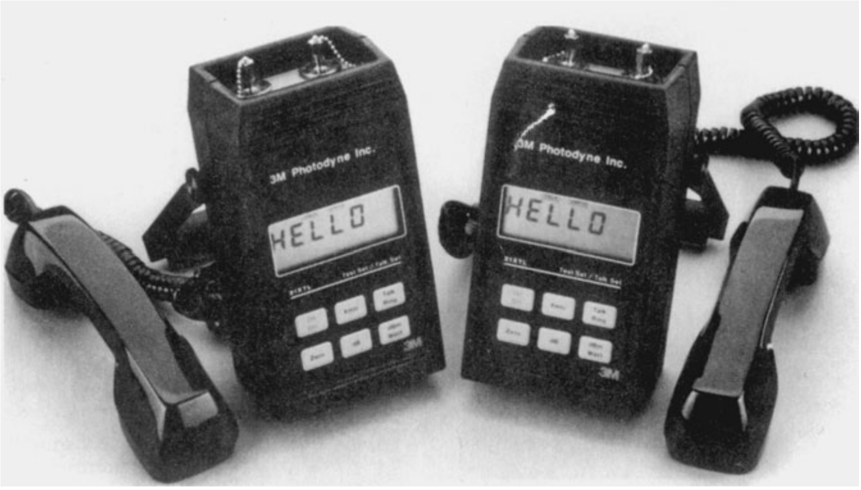


FIGURE 2.55 Photodyne 21XTL fiber-optic test set/talk set. (*Courtesy 3M Corporation*)

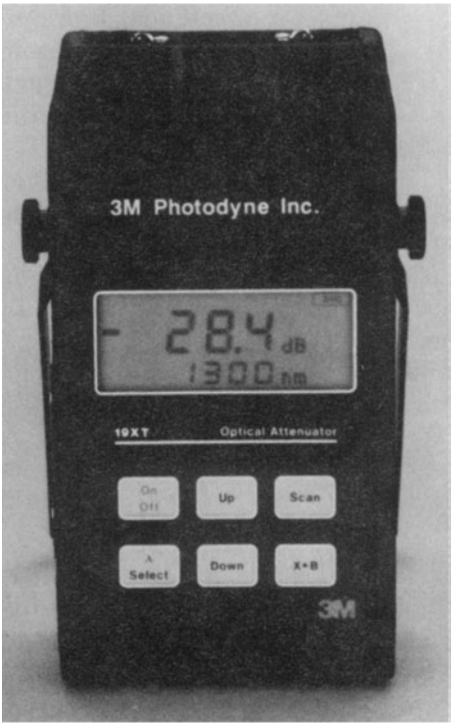


FIGURE 2.56 Photodyne 19XT optical attenuator. (*Courtesy 3M Corporation*)

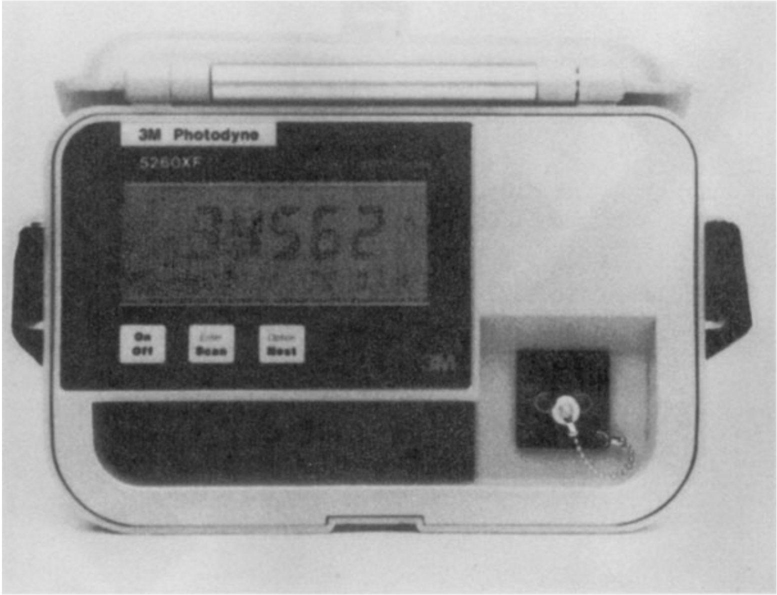


FIGURE 2.57 Photodyne 5200 series optical fault finder. (*Courtesy 3M Corporation*)

can easily be integrated into the troubleshooting routines of fiber-optic crews. They offer a variety of features, such as:

- Fast, accurate analysis of faults
- Autoranging for greatest accuracy
- Reflective and nonreflective fault detection
- Multiple-fault detection capabilities
- Go/no-go splice qualification
- Variable fault threshold setting (0.5 to 6.0 dB)
- Fault location up to 82 km
- Automatic self-test and performance check
- AC or rechargeable battery operation
- Large easy-to-read liquid-crystal display (LCD)

Fiber Identifiers

The Photodyne 8000XG fiber identifier (Fig. 2.58) is designed for fast, accurate identification and traffic testing of fiber-optic lines without cutting the fiber line or interrupting normal service. Ideal for use during routine maintenance and line modification, this small handheld unit can be used to locate any particular fiber line, identify live fibers, and determine whether or not traffic is present. Features of the 8000XG include:

- Lightweight design, portability, and battery operation
- Automatic self-tests after each fiber insertion

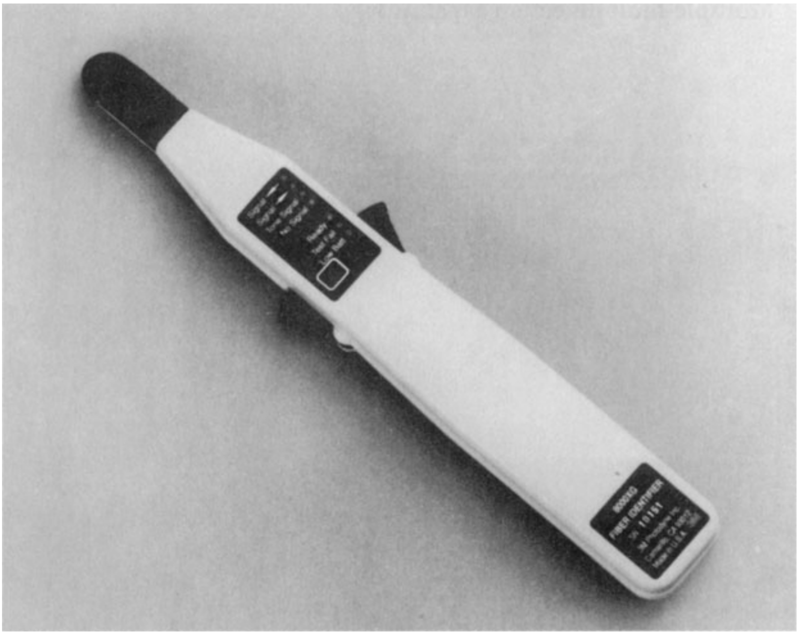


FIGURE 2.58 Photodyne 8000XG fiber identifier. (*Courtesy 3M Corporation*)

- Mechanically damped fiber action
- Operation over 850- to 1550-nm range
- Transmission direction indicators
- 1- and 2-kHz tone detection
- Low insertion loss at 1300 and 1550 nm
- Completely self-contained operation

NETWORKING WITH ELECTRO-OPTIC LINKS

The following examples describe a number of products utilizing communication through fiber electrooptic modules. The function of the fiber is to replace wire. This can be achieved by interconnecting the electrooptic modules in a variety of ways. Figure 2.59 shows a programmable controller communication network through-coaxial-cable bus branched to four remote input/output stations. The programmable controller polls each of the input/output stations in sequence. All input/output stations hear the programmable controller communication, but only the one currently being addressed responds.

Hybrid Wire/Fiber Network

Figure 2.60 shows an electrooptic module used to replace a troublesome section of coaxial cable subject to noise, grounding problems, lightning, or hazardous environment. When fiber is used in this mode, the electrooptic module should be placed as close to the input/

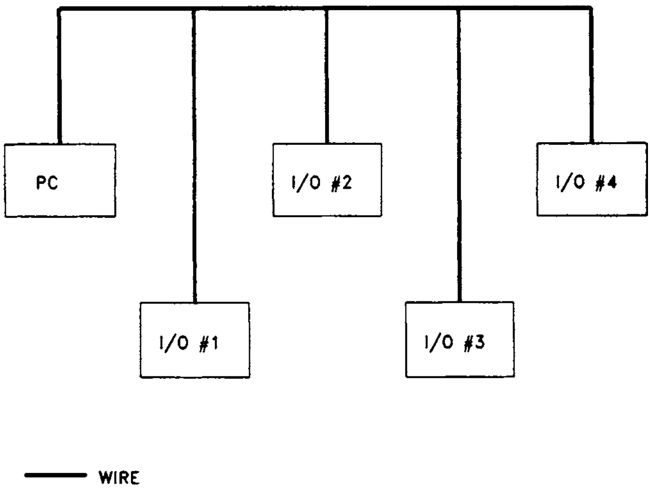


FIGURE 2.59 Programmable controller communication network.

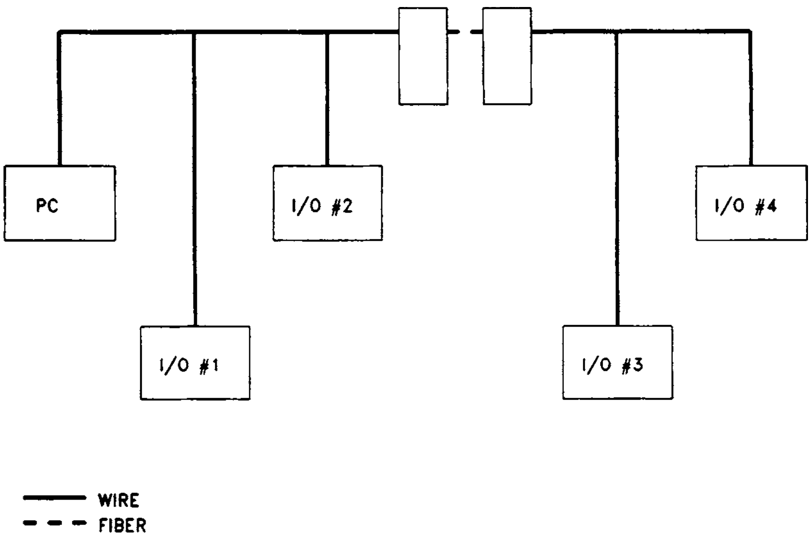


FIGURE 2.60 Hybrid wire/fiber network.

output drop as possible in order to minimize the effect of potential electrical noise problems over the section of coax cable connecting them.

Daisy Chain Network

The daisy chain configuration (Fig. 2.61) is an economical choice for long straight-line installations (e.g., conveyer or mine shaft applications). The signal generated at the programmable controller is converted to light and transmitted outward. At each transmitted section, the electrical signal is reconstructed and a light signal is regenerated and sent down the chain.

Active Star Network

The electrooptical programmable controller links may be joined to an electrooptic module with four-port and eight-port active star couplers for a hub-and-spoke-type system. Figure 2.62 shows two active star couplers joined via fiber, a configuration that might be appropriate where clusters of input/output racks reside in several locations separated by some distance. The star coupler then becomes the distributor of the light signals to the racks in each cluster, as well as a potential repeater to a star in another cluster.

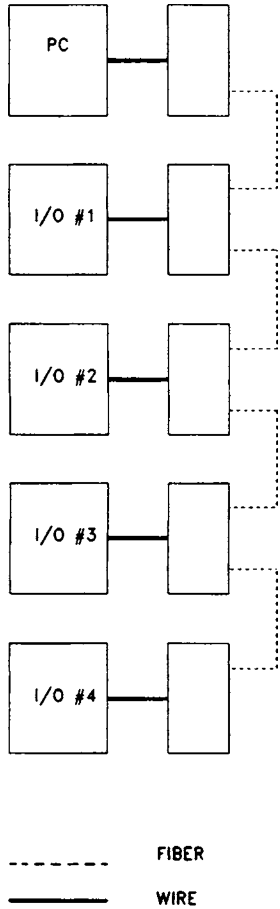


FIGURE 2.61 Daisy chain network.

Hybrid Fiber Network

Star and daisy chain network structures can be combined to minimize overall cabling costs (Fig. 2.63). The fiber network configuration exactly duplicates the coax network. The final decision on exactly which configuration to choose depends on the following criteria:

- Economical cabling layout
- Length of cable runs versus cost of electronics
- Location of end devices and power availability
- Power-out considerations

Other considerations specific to the programmable controller being used can be summarized as follows:

- *Pulse width.* Total acceptable pulse-width distortion, which may limit the number of allowable repeater sites, is Allowable distortion (ns) = allowable percent distortion × period of signal (ns).
- *Signal propagation delay.* Allowable system signal propagation delay may limit the overall distance of the fiber network.

Example.

Electrooptic module distortion = 50% allowable distortion × 1 Mbaud transmission
= 50 ns

Calculate.

Allowable distortion = 50% allowable distortion × 10^{-6}
= 500 ns

Maximum number of repeater sites = 500 ns/50 ns
= 10

(Note: A star coupler counts as a repeater site.)

Distance Calculation (Signal Propagation Delay)	
Delay of light in fiber-optic regeneration	1.5 ns/ft
Delay in module	50 ns

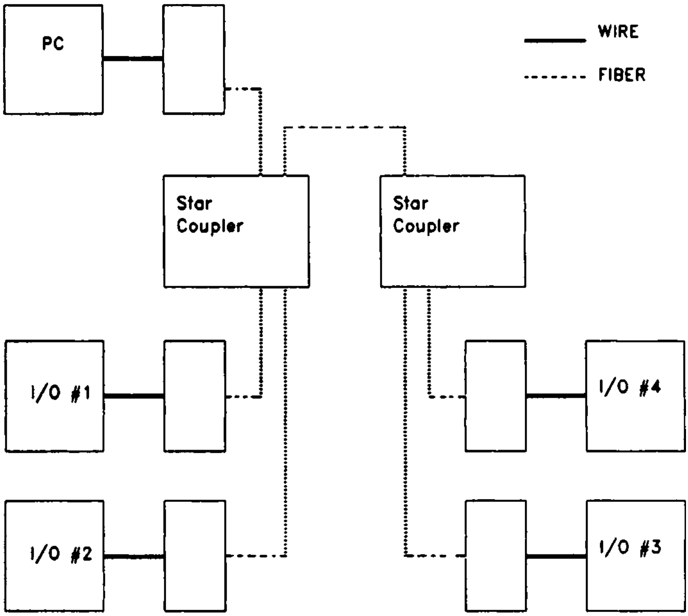


FIGURE 2.62 Two active star couplers.

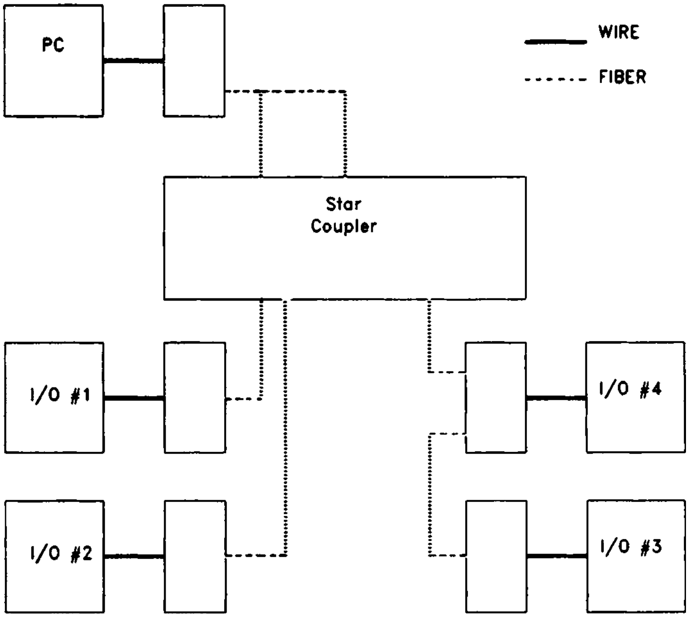


FIGURE 2.63 Hybrid fiber network.

Manufacturers of programmable controllers will provide the value of the system delay. This value must be compared with the calculated allowable delay. If the overall fiber system is longer than the published maximum length of the wired system, the system must be reconfigured.

Fiber-Optic Sensory Link for Minicell Controller

A minicell controller is typically used to coordinate and manage the operation of a manufacturing cell, consisting of a group of automated programmable machine controls (programmable controllers, robots, machine tools, etc.) designed to work together and perform a complete manufacturing or process-related task. A key benefit of a minicell controller is its ability to adjust for changing products and conditions. The minicell controller is instructed to change data or complete programs within the automation work area. A minicell controller is designed to perform a wide variety of functions such as executing programs and data, uploading/downloading from programmable controllers, monitoring, data analysis, tracking trends, generating color graphics, and communicating in the demanding plant floor environment. Successful minicell controllers use fiber-optic links that can interface with a variety of peripheral devices. A minicell controller can be used in a variety of configurations, depending on the optical-fiber lengths, to provide substantial system design and functional flexibility (Fig. 2.64).

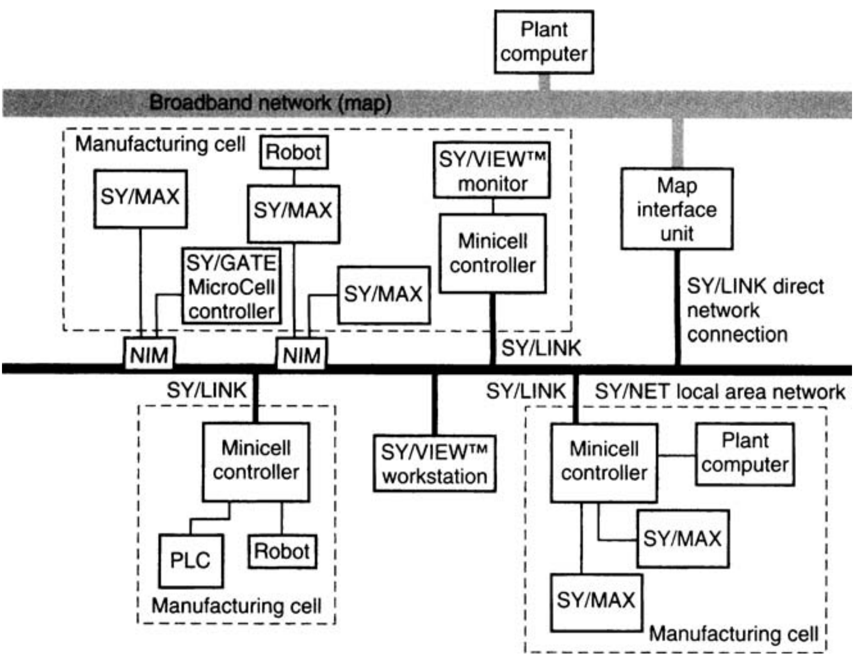


FIGURE 2.64 Minicell controller.

VERSATILITY OF FIBER OPTICS IN INDUSTRIAL APPLICATIONS

A constant concern in communication is the ever-increasing amount of information that must be sent with greater efficiency over a medium requiring less space and less susceptibility to outside interferences. As speed and transmission distance increase, the problems caused by electromagnetic interference, radio-frequency interference, crosstalk, and signal distortion become more troublesome. In terms of signal integrity, just as in computer-integrated manufacturing data acquisition and information-carrying capacity, fiber optics offers many advantages over copper cables. Furthermore, optical fibers emit no radiation and are safe from sparking and shock. These features make fiber optics the ideal choice for many processing applications where safe operation in hazardous or flammable environments is a requirement.

Accordingly, fiber-optic cables have the following advantages in industrial applications:

- Wide bandwidth
- Low attenuation
- Electromagnetic immunity
- No radio-frequency interference
- Small size
- Light weight
- Security
- Safety in hazardous environment

High-Clad Fiber-Optic Cables

Large-core, multimode, step-index high-clad silica fiber-optic cables make fiber-optic technology user-friendly and help designers of sensors, controls, and communications realize substantial cost saving. The coupling efficiency of high-clad silica fibers allows the use of less expensive transmitters and receivers. High-clad silica polymer technology permits direct crimping onto the fiber cladding. Field terminations can be performed in a few minutes or less, with minimal training. The following lists describe the structure and characteristics of several fiber-optic cables used in industry:

Simplex fiber-optic cables (Fig. 2.65) are used in:

- Light-duty indoor applications
- Cable trays
- Short conduits
- Loose tie wrapping
- Subchannels for breakout cables

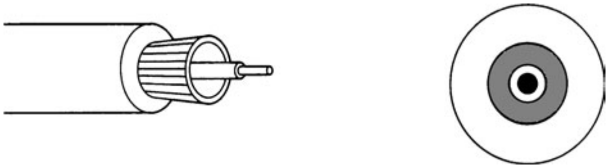


FIGURE 2.65 Simplex fiber-optic cable.

Zipcord fiber-optic cables (Fig. 2.66) are used in:

- Light-duty (two-way) transmission
- Indoor runs in cable trays
- Short conduits
- Tie wrapping

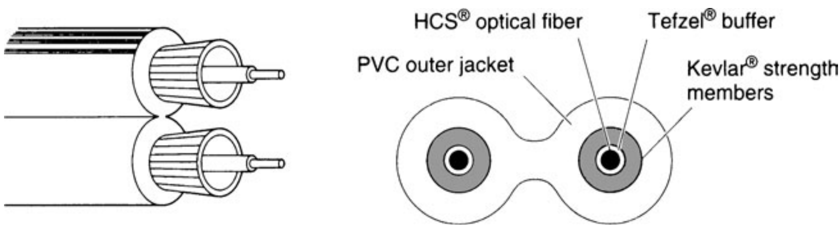


FIGURE 2.66 Zipcord fiber-optic cable.

Multichannel fiber-optic cables (Fig. 2.67) are used in:

- Outdoor environments
- Multifiber runs where each channel is connectorized and routed separately
- Aerial runs
- Long conduit pulls
- Two, four, and six channels (standard)
- Eight to eighteen channels

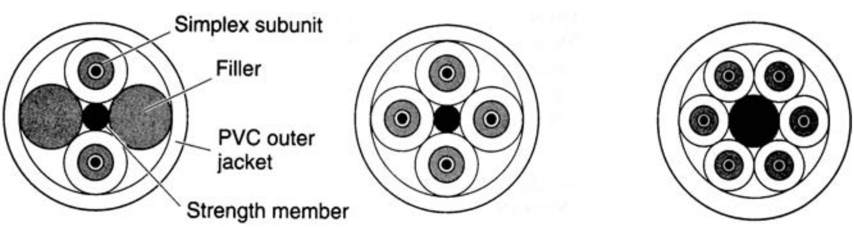


FIGURE 2.67 Multichannel fiber-optic cable.

Heavy-duty duplex fiber-optic cables (Fig. 2.68) are used in:

- Rugged applications
- Wide-temperature-range environments
- Direct burial
- Loose tube subchannel design
- High-tensile-stress applications

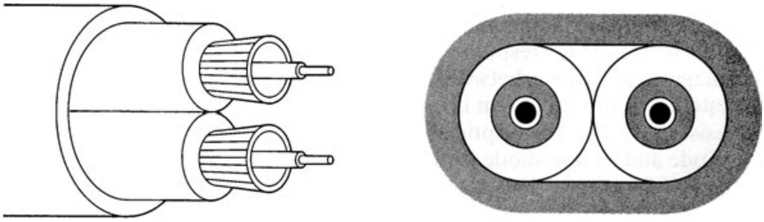


FIGURE 2.68 Heavy-duty fiber-optic cable.

Table 2.1 summarizes fiber-optic cable characteristics.

TABLE 2.1 Fiber-Optic Cable Characteristics

Characteristics	Simplex	Zipcord	Duplex	Channels 2	4	6
Cable diameter, mm	2.5	2.5 × 5.4	3.5 × 6	8	8	10
Weight, kg/km	6.5	11	20	41	41	61
Jacket type	PVC	PVC	PVC	PE	PE	PE
Jacket color	Orange	Orange	Orange	Black	Black	Black
Pull tension	330 N	490 N	670 N	890 N	1150 N	2490 N
Max. long-term tension	200 N	310 N	400 N	525 N	870 N	1425 N
Max. break strength	890 N	1340 N	1780 N	2370 N	4000 N	11,000 N
Impact strength at 1.6 N • m	100	100	150	200	200	200
Cyclic flexing, cycles	>5000	>5000	>5000	>2000	>2000	>2000
Minimum bend radius, mm	25	25	25	50	50	75
Cable attenuation	0.6 dB/km at 820 nm					
Operating temperature	–40 to +85°C					
Storage temperature	–40 to +85°C					

Fiber-Optic Ammeter

In many applications, including fusion reactors, radio-frequency systems, and telemetry systems, it is often necessary to measure the magnitude and frequency of current flowing through a circuit in which high DC voltages are present. A fiber-optic current monitor (Fig. 2.69) has been developed at the Princeton Plasma Physics Laboratory (PPPL) in response to a transient voltage breakdown problem that caused failures of Hall-effect devices used in the Tokamak fusion test reactor's natural-beam heating systems.

The fiber-optic current monitor measures low current in a conductor at very high voltage. Typical voltages range between tens of kilovolts and several hundred kilovolts. With a dead band of approximately 3 mA, the circuit derives its power from the conductor being measured and couples information to a (safe) area by means of fiber optics. The frequency response is normally from direct current to 100 kHz, and a typical magnitude range is between 5 and 600 mA.

The system is composed of an inverting amplifier, a current regulator, transorbs, diodes, resistors, and a fiber-optic cable. Around an inverting amplifier, a light-emitting diode and a photodiode form an optical closed feedback loop. A fraction of the light emitted by the LED is coupled to the fiber-optic cable.

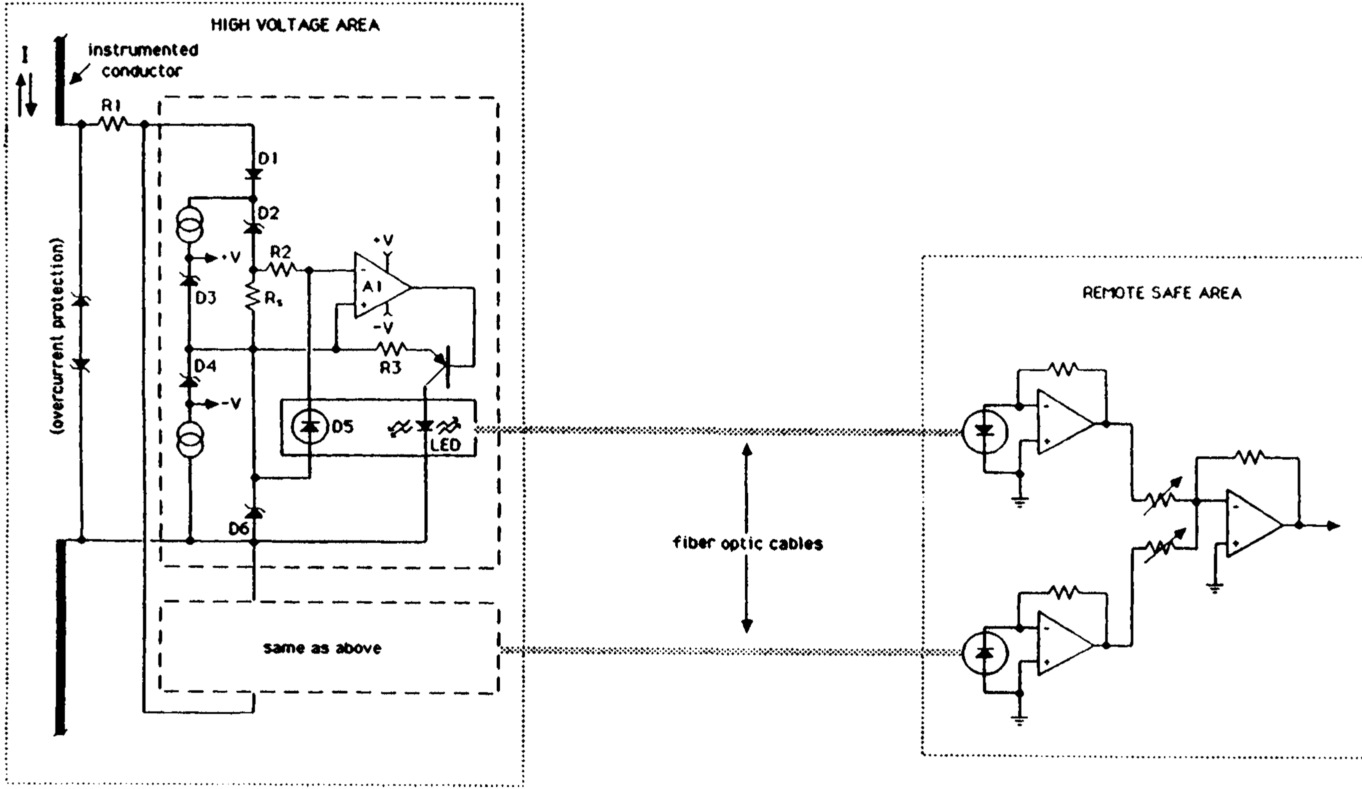


FIGURE 2.69 Fiber-optic current monitor.

As the current flows through the first diode, it splits between the 1.5-mA current regulator and the sampling resistor. The voltage across the sampling resistor causes a small current to flow into the inverting amplifier summing junction and is proportional to the current in the sampling resistor. Since photodiodes are quite linear, the light power from the LED is proportional to the current through the sampling resistor. The light is split between the local photodiode and the fiber cable. A photodiode, located in a remote safe area, receives light that is linearly proportional to the conductor current (for current greater than 5 mA and less than 600 mA).

To protect against fault conditions, the design utilizes two back-to-back transorbs in parallel with the monitor circuit. The transorbs are rated for 400 A for 1 ms. The fiber-optic ammeter is an effective tool for fusion research and other applications where high voltage is present.

REFERENCES

1. Berwick, M., J. D. C. Jones, and D. A. Jackson, "Alternating Current Measurement and Non-Invasive Data Ring Utilizing the Faraday Effect in a Closed Loop Fiber Magnetometer," *Optics Lett.*, 12(294) (1987).
2. Cole, J. H., B. A. Danver, and J. A. Bucaro, "Synthetic Heterodyne Interferometric Demodulation," *IEEE J. Quant. Electron.*, QE-18(684) (1982).
3. Dandridge, A., and A. B. Tveten, "Phase Compensation in Interferometric Fiber Optic Sensors," *Optics Lett.*, 7(279) (1982).
4. Desforges, F. X., L. B. Jeunhomme, Ph. Graindorge, and G. L. Baudec, "Fiber Optic Microswitch for Industrial Use," presented at SPIE O-E Fiber Conf., San Diego, no. 838–41 (1987).
5. Favre, F., and D. LeGuen, "High Frequency Stability in Laser Diode for Heterodyne Communication Systems," *Electron. Lett.*, 16(179) (1980).
6. Giallorenzi, T. G., "Optical Fiber Interferometer Technology and Hydrophones in Optical Fiber Sensors," NATO ASI Series E, No. 132, Martinus Nijhoff Dordrecht, 35–50 (1987).
7. Hocker, G. B., "Fiber Optic Sensing for Pressure and Temperature," *Appl. Optics*, 18(1445) (1979).
8. Jackson, D. A., A. D. Kersey, and A. C. Lewin, "Fiber Gyroscope with Passive Quadrature Demodulation," *Electron. Lett.*, 20(399) (1984).
9. Optical Society of America, *Optical Fiber Optics.* Summaries of papers presented at the Optical Fiber Sensors Topical Meeting, January 27–29, 1988, New Orleans, La. (IEEE, Catalog No. 88CH2524–7).
10. Popovic, R. S., "Hall Effect Devices," *Sensors and Actuators*, 17, 39–53 (1989).
11. Saxena, S. C., and S. B. Lal Seksena, "A Self-Compensated Smart LVDT Transducer," *IEEE Trans. Instrum. Meas.*, 38, 748–783 (1989).
12. Wong, Y. L., and W. E. Ott, *Function Circuits and Applications*, McGraw-Hill, New York, 1976.

This page intentionally left blank

CHAPTER 3
NETWORKING OF SENSORS AND CONTROL SYSTEMS IN MANUFACTURING

INTRODUCTION

Central to the development of any computer-integrated manufacturing facility is the selection of the appropriate automated manufacturing system and the sensors and control systems to implement it. The degree to which a CIM (computer-integrated manufacturing) configuration can be realized depends on the capabilities and cost of available equipment and the simplicity of information flow.

When designing an error-free manufacturing system, the manufacturing design group must have an appreciation for the functional limits of the automated manufacturing equipment of interest and the ability of the sensors to provide effective information flow, since these parameters will constrain the range of possible design configurations. Obviously, it is not useful to design a manufacturing facility that cannot be implemented because it exceeds the equipment's capabilities. It is desirable to match automated manufacturing equipment to the application. Although sensors and control systems are—by far—less costly than the automated manufacturing equipment, it is neither useful nor cost-effective to apply the most sophisticated sensors and controls, with the highest performance, to every possible application. Rather, it is important that the design process determines the preferred parameter values.

The preferred values must be compatible with available equipment and sensors and control systems, and should be those appropriate for the particular factory. The parameters associated with the available equipment and sensors and control systems drive a functional process of modeling the manufacturing operation and facility. The parameters determine how the real-world equipment constraints will be incorporated into the functional design process. In turn, as many different functional configurations are considered, the cost-benefit relations of these alternatives can be evaluated and preferred parameter values determined. So long as these preferred values are within the limits of available automated manufacturing equipment and sensory and control systems, the design group is assured that the automated manufacturing equipment can meet its requirements. To the degree that optimum design configurations exceed present equipment capabilities, original equipment manufacturers (OEMs) are motivated to develop new equipment designs and advanced sensors and control systems.

Sensors and control systems, actuators/effectors, controllers, and control loops must be considered in order to appreciate the fundamental limitations associated with manufacturing equipment for error-free manufacturing. Many levels of factory automation are associated with manufacturing equipment; the objective at all times should be to choose the levels of automation and information flow appropriate for the facility being designed, as revealed through cost-benefit studies. Manufacturing facilities can be designed by describing each manufacturing system—and the sensors and controls to be used in it—by a set of functional parameters. These parameters are:

- The number of product categories for which the automated manufacturing equipment, sensors, and control systems can be used (with software downloaded for each product type)
- The mean time between operator interventions (MTOI)
- The mean time of intervention (MTI)
- The percentage yield of product of acceptable quality
- The mean processing time per product

An *ideal equipment unit* would be infinitely flexible so it could handle any number of categories desired, would require no operator intervention between setup times, would produce only product of acceptable quality, and would have unbounded production capabilities.

The degree to which real equipment containing sensors and control systems can approach this ideal depends on the physical constraints associated with the design and operation of the equipment and the ability to obtain instantaneous information about equipment performance through sensors and control systems. The performance of the equipment in each of the five parameters stated earlier is related to the details of the equipment's operation in an error-free environment. Relationships must be developed between the physical description of the equipment's operation and the functional parameters that will be associated with this operation. The objective is to link together the physical design of the equipment and its functional performance through sensory and control systems in the factory setting.

This concept provides insight into an area in which future manufacturing system improvements would be advantageous, and also suggests the magnitude of the cost-benefit payoffs that might be associated with various equipment designs. It also reveals the operational efficiency of such systems.

An understanding of the relationships between the equipment characteristics and the performance parameters based on sensors and control systems can be used to select the best equipment for the parameter requirements associated with a given factory configuration. In this way, the manufacturing design team can survey alternative types of available equipment and select the units most appropriate for each potential configuration.

NUMBER OF PRODUCTS IN A FLEXIBLE SYSTEM

The first parameter listed earlier—the number of product categories for which the manufacturing system can be used—represents the key concern in flexible manufacturing. A unit of automated manufacturing equipment is described in terms of the number of product categories for which it can be used with only a software download to distinguish among product types. A completely fixed automated manufacturing system that cannot respond to computer control might be able to accommodate only one product category without a manual setup. On the other hand, a very flexible manufacturing system would be able to accommodate a wide range of product categories with the aid of effective sensors and control systems. This parameter will thus be defined by the breadth of the processes that can be performed by an automated manufacturing equipment unit and the ability of the unit to respond to external control data to shift among these operations.

The most effective solution will depend on the factory configuration that is of interest. Thus, OEMs are always concerned with anticipating future types of factories in order to ensure that their equipment will be an optimum match to the intended configuration. This will also ensure that the *concept of error-free manufacturing can be implemented with a high degree of spontaneity*. There is a continual trade-off between flexibility and cost. In general, more flexible and "smarter" manufacturing equipment will cost more. Therefore, the objective in a particular setting will be to achieve just the required amount of flexibility, without any extra capability built into the equipment unit.

SENSORS TRACKING THE MEAN TIME BETWEEN OPERATOR INTERVENTIONS

The MTOI value should be matched to the factory configuration in use. In a highly manual operation, it may be acceptable to have an operator intervene frequently. On the other hand, if the objective is to achieve operator-independent manufacturing between manual setups, then the equipment must be designed so the MTOI is longer than the planned duration between manual setups. The manufacturer of automated equipment with adequate sensors and control systems must try to assess the ways in which factories will be configured and produce equipment that can satisfy manufacturing needs without incurring any extra cost due to needed features.

SENSORS TRACKING THE MEAN TIME OF INTERVENTION

Each time an intervention is required, it is desirable to compare the intervention interval with that for which the system was designed. If the intervention time becomes large with respect to the planned mean time between operator interventions, then the efficiency of the automated manufacturing equipment drops rapidly in terms of the fraction of time it is available to manufacture the desired product.

SENSORS TRACKING YIELD

In a competitive environment, it is essential that all automated manufacturing equipment emphasize the production of quality product. If the automated manufacturing equipment produces a large quantity of product that must be either discarded or reworked, then the operation of the factory is strongly affected, and costs will increase rapidly. The objective, then, is to determine the product yields required for given configurations and to design automated manufacturing equipment containing sensors and control systems that can achieve these levels of yield. Achieving higher yield levels will, in general, require additional sensing and adaptability features for the equipment. These features will enable the equipment to adjust and monitor itself and, if it gets out of alignment, to discontinue operation.

SENSORS TRACKING THE MEAN PROCESSING TIME

If more product units can be completed in a given time, the cost of automated manufacturing equipment with sensors and control systems can be more widely amortized. As the mean processing time is reduced, the equipment can produce more product units in a given time, reducing the manufacturing cost per unit. Again, automated manufacturing equipment containing sensory and control systems generally becomes more expensive as the processing time is reduced. Tradeoffs are generally necessary among improvements in the five parameters and the cost of equipment. If high-performance equipment is to be employed, the factory configuration must make effective use of the equipment's capabilities to justify its higher cost. On the other hand, if the factory configuration does not require the highest parameters, then it is far more cost-effective to choose equipment units that are less sophisticated but adequate for the purposes of the facility. This interplay between parameter values and equipment design and cost is an essential aspect of system design.

Table 3.1 illustrates the difference between available parameter values and optimum parameter values, where the subscripts for equipment E represent increasing levels of complexity. The table shows the type of data that can be collected to evaluate cost and benefits. These data have a significant impact on system design and performance which, in turn, has a direct impact on product cost. Given the type of information in Table 3.1, system designers can evaluate the effects of utilizing various levels of sensors and control systems on new equipment and whether they improve performance enough to be worth the research and development and production investment.

One of the difficulties associated with manufacturing strategies in the United States is that many companies procure manufacturing equipment only from commercial vendors and do not consider modifying it to suit their own needs. Custom modification can produce pivotal manufacturing advantage, but also requires the company to expand both its planning scope and product development skills. The type of analysis indicated in Table 3.1 may enable an organization to determine the value and return on investment of customizing manufacturing equipment to incorporate advanced sensors and control systems. Alternatively, enterprises with limited research and development resources may decide to contract for development of the optimum equipment in such a way that the sponsor retains proprietary rights for a period of time.

TABLE 3.1 Values of Available Parameters

Equipment	MTOI, min	MTI, min	Yield, %	Process, min	R&D expense, thousands of dollars	Equipment cost, thousands of dollars
			Production function A			
E_1	0.1	0.1	90	12		50
E_2	1.0	0.1	85	8		75
E_3	10	1.0	80	10		85
E_4	18	1.0	90	8	280	155
			Production function B			
E_1	1.0	0.1	95	10		150
E_2	10	0.5	90	2		300
			Production function C			
E_1	0.1	0.1	98	3		125
E_2	5.0	1.0	98	2		250
E_3	8.0	2.0	96	1		300
E_4	20	2.0	96	1	540	400

NETWORK OF SENSORS DETECTING MACHINERY FAULTS

A comprehensive detection system for automated manufacturing equipment must be seriously considered as part of the manufacturing strategy. A major component of any effort to develop an intelligent and flexible automatic manufacturing system is the concurrent development of automated diagnostic systems, with a network of sensors, to handle machinery maintenance and process control functions. This will undoubtedly lead to significant gains in productivity and product quality. Sensors and control systems are one of the enabling technologies for the "lights-out" factory of the future.

A flexible manufacturing system often contains a variety of manufacturing work cells. Each work cell in turn consists of various workstations. The flexible manufacturing cell may consist of a CNC lathe or mill whose capabilities are extended by a robotic handling device, thus creating a highly flexible machining cell whose functions are coordinated by its own computer. In most cases, the cell robot exchanges workpieces, tools (including chucks), and even its own gripping jaws in the cell (Fig. 3.1).

Diagnostic Systems

A diagnostic system generally relies on copious amounts of *a priori* and *a posteriori* information. *A priori* information is any previously established fact or relationship that the system can exploit in making a diagnosis. *A posteriori* information is the information concerning the problem at hand for which the diagnosis will be made. The first step in collecting data is to use sensors and transducers to convert physical states into electrical signals. After processing, a signal will be in an appropriate form for analysis (perhaps as

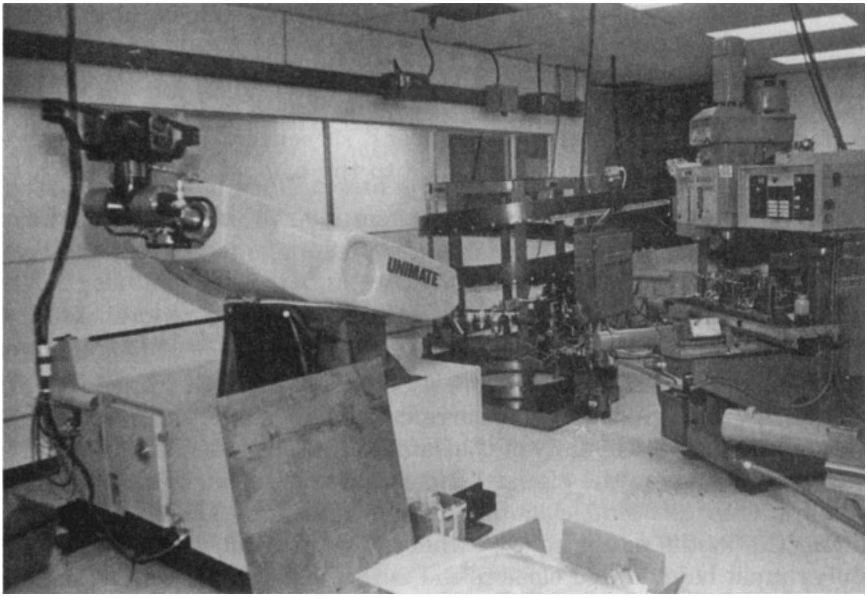


FIGURE 3.1 Flexible machining cell.

a table of values, a time-domain waveform, or a frequency spectrum). Then the analysis, including correlations with other data and trending, can proceed.

After the data have been distilled into information, the deductive process begins, leading finally to the fault diagnosis. Expert systems have been used effectively for diagnostic efforts, with the diagnostic system presenting either a single diagnosis or a set of possibilities with their respective probabilities, based on the *a priori* and *a posteriori* information.

Resonance and Vibration Analysis

Resonance and vibration analysis is a proven method for diagnosing deteriorating machine elements in steady-state process equipment such as turbomachinery and fans. The effectiveness of resonance and vibration analysis in diagnosing faults in machinery operating at variable speed is not proved, but additional study has indicated good potential for its application in robots. One difficulty with resonance and vibration analysis is the attenuation of the signal as it travels through a structure on the way to the sensors and transducers. Moreover, all motion of the machine contributes to the motion measured by the sensors and transducers, so sensors and transducers must be located as close as possible to the component of concern to maximize the signal-to-noise ratio.

Sensing Motor Current for Signature Analysis

Electric motors generate back electromotive force (emf) when subjected to mechanical load. This property makes a motor a transducer for measuring load vibrations via current fluctuations. Motor current signature analysis uses many of the same techniques as vibration analysis for interpreting the signals. But motor current signature analysis is nonintrusive because motor current can be measured anywhere along the motor power cables, whereas a vibration sensor or transducer must be mounted close to the machine element of interest. The limited bandwidth of the signals associated with motor drive signature analysis, however, may restrict its applicability.

Acoustics

A good operator can tell from the noise that a machine makes whether a fault is developing or not. It is natural to extend this concept to automatic diagnosis. The operator, obviously, has access to subtle, innate pattern recognition techniques, and thus is able to discern sounds within myriad background noises. Any diagnostic system based on sound would have to be able to identify damage-related sounds and separate the information from the ambient noise. Acoustic sensing (looking for sounds that indicate faults) is a nonlocal noncontact inspection method. Any acoustic technique is subject to outside disturbances, but is potentially a very powerful tool, provided that operating conditions are acoustically repeatable and that the diagnostic system can effectively recognize acoustic patterns.

Temperature

Using temperature as a measurement parameter is common, particularly for equipment running at high speed, where faults cause enough waste heat to raise temperature significantly. This method is generally best for indicating that a fault has occurred, rather than the precise nature of the fault.

Sensors for Diagnostic Systems

Assuming an automated diagnostic system is required, the necessary sensors are normally mounted permanently at their monitoring sites. This works well if data are required continuously, or if there are only a few monitoring locations. However, for those cases where many sites must be monitored and the data need not be continuously received during operation of the flexible manufacturing cell, it may be possible to use the same sensor or transducer, sequentially, in the many locations.

The robot is well-suited to gather data at multiple points with a limited number of sensors and transducers. This would extend the mandate of the robot from simply moving workpieces and tools within the flexible manufacturing cell (for production) to include moving sensors (for diagnostic. inspection).

Within the flexible manufacturing cell, a robot can make measurements at sites inside its work space by taking a sensor or transducer from a tool magazine, delivering the sensor or transducer to a sensing location, detaching it during data collection, and then retrieving it before moving to the next sensing position.

Sensor mobility does, however, add some problems. First, the robot will not be able to reach all points within the flexible manufacturing cell because its work space is only a subspace of the volume taken up by the flexible manufacturing cell. The manipulator may be unable to assume an orientation desired for a measurement even inside its work space. Also, the inspection procedure must limit the robot's influence on the measurement as much as possible. Finally, sensors require connectors on the robot end effectors for signals and power. The end effector would have to be able to accommodate all the types of sensors to be mounted on it.

Quantifying the Quality of a Workpiece

If workpiece quality can be quantified, then quality can become a process variable. Any system using product quality as a measure of its performance needs tight error checks so as not to discard product unnecessarily while the flexible manufacturing cell adjusts its operating parameters. Such a system would depend heavily, at first, on the continued supervision of an operator who remains in the loop to assess product quality. Since it is forbidden for the operator to influence the process while it is under automatic control, it is more realistic for the operator to look for damage to product after each stage of manufacture within the cell. In that way, the flexible manufacturing cell receives diagnostic information about product deficiencies close to the time that improper manufacture occurred.

In the future, these quality assessments will be handled by the flexible manufacturing cell itself, using sensors and diagnostic information for process control. Robots, too, will be used for maintenance and physical inspection as a part of the regular operation of the flexible manufacturing cell. In the near term, the flexible manufacturing cell robot may be used as a sensor-transfer device, replacing inspectors who would otherwise apply sensors to collect data.

Evaluation of an Existing Flexible Manufacturing Cell Using a Sensing Network

A study was conducted at the Mi-TNO in the Netherlands of flexible manufacturing cells for low-volume orders (often called job production, ranging from 1 to 100 parts per order). The automated manufacturing equipment used in the study consisted of two free-standing flexible manufacturing cells. The first cell was a turning-machine cell; the second, a milling-machine cell. The turning cell contained two armed gantry robots for material handling. The study was mainly conducted to assess the diagnostics for flexible manufacturing systems

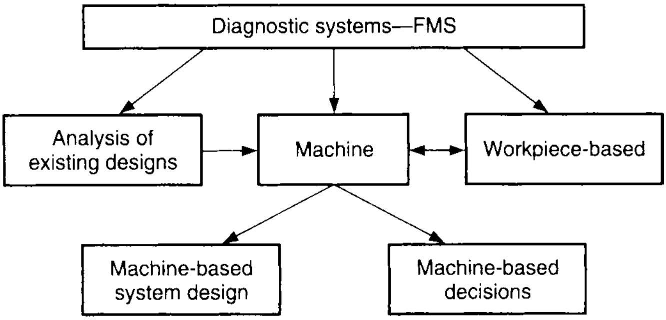


FIGURE 3.2 Diagnostics for an FMS.

(FMS). In considering the approach to setting up diagnostics for an FMS, it was decided to divide development of the diagnostics program into three major blocks (Fig. 3.2):

- Analyzing the existing design
- Setting up diagnostics that are machine-based
  - Choosing important points in the flexible manufacturing cells where critical failure can occur and where sensors are mounted
  - Setting up a diagnostic decision system for the hardware system
- Establishing a workpiece-based diagnostic system that is actually part of quality control

The flexible manufacturing cells were divided into control volumes (as shown in Fig. 3.3 for the turning cell). Regarding the turning cell, for example, the hardware control volumes

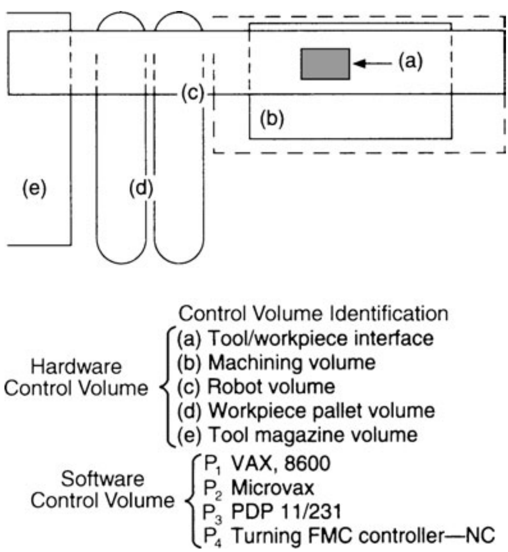


FIGURE 3.3 Hardware control volumes in a manufacturing cell.

were denoted *a*, *b*, *c*, *d*, and *e*, and software control volumes p1, p2, p3, and p4. The failures within each of these control volumes were further categorized as:

- Fault detected by an existing sensor
- Fault that could have been detected if a sensor had been present
- Operator learning phase problem
- Failure due to manufacturer problem
- Software logic problem
- Repeat of a problem
- System down for an extended period

Software Problems. It is often assumed that disturbances in the cell software control system can be detected and evaluated relatively easily. Software diagnostics are common in most turnkey operations; however, it has been shown that software diagnostics are far from perfect. Indeed, software problems are of particular concern when a revised program is introduced into a cell that has been running smoothly (existing bugs having been ironed out). The availability of the cell plummets in these situations, with considerable loss of production capabilities and concomitant higher cost. The frequency of software faults increases dramatically when revised packages are introduced to upgrade the system (Fig. 3.4). This

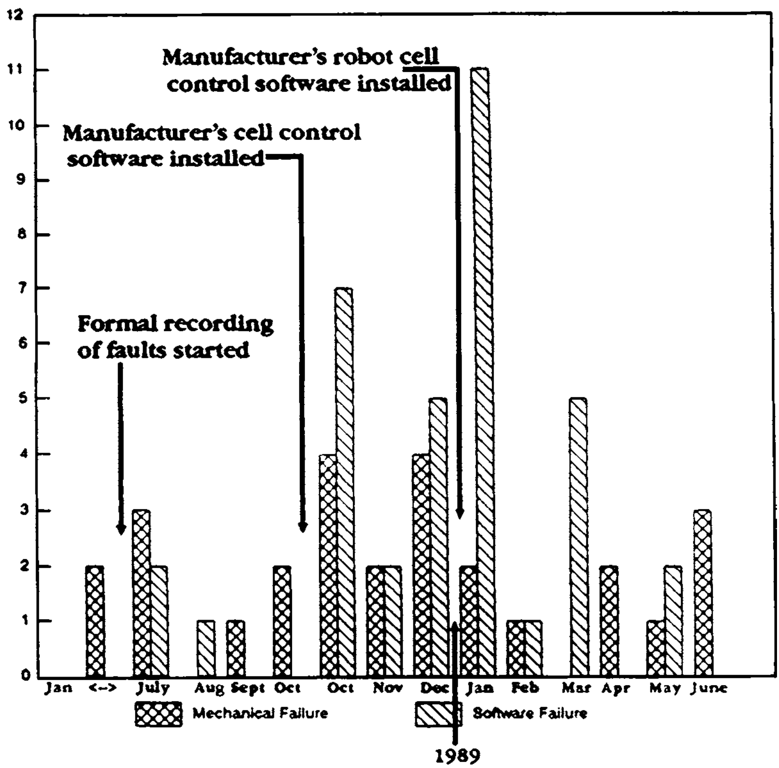


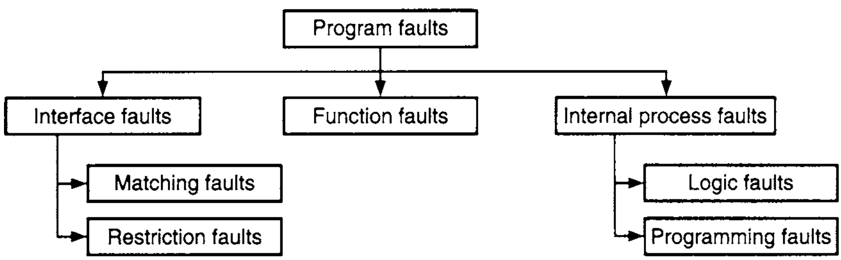
FIGURE 3.4 Frequency of software faults.

is mainly due to human error on the part of either the vendor or the operator. Two possible approaches to software defect prevention and detection are:

- Investigate software methodologies and procedures and recommend alternative languages or more tools as defect prevention measurements. This is tedious and uses ambiguous results because such investigations are not based on data.
- Analyze the problems that result from the current design and develop a solution for each class of problem. This produces less ambiguous solutions and is typically used to solve only immediate problems, thereby producing only short-term solutions.

To identify the types of faults that occur in programs, it is necessary to know what caused the problem and what remedial actions were taken. Program faults can be subdivided into the categories shown next and restated in Fig. 3.5:

- Wrong names of global variables or constants
- Wrong type of structure or module arguments



Matching faults

- Wrong names of global variables or constants
- Wrong type of structure or module arguments
- Wrong number of hardware units
- Wrong procedures for writing data to hardware

Restriction faults

- Omission of procedures to prevent invalid input or output of data
- Wrong limit value for validity check of arguments

Function faults

- Omission of saving data to global variables
- Unnecessary calling modules
- Wrong limit value for judging whether or not hardware is set
- Reference to undefined local variables
- Omission of loop variable incrementation
- Logic expressions that are always true

Programming faults

- Comparison of local variables of different types
- Omission of comment marks

FIGURE 3.5 Program fault categories.

- Wrong number of hardware units
- Wrong procedures for writing data to hardware

Restriction faults

- Omission of procedures to prevent invalid input or output of data
- Wrong limit value for validity check of arguments

Function faults

- Omission of saving data to global variables
- Unnecessary calling modules
- Wrong limit value for judging whether or not hardware is set
- Reference to undefined local variables
- Omission of loop variable incrementation
- Logic expressions that are always true

Programming faults

- Comparison of local variables of different types
- Omission of comment marks

This categorization provides significant insight into the location of fault conditions, the reasons for their occurrence, and their severity. If the faults are in either a hardware (mechanical) or software category, then the frequency of failure by month can be summarized as indicated in Fig. 3.4. Two major and unexpected milestones in the program represented in Fig. 3.4 are the routine introduction of revised cell control software and revised robot control software. In both cases, a substantial increase occurred in the downtime of the flexible manufacturing cell. In an industrial environment, this would have been very costly.

In this study, it was found that interface faults (mismatched data transfer between modules and hardware) were the major cause of downtime. Also, in the new machine software, it was found that faults occurred because the software had not been properly matched to the number of tool positions physically present on the tool magazine. Once, such a fault actually caused a major collision within the machining volume.

Detecting Tool Failure. An important element in automated process control is real-time detection of cutting tool failure, including both wear and fracture mechanisms. The ability to detect such failures on line allows remedial action to be undertaken in a timely fashion, thus ensuring consistently high product quality and preventing potential damage to the process machinery. The preliminary results from a study to investigate the possibility of using vibration signals generated during face milling to detect both progressive (wear) and catastrophic (breakage) tool failure are discussed next.

Experimental Technique. The experimental studies were carried out using a 3-hp vertical milling machine. The cutting tool was a 381-mm-diameter face mill employing three Carboloy TPC-322E grade 370 tungsten carbide cutting inserts. The standard workpiece was a mild steel plate with a length of 305 mm, a height of 152 mm, and a width of 13 mm. While cutting, the mill traversed the length of the workpiece, performing an interrupted symmetric cut. The sensor sensed the vibration generated during the milling process on the workpiece clamp. The vibration signals were recorded for analysis. Inserts with various magnitudes of wear and fracture (ranging from 0.13 mm to 0.78 mm) were used in the experiments.

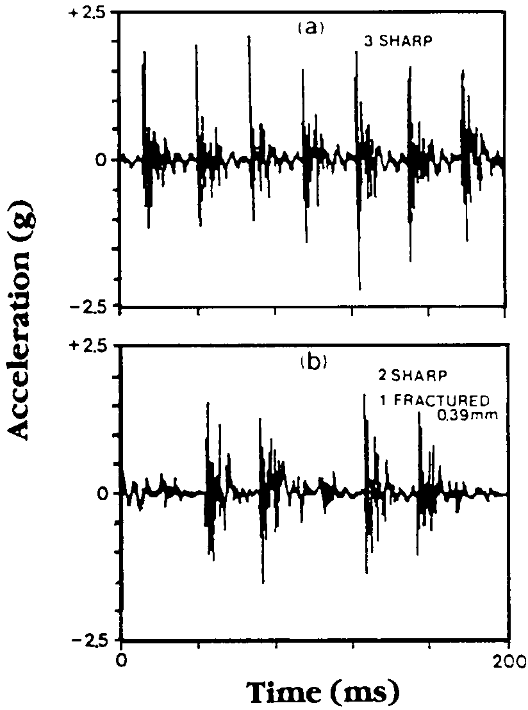


FIGURE 3.6 Typical acceleration level versus time plot.

Manufacturing Status of Parts. Figure 3.6 shows typical acceleration versus time histories. Figure 3.7*a* is the acceleration for three sharp inserts. Note that the engagement of each insert in the workpiece is clearly evident and that all engagements share similar characteristics, although they are by no means identical.

Figure 3.7*b* shows the acceleration for the combination of two sharp inserts and one insert with a 0.39-mm fracture. The sharp inserts produce signals consistent with those shown in Fig. 3.6, while the fractured insert produces a significantly different output.

The reduced output level for the fractured insert is a result of the much smaller depth of cut associated with it. It would seem from the time-domain data that use of either an envelope detection or a threshold crossing scheme would provide the ability to automate the detection of tool fracture in a multi-insert milling operation.

Figure 3.7 shows typical frequency spectra for various tool conditions. It is immediately apparent that, in general, fracture phenomena are indicated by an increase in the level of spectra components within the range of 10,000 to 17,000 Hz. A comparison of Fig. 3.7*a* and *b* indicates a notable increase in the spectra around 11 kHz when a single insert fracture of 0.39 mm is present. For two fractured inserts (Fig. 3.7*c*), the peak shifts to around 13.5 kHz. For three fractured inserts (Fig. 3.7*d*), both the 13.5-kHz peak and an additional peak at about 17 kHz are apparent.

Comparing Figs. 3.7*e* and *c*, it is seen that, in general, increasing the size of the insert fracture results in an increase in the spectral peak associated with the failure condition. Usually, the assemblies used to obtain the spectral data are not synchronously averaged. Therefore, it may be possible to improve the spectral data by utilizing a combination of

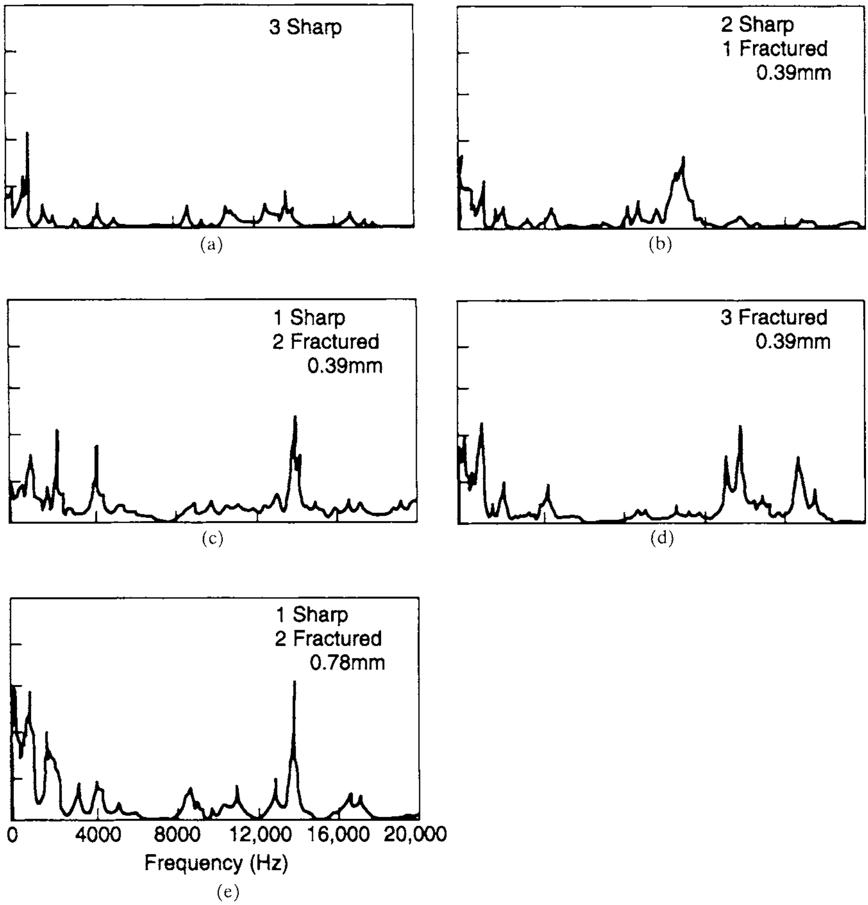


FIGURE 3.7 Acceleration level versus time for various inserts.

synchronous averaging and delayed triggering to ensure that data representative of each insert in the cutter is obtained and processed.

In general, the acceleration-time histories for the worn inserts do not produce noticeably different engagement signals evident in a case of fracture. However, by processing the data in a slightly different manner, it is possible to detect evidence of tool wear.

Figure 3.8 shows the amplitude probability density (APD) as a function of time for several tool conditions. The data are averaged for eight assemblies. It thus seems possible that insert wear could be detected using such features as the location of the peak in the APD, the magnitude of the peak, and the area under specific segments of the distribution.

As with fracture, the presence of insert wear resulted in a significant increase in the spectral components within the 10- to 13-kHz band. Although this would seem to indicate that the presence of flank wear could be detected by simple spectral analysis, it is not yet clear if this method would be sufficiently discriminating to permit reliable determination of the magnitude of flank wear.

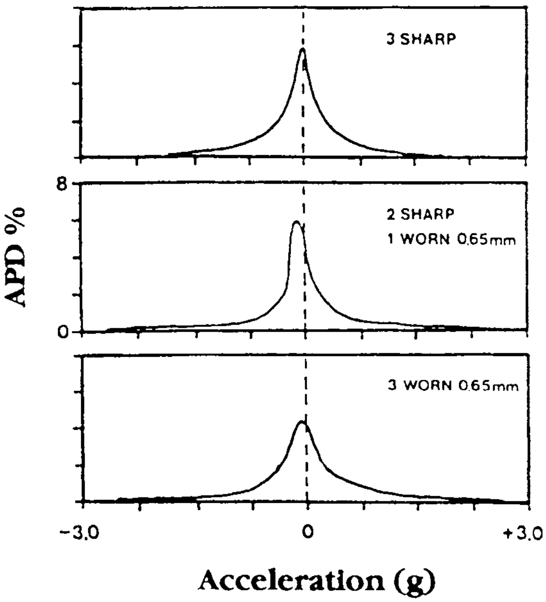


FIGURE 3.8 Amplitude probability density (APD) for several tool conditions.

UNDERSTANDING COMPUTER COMMUNICATIONS AND SENSORS' ROLE

The evolution in computer software and hardware has had a major impact on the capability of computer-integrated manufacturing concepts. The development of smart manufacturing equipment and sensors and control systems, as well as methods of networking computers, has made it feasible to consider cost-effective computer applications that enhance manufacturing. In addition, this growth has changed approaches to the design of manufacturing facilities.

Messages are exchanged among computers according to various protocols. The open system interconnect (OSI) model developed by the International Standards Organization (ISO) provides such a framework. Figure 3.9 illustrates how two users might employ a computer network. As illustrated, the transfer of information takes place from User 1 to User 2. Each message passes through a series of layers associated with message processing:

Sequence of layers—User 1	Sequence of layers—User 2
(PhL)1 Physical layer	(PhL)1 Physical layer
(DL)2 Data link layer	(DL)2 Data link layer
(NL)3 Network layer	(NL)3 Network layer
(TL)4 Transport layer	(TL)4 Transport layer
(SL)5 Session layer	(SL)5 Session layer
(PL)6 Presentation layer	(PL)6 Presentation layer
(AL)7 Application layer	(AL)7 Application layer

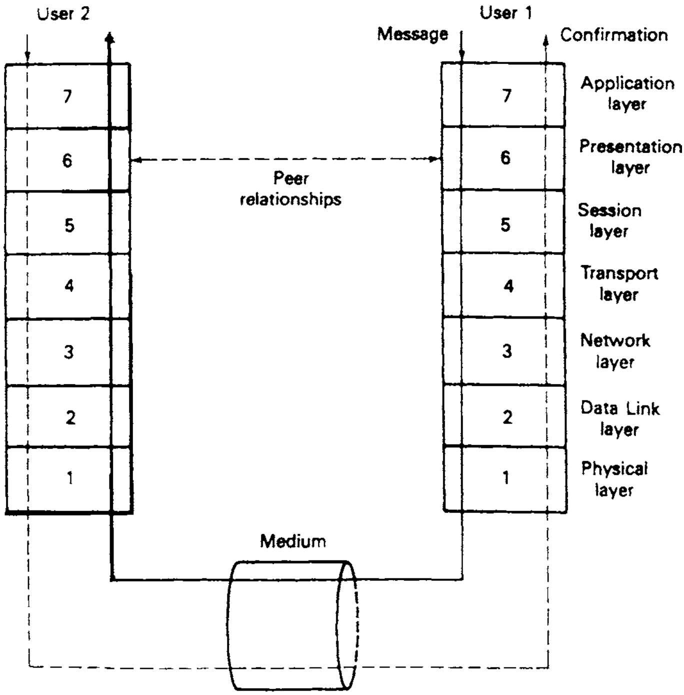


FIGURE 3.9 Message movement within a computer network.

A message is sent from User 1 to User 2, with message acknowledgment sent back from User 2 to User 1. The objective of this communication system is to transfer a message from User 1 to User 2 and to confirm message receipt. The message is developed at application layer (AL)7, and passes from there to presentation layer (PL)6, from there to session layer (SL)5, and so forth until the message is actually transmitted over the communication path. The message arrives at physical layer (PhL)1 for User 2 and then proceeds from physical layer (PhL)1 to data link layer (DL)2, to network layer (NL)3, and so forth, until User 2 has received the message. In order for Users 1 and 2 to communicate with one another, every message must pass though all the layers.

The layered approach provides a structure for the messaging procedure. Each time a message moves from User 1 [application layer (AL)7 to physical layer (PhL)1] additional processing and addressing information is added to the beginning or end of the message. As the original message moves on, new information is added to ensure correct communication. Then, as the message moves to User 2 [from physical layer (PhL)1 to application layer (AL)7], this additional information is removed, as illustrated in Fig. 3.9.

The layers work together to achieve "peer" communication. Any information or operation instruction added at session layer (SL)5 for User 1 will be addressing session layer (SL)5 for User 2. The layers thus work as peers; each layer has a given operational or addressing task to make sure the message is correctly communicated from User 1 to User 2. Each layer associates only with the layers above and below itself. The layer receives messages from one direction, processes the messages, and then passes them on to the next layer.

Application Layer Communication

Communication begins when User 1 requests that a message be transferred from location 1 to location 2. In developing this message, it may be necessary to have application software that will provide supporting services to the users. This type of service is provided by application layer (AL)7. Common software application tools enable, for example, the transfer of files or the arranging of messages in standard formats.

Presentation Layer Communication

The initial message is passed down from the application layer (AL)7 to presentation layer (PL)6, where any necessary translation is performed to develop a common message syntax that User 2 will understand. If the two different users apply different computer or equipment languages, it will be necessary to define the differences in such a way that the users can communicate with one another. The basic message that began with User 1 is translated to a common syntax that will result in understanding by User 2. Additional information is added to the message at presentation layer (PL)6 to explain to User 2 the nature of the communication that is taking place. An extended message begins to form.

Session Layer Communication

The message now passes from presentation layer (PL)6 to session layer (SL)5. The objective of the session layer is to set up the ability for the two users to converse, instead of just passing unrelated messages back and forth. The session layer will remember that there is an ongoing dialog and will provide the necessary linkages between the individual messages so an extended transfer of information can take place.

Transport Layer Communication

The message then passes from session layer (SL)5 to transport layer (TL)4, which controls the individual messages as part of the communication sequence. The purpose of the transport layer is to make sure individual messages are transferred from User 1 to User 2 as part of the overall communication session defined by session layer (SL)5. Additional information is added to the message so the transport of this particular portion of the communication exchange results.

Network Layer Communication

The message is then passed to network layer (NL)3, divided into packets, and guided to the correct destination. The network layer operates so that the message (and all accompanying information) is tracked and routed correctly so it will end up at User 2. This includes making sure addresses are correct and that any intermediate stops between User 1 and User 2 are completely defined.

Data Link Layer Communication by Fiber Optics or Coaxial Cable

The system now passes a message to data link layer (DL)2, which directs each frame of the message routing in transit. Each frame is loaded into the communication system in preparation for transmission. The message is prepared to leave the User 1 environment and to move into the communication medium. The next step is from data link layer (DL)2 to physical layer (PhL)1. At this point, the frame is converted into a series of digital or analog electronic signals that can be placed on the communication medium itself—wire, fiber optics, coaxial cables, or other means—to achieve the transmission of the message from User 1 to User 2.

Physical Layer Communication

The electronic signal arrives at the correctly addressed location for User 2 and is received by physical layer (PhL)1. The physical layer then converts the electronic signal back to the original frame that was placed on the medium. This extended message is passed up to the data link layer, which confirms that error-free communication has taken place and that the frame has been received at the correct location. Data link layer (DL)2 directs the frame routing in transit. When the full frame is received at data link layer (DL)2, the routing information is removed and the remaining information is transferred to network layer (NL)3. Network layer (NL)3 confirms the appropriate routing and assembles the packets. Then, the routing information is stripped from the message. The message is then passed to transport layer (TL)4, which controls the transport of the individual messages. Transport layer (TL)4 confirms that the correct connection has been achieved between User 1 and User 2 and receives the information necessary to achieve the appropriate connection for this particular exchange.

The remaining information is often passed to session layer (SL)5, which interprets whether the message is part of a continuing communication and, if so, identifies it as part of an ongoing conversation. The information is then passed to presentation layer (PL)6, which performs any necessary translation to make sure User 2 can understand the message as it has been presented. The message is then passed to application layer (AL)7, which identifies the necessary support programs and software necessary for interpretation of the message by User 2. Finally, User 2 receives the message and understands its meaning.

Adding and Removing Information in Computer Networks Based on Open System Interconnect (OSI)

This step-by-step passing of the message from User 1 "down" to the actual communication medium and "up" to User 2 involves adding information to the original message prior to the transfer of information and removing this extra information on arrival (Fig. 3.10). At User 1, additional information is added step by step until the communication medium is reached, forming an extended message. When this information arrives at User 2, the additional information is removed step by step until User 2 receives the original message. As noted, a peer relationship exists between the various levels. Each level communicates only with the level above or below it, and the levels perform a matching service. Whatever information is added to a message by a given level at User 1 is removed from the message by the matching level associated with User 2. Communication is made possible by having User 1's physical layer (PhL)1 communicate with User 2's physical layer (PhL)1, data link layer (DL)2 communicate with data link (DL)2, network layer (NL)3 communicate with network layer (NL)3, and so forth to achieve an exchange between User 1 and User 2.

The two users see only the message that is originated and delivered; they do not see all of the intermediate steps. This is analogous to the steps involved in making a telephone call or sending a letter, in which the two users know only that they have communicated with one another, but do not have any specific knowledge of the details involved in passing a message

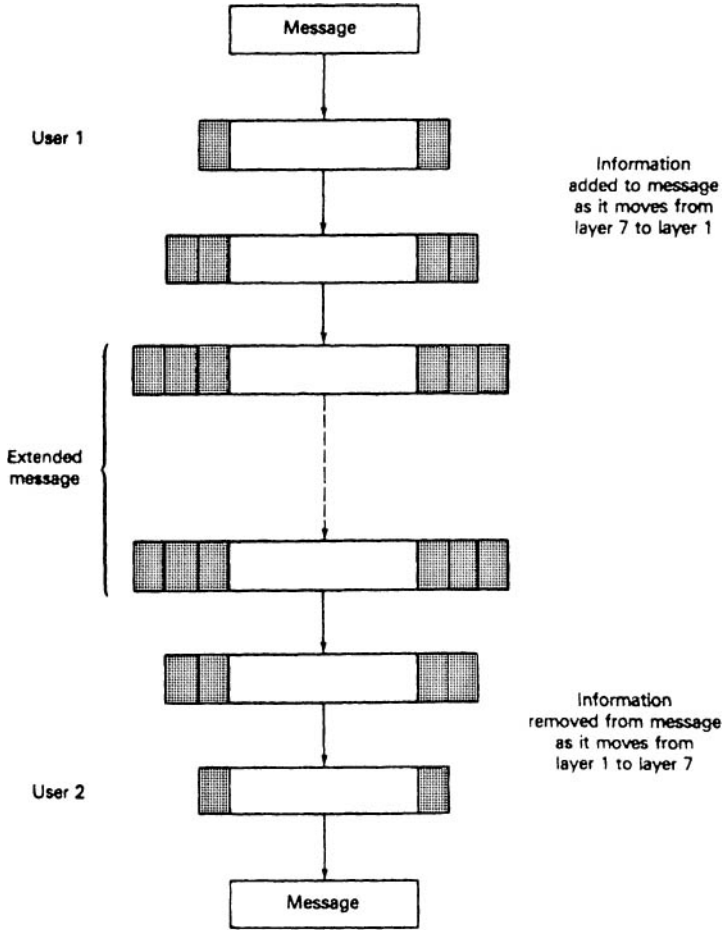


FIGURE 3.10 Open system interconnect (OSI) layered model.

from one location to another. This orderly and structured approach to communications is useful because it separates the various tasks that must take place. It provides a means for assuring that the methods for processing and addressing messages are always the same at every node. Whether a message is being sent or is being received, a sequential processing activity always takes place.

UNDERSTANDING NETWORKS IN MANUFACTURING

A seemingly simple problem in the development of computer networks is to establish the ability to interconnect *any* two computer system elements. This might involve a computer terminal, a modem, a bar-code scanner, a printer, sensors, and other system elements that must exchange information in manufacturing. It might seem reasonable that such one-to-one interconnection would follow a well-defined strategy and make maximum use of standards. Unfortunately, for historical reasons and because of the wide diversity of the equipment units available today, this situation does not hold. In fact, achieving interconnection between typical system elements can be a disconcerting experience.

RS-232–Based Networks

One of the most common approaches used to interconnect computer system elements in manufacturing is associated with a strategy that was never really intended for this purpose. As noted by Campbell (1984), "In 1969, the EIA (Electronic Industries Association), Bell Laboratories, and manufacturers of communication equipment cooperatively formulated and issued 'EIA RS-232.'" Almost immediately it underwent minor revisions to become "RS-232-C." The RS-232 interface was developed to allow data equipment terminals to be connected to modems so that data could be transmitted over the telephone network. The entire purpose of this standard was to assure that the use of telephone lines for computer communications would be handled in a way acceptable to the telephone company.

Thus, in its general application today, RS-232 is not a standard. It is more a guideline for addressing some of the issues involved in interfacing equipment. Many issues must be resolved on an individual basis, which leads to the potential for difficulties. Essentially, a vendor's statement that a computer system element is RS-232–compatible provides a starting point to consider how the equipment unit might be interconnected. However, the detailed aspects of the interconnection require further understanding of the ways in which equipment units are intended to communicate. Campbell (1984) is a helpful introduction to applying RS-232 concepts.

In a sense, the history of the RS-232 interface illustrates the difficulties associated with creating a well-defined means for allowing the various elements of a computer network to interface. Past experience also indicates how difficult it is to develop standards that will apply in the future to all the difficult situations that will be encountered. As it has evolved, the RS-232 approach to the communications interface is an improvement over the "total anarchy" at position A in Fig. 3.11, but it still leads to a wide range of problems.

An entire computer network can be configured by using combinations of point-to-point RS-232 connections. In fact, a number of networks of this type are in common use. Such networks require that multiple RS-232 interfaces be present on each equipment unit, as is often the case. Each particular interconnection must be customized for the two equipment units being considered. Thus, a system integrator not only must decide on the elements of the system and how they should perform in a functional sense, but also must develop a detailed understanding of the ways in which RS-232 concepts have been applied to the particular equipment units used for the network. The system integrator incurs a substantial expense in achieving the required interconnections.

It is essential to realize that the RS-232 pseudo-standard only addresses the ability to transfer serially information bit by bit from one system element to another. The higher level communications protocol in Fig. 3.11 is not considered. RS-232 provides the means for running wires or fiber-optic cables from one element to another in order to allow digital signals to be conveyed between system elements. The meaning associated with these bits of information is completely dependent on the hardware and software that is implemented in the system elements. RS-232 is a widely used approach for computer elements to transfer information. Using RS-232 is certainly much better than using no guidelines at all. However, because RS-232 does not completely define all of the relationships that must exist in communication links, it falls far short of being a true standard or protocol.

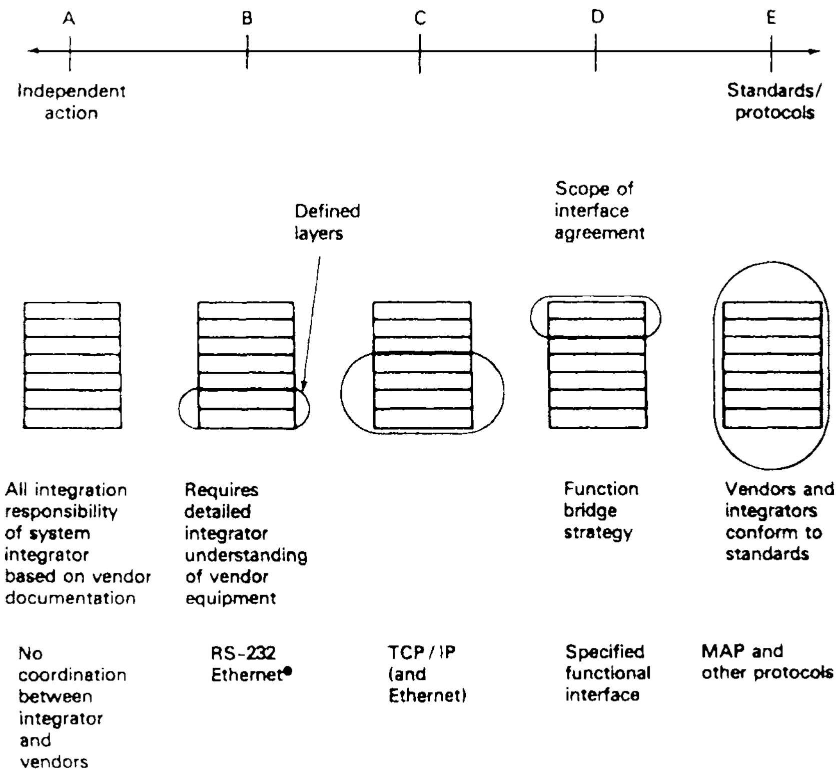


FIGURE 3.11 Levels of integration of computer networks.

Ethernet*

As illustrated in Fig. 3.11, one approach to local area networks (LANs) is to define a protocol for the first two layers of a communication strategy and then allow individual users to define the upper layers. This approach has been widely applied using a method referred to as *Ethernet*[7,8,9].

In every computer communication system, there must be a means of scheduling for each node to transmit onto the network and listen to receive messages. This may be done on a statistical basis. For example, when a unit needs to transmit over the network, it makes an effort to transmit. If another node tries to transmit at the same time, both nodes become aware of the conflict, wait for a random length of time, and try again. It might seem that this would be an inefficient means of controlling a network, since the various nodes are randomly trying to claim the network for their own use, and many collisions may occur. As it turns out, for lower communication volumes, this method works very well. As the number of nodes on the system and the number of messages being exchanged increases, however, the number of collisions between active nodes goes up and reduces the effectiveness of the system (Fig. 3.12).

*Ethernet is a trademark of Xerox Corp.

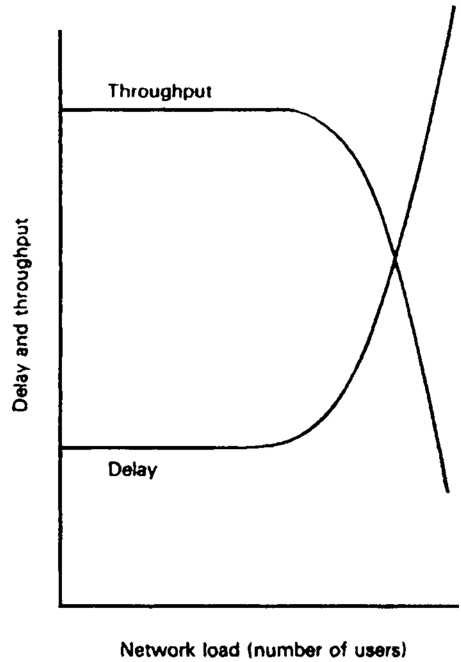


FIGURE 3.12 Delay and throughput versus network load (number of users).

This type of access control for a computer network is referred to as *carrier-sense multiple-access with collision detection* (CSMA/CD). Ethernet and similar solutions are widely applied to create CSMA/CD networks, particularly in settings in which a maximum waiting time for access to the network does not have to be guaranteed. This type of network is simple to install, and a wide range of hardware and software products are available for support. On the other hand, as indicated in Fig. 3.12, network performance can degrade significantly under high load; therefore, the utility of an Ethernet-oriented network will depend on the particular configuration and loads expected for the network.

Transmission Control Protocol (TCP)/Internet Protocol (IP)

TCP/IP applies to the transport and network layers indicated in Fig. 3.9. TCP/IP thus provides a means for addressing intermediate protocol levels, and in fact is often combined with Ethernet in a communication approach that defines both the lower and middle aspects of the system. TCP/IP functions by dividing any message provided to these middle layers into *packets* of 64 kbytes and then sending packets one at a time to the communication network. TCP/IP must also reassemble the packets in the correct order at the receiving user.

TCP/IP provides a common strategy to use for networking. It allows extension of the Ethernet lower layers to a midlayer protocol on which the final application and presentation layers may be constructed.

MANUFACTURING AUTOMATION PROTOCOL

The manufacturing automation protocol (MAP) is one of the protocols that has been developed for computer communication systems. MAP was developed specifically for use in a factory environment. General Motors Corp. has been the leading advocate of this particular protocol. When faced with a need for networking many types of equipment in its factory environment, General Motors decided that a new type of protocol was required. Beginning in 1980, General Motors began to develop a protocol that could accommodate the high data rate expected in its future factories and provide the necessary noise immunity expected for this environment. In addition, the intent was to work within a mature communications technology and to develop a protocol that could be used for all types of equipment in General Motors factories. MAP was developed to meet these needs. The General Motors' effort has drawn on a combination of Institute of Electrical and Electronics Engineers (IEEE) and ISO standards, and is based on the open system interconnect (OSI) layered model, as illustrated in Fig. 3.10.

Several versions of MAP have been developed. One difficulty of several has been obtaining agreement among many different countries and vendor groups on a specific standard. Another problem is that the resulting standards are so broad that they have become very complex, making it difficult to develop the hardware and software to implement the system and consequently driving up related costs. The early version of MAP addressed some of the OSI layers to a limited degree, and made provision for users to individualize the application layer for a particular use. The latest version of MAP makes an effort to define more completely all the application layer software support as well as the other layers. This has led to continuing disagreements and struggles to produce a protocol that can be adopted by every vendor group in every country to achieve MAP goals.

Because of its complexity, MAP compatibility among equipment units, or *interoperability*, has been a continuing difficulty. MAP has not been applied as rapidly as was initially hoped for by its proponents because of the complexity, costs, and disagreements on how it should be implemented. Assembling a complete set of documentation for MAP is a difficult activity that requires compiling a file of standards organization reports, a number of industry organization reports, and documentation from all the working committees associated with ISO.

Broadband System for MAP Protocol

The MAP protocol was developed with several alternatives for physical layer (PhL)1. MAP can be implemented through what is called a *broadband* system (Fig. 3.13). So that the manufacturing units can talk to one another, transmitted messages are placed on the cable; a *head-end remodulator* retransmits these messages and directs them to the receiving station. The broadband version of MAP has the highest capabilities because it allows several types of communications to take place on the same cabling at the same time. On the other hand, because of its greater flexibility, a broadband system is more complex and more expensive to install. It requires modems and MAP interface equipment for each item of equipment and a head-end remodulator to serve the entire network. The main cable used for broadband is unwieldy (approximately 25 mm in diameter) and is appropriate only for wiring very large factories. Multiple drop cables can be branched off the main cable for the different MAP nodes. Broadband communication can achieve a high data rate of 10 Mb/s and can split the frequency spectrum to allow several different communications to take place simultaneously. As indicated in Fig. 3.14, the three transmit frequencies and the three receive frequencies are separated from one another in the frequency domain. Three different channels can coexist on the MAP network. The head-end remodulator transfers messages from the low frequencies to the high frequencies.

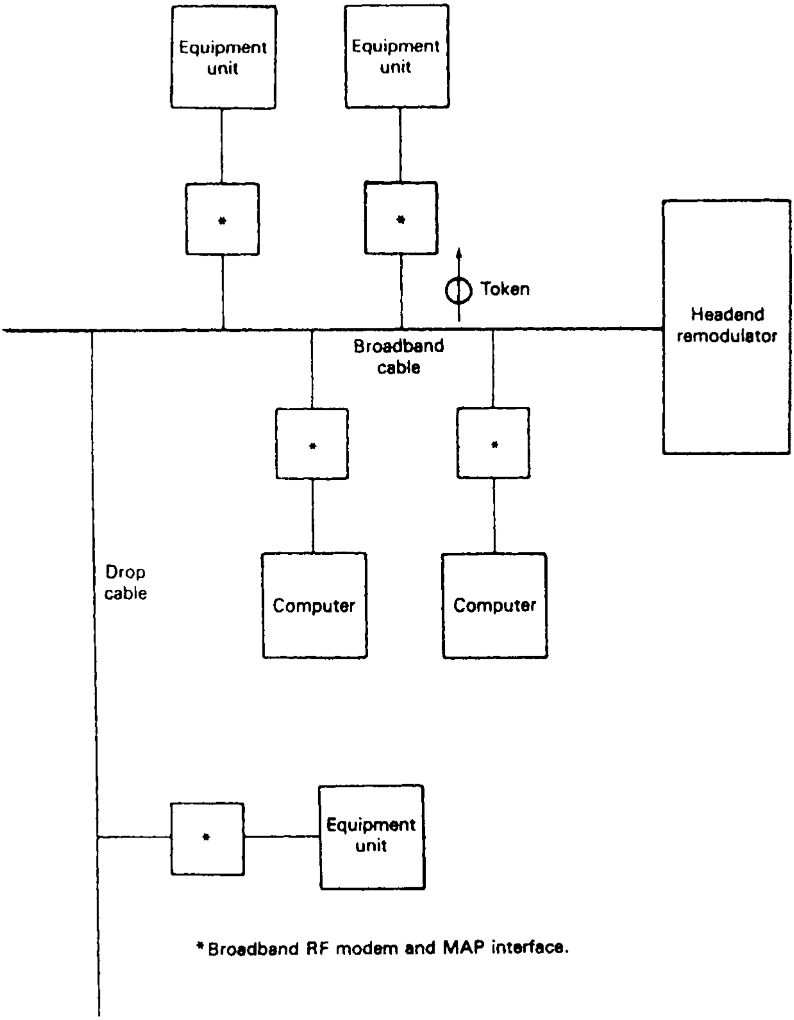


FIGURE 3.13 Broadband system.

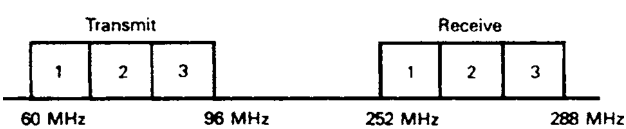


FIGURE 3.14 Broadband frequencies.

Carrier-Band System for MAP Protocol

Another type of MAP network makes use of a carrier-band approach, which uses somewhat less expensive modems and interface units and does not require heavy duty cable (Fig. 3.15). For a small factory, a carrier-band version of MAP can be much more cost-effective. The carrier-band communication can also achieve a high data rate of 5 to 10 Mb/s, but only one channel can operate on the cable at a given time. A single channel is used for both transmission and reception.

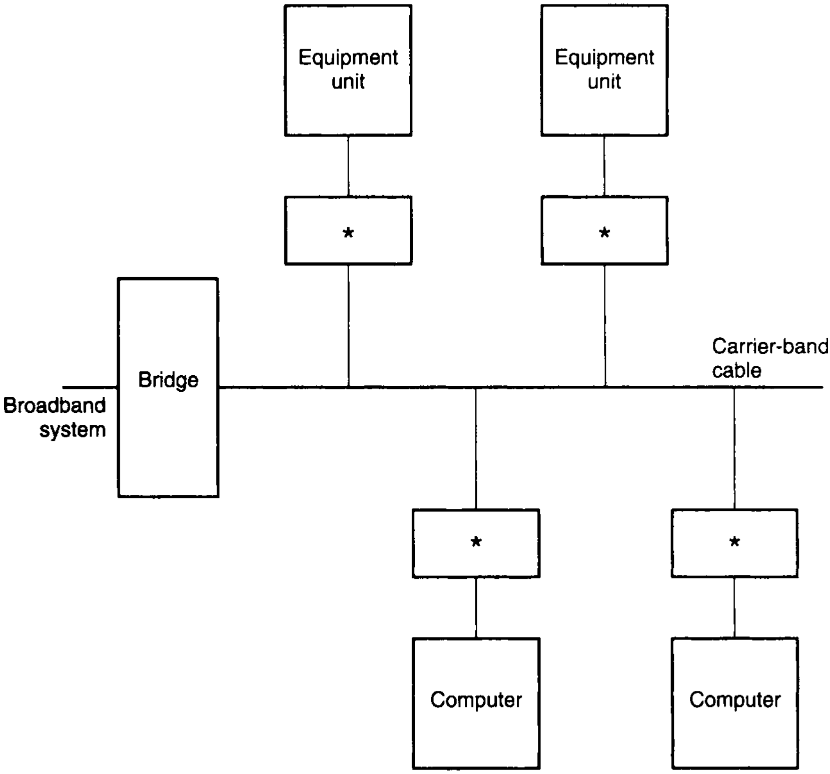


FIGURE 3.15 Carrier-band system for MAP protocol.

Bridges MAP Protocol

It is possible to use devices called *bridges* to link a broadband factory-wide communication network to local carrier-band networks (Fig. 3.16). The bridge transforms the message format provided on one side to the message formats required on the other. In this sense, a bridge transforms one operating protocol to another.

Token System for MAP Protocol

In developing MAP, General Motors was concerned with assuring that every MAP mode would be able to claim control of the network and communicate with other nodes within

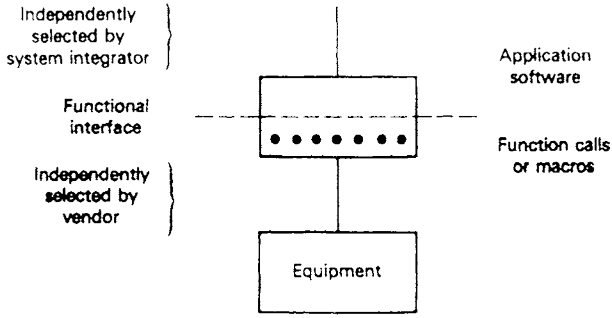


FIGURE 3.16 Bridges system for MAP protocol.

a certain maximum waiting time. Within this waiting time, every node would be able to initiate its required communications. To do this, MAP implements a *token passing* system (Fig. 3.17). The token in this case is merely a digital word that is recognized by the computer. The token is rotated from node address to node address; a node can claim control of the network and transmit a message only when it holds the token. The token prevents message collisions and also ensures that, for a given system configuration, the maximum waiting time is completely defined (so long as no failures occur). The token is passed around a logic ring defined by the sequence of node addresses, not necessarily by the physical relationships.

The token is a control word, and each MAP node can initiate communication only if it passes the token. It is interesting to note that MAP nodes that are not a part of the logic ring will never possess the token, but they may still respond to a token holder if a message is addressed to them. Token management is handled at data link layer (DL)2 of the OSI model. This layer controls the features of the token application with respect to how long a token can be held, the sequence of addresses that is to take place, and the amount of time that is allowed for retrying communications before failure is assumed. If the logic ring is broken at some point—for example, if one equipment unit is no longer able to operate—the other nodes will wait a certain length of time and then will reform the token passing scheme. They will do this by an algorithm through which the token is awarded

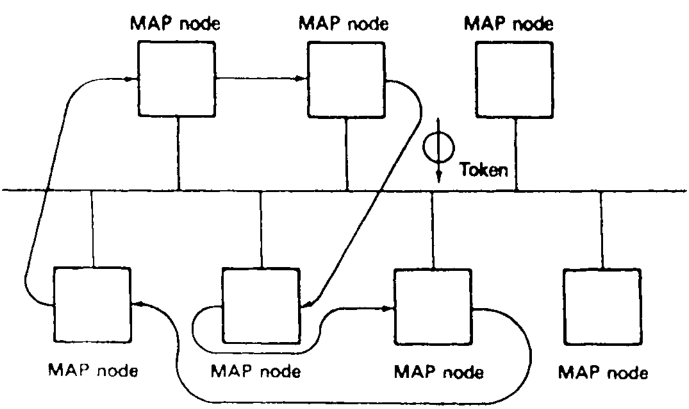


FIGURE 3.17 Token system for MAP protocol.

to the highest station address that contends. The rest of the stations on the ring are then determined by the highest-address successors. This process is repeated until the token ring is re-formed.

Physical layer (PhL)1 involves encoding and modulation of the message so the digital data are transferred into analog and digital communication signals. Each MAP application requires a modem for this purpose. The modem takes the extended message that has been developed at higher layers and modifies it so it can be used to provide an electronic signal to the communications medium. The medium itself provides the means for transferring the signal from User 1 to User 2.

MAP continues to be an important protocol approach for sensors and control systems in computer-integrated manufacturing. For very large factories, the broadband option is available, and for smaller factories, a carrier-band system is also available. A number of vendors now produce the hardware and software necessary to establish a MAP network. However, such networks typically are quite high in cost and, because of the complexity of the protocol, can also be difficult to develop and maintain. Thus, MAP is only one of several solutions available to planning and implementation teams.

MULTIPLE-RING DIGITAL COMMUNICATION NETWORK—AbNET

An optical-fiber digital communication network has been proposed to support the data acquisition and control functions of electric power distribution networks. The optical-fiber links would follow the power distribution routes. Since the fiber can cross open power switches, the communication network would include multiple interconnected loops with occasional spurs (Fig. 3.18). At each intersection a node is needed. The nodes of the communication network would also include power distribution substations and power controlling units. In addition to serving data acquisition and control functions, each node would act as a repeater, passing on messages to the next node.

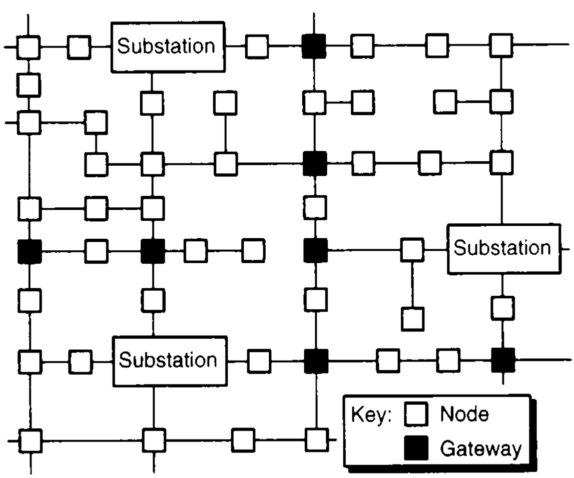


FIGURE 3.18 Multiple-ring digital communication network—AbNET.

Network topology is arbitrary, governed by the power system. The token ring protocols used in single-ring digital communication networks are not adequate. The multiple-ring communication network would operate on the new AbNET protocol, which has already been developed, and would feature fiber optics for this more complicated network.

Initially, a message inserted anywhere in the network would pass from node to node throughout the network, eventually reaching all connected nodes. On the first reception of a message, each node would record an identifying number and transmit the message to the next node. On second reception of the message, each node would recognize the identifying number and refrain from retransmitting the message. This would prevent the endless repetition and recirculating of messages. This aspect of the protocol resembles the behavior of cells in the immune system, which learn to recognize invading organisms on first exposure and kill them with antibodies when they encounter the organisms again. For this reason, the protocol is called *AbNET* after the microbiologists' abbreviation *Ab* for *antibodies*. The AbNET protocols include features designed to maximize the efficiency and fault-tolerant nature of the approach. Multiple service territories can be accommodated, interconnected by *gateway* nodes (Fig. 3.18).

The AbNET protocol is expected to enable a network to operate as economically as a single ring that includes an *active monitor* node to prevent the recirculation of messages. With AbNET, the performance of the proposed network would probably exceed that of a network that relies on a central unit to route messages. Communications would automatically be maintained in the remaining intact parts of the network even if fibers were broken.

For the power system application, the advantages of optical-fiber communications include electrical isolation and immunity to electrical noise. The AbNET protocols augment these advantages by allowing an economical system to be built with topology-independent and fault-tolerant features.

UNIVERSAL MEMORY NETWORK

The universal memory network (UMN) is a modular, digital data communication system that enables computers with differing bus architectures to share 32-bit-wide data between locations up to 3 km apart with less than 1 ms of latency (Fig. 3.19). This network makes it possible to design sophisticated real-time and near-real-time data processing systems without the data transfer bottlenecks that now exist when computers use the usual communication protocols. This enterprise network permits the transmission of the volume of data equivalent to an average encyclopedia each second (40 Mbyte/s). Examples of facilities that can benefit from the universal memory network include telemetry stations, real-time-monitoring through laser sensors, simulation facilities, power plants, and large laboratories (e.g., particle accelerators), or any facility that shares very large volumes of data. The main hub of the universal memory network uses a *reflection center*—a subsystem containing a central control processor (the reflection controller) and a data bus (the reflection bus) equipped with 16 dual memory parts. Various configurations of host computers, workstations, file servers, and small networks or subnetworks of computers can be interconnected by providing memory speed-bandwidth connectivity. The reflection center provides full duplex communication between the ports, thereby effectively combining all the memories in the network into dual-ported random-access memory. This dual-port characteristic eliminates the CPU overhead on each computer that is incurred with Ethernet.

The reflection bus carries write transfers only and operates at a sustained data rate of 40 Mbyte/s. This does not include address, error correction, and coordination information, which makes actual universal memory network bus traffic approach 100 Mbyte/s. The

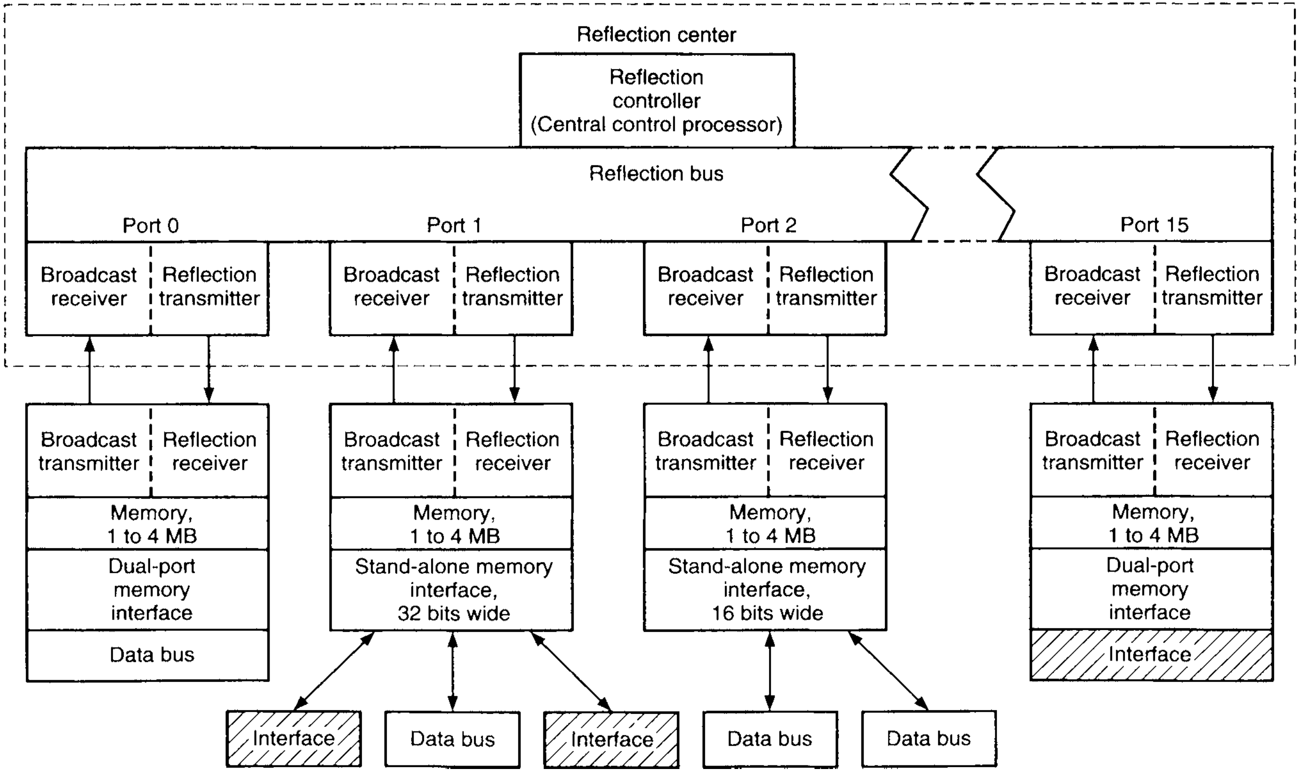


FIGURE 3.19 Universal memory network (UMN)

universal memory network can be implemented in copper cables for distances up to 15 m and in fiber optics for distances up to 3 km. A combination of both for media can be used in the same network. Multiple reflection centers can be interconnected to obtain configurations requiring more ports.

In addition to the reflection center of the main hub, the universal memory network includes smaller subsystems called *shared memory interfaces* (SMIs), which make it possible for computers based on different bus architectures (e.g., SELBus, DEC BI, Multi Bus, VME, and other selected buses) to communicate via the reflection bus. Each host computer is attached to the reflection center by a bus-interface circuit card, which translates the read and write transfers of the host computer to and from the reflection-bus standard. This translation centers around the ordering of bits and conversation used by various vendor architectures to a common strategy required by the 100-ns cycle time of the reflection bus.

The standard memory interface enhances the modular nature of the network. It provides computer memory access to processors of lower cost and enables a large number of workstations to be supported from one reflection center. For example, one reflection center can support up to 12 SMI memory interfaces, each with the capacity to support between 8 and 16 workstations, depending on local hardware configurations. Multiple reflection centers can be interconnected to support even more workstations.

REFERENCES

1. Ames, J. G., "Which Network Is the Right One," *Manufacturing Engineering*, 56 (May) (1988).
2. Borggraaf, P., "Semiconductor Factory Automation: Current Theories," *Semiconductor International*, 88 (October) (1985).
3. Bux, W., "Local Area Subnetworks: A Performance Comparison," *IEEE Trans. Communications*, COM-29(10):1465 (1981).
4. Campbell, J., *The RS-232 Solution*, Sybex, Inc. Alameda, Ca., 1984.
5. Kaminski, A. M., Jr., "Protocols for Communicating in the Factory," *IEEE Spec.*, 65 (April) (1986).
6. Kleinrock, L., and S. S. Lam, "Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation," *IEEE Trans. Communications*, COM-23(4):410 (1975).
7. Metcalfe, R. M., and D. R. Boggs, "Ethernet: Distributed Packet-Switching for Local Computer Networks," *Communications of the ACM*, 19:395 (1976).
8. Shock, J. F., and J. A. Hupp, "Measured Performance of an Ethernet Local Network," *Communications of the ACM*, 23:711 (1980).
9. Tanenbaum, A. S., *Computer Networks*, 2d ed., Prentice-Hall, Englewood Cliffs, N.J., 1988.
10. Talvage, J., and R. G. Hannam, *Flexible Manufacturing Systems in Practice: Application Design and Simulation*, Marcel Decker New York: 1988.
11. Voelcker, J., "Helping Computers Communicate," *IEEE Spectrum*, 61 (March) (1986).
12. Warndorf, P. R., and M. E. Merchant, "Development and Future Trends in Computer Integrated Manufacturing in the USA," *International Journal in Technology Management*, 1(1–2):162 (1986).
13. Wick, C., "Advances in Machining Centers," *Manufacturing Engineering*, 24 (October) (1987).
14. Wittry, E. J., *Managing Information Systems: An International Approach*, Society of Manufacturing Engineers, Dearborn, Mich., 1987.

This page intentionally left blank

CHAPTER 4
THE ROLE OF SENSORS AND CONTROL TECHNOLOGY IN COMPUTER-INTEGRATED MANUFACTURING

INTRODUCTION

According to various studies conducted in the United States, nearly 50 percent of the productivity increase during the period 1950–1990 was due to technological innovation. That is, the increase was due to the introduction of high-value-added products and more efficient manufacturing processes, which in turn have caused the United States to enjoy one of the highest living standards in the world. However, unless the United States continues to lead in technological innovation, the relative living standard of the country will decline over the long term.

This clearly means that the United States has to invest more in research and development, promote scientific education, and create incentives for technological innovation. In the R&D arena, the United States has been lagging behind other nations: about 1.9 percent of the U.S. gross national product (GNP) (versus about 2.6 percent of the GNP in Japan and West Germany) goes for R&D. The introduction of computer-integrated manufacturing (CIM) strategy in U.S. industry has begun to provide a successful flow of communication, which may well lead to a turnaround. *Sensors and control systems in manufacturing* are a powerful tool for implementing CIM. Current world business leaders view CIM as justifying automation to save our standard of living. Successful implementation of CIM depends largely on creative information gathering through sensors and control systems, with information flow as feedback response. Information gathering through sensors and control systems is *imbedded* in CIM. CIM provides manufacturers with the ability to react more quickly to market demands and to achieve levels of productivity previously unattainable.

Effective implementation of sensors and control subsystems within the CIM manufacturing environment will enable the entire manufacturing enterprise to work together to achieve new business goals.

CIM PLAN

This chapter will address implementation of a CIM plan through the technique of modeling.

A model can be defined as a tentative description of a system or theory; the model accounts for many of the system's known properties. An *enterprise model* can be defined (in terms of its functions) as the function of each area, the performance of each area, and the performance of these areas interactively. The creation of a model requires an accurate description of the needs of an enterprise.

In any manufacturing enterprise, there is a unique set of business processes that are performed in order to design, produce, and market the enterprise's products. Regardless of how unique an enterprise or its set of processes is, it shares with others the same set of high-level objectives. To attain the objectives, the following criteria must be met:

- Management of manufacturing finances and accounting
- Development of enterprise directives and financial plans
- Development and design of product and manufacturing processes utilizing adequate and economical sensors and control systems
- Management of manufacturing operations
- Management of external demands

CIM Plan in Manufacturing

In manufacturing, CIM promotes customer satisfaction by allowing order entry from customers, faster response to customer enquiries and changes, via electronic sensors, and more accurate sales projections.

CIM Plan in Engineering and Research

In engineering and research, CIM benefits include quicker design, development, prototyping, and testing; faster access to current and historical product information; and paperless release of products, processes, and engineering changes to manufacturing.

CIM Plan in Production Planning

In production planning, CIM offers more accurate, realistic production scheduling that requires less expediting, canceling, and rescheduling of production and purchase orders.

In plant operations, CIM helps control processes, optimize inventory, improve yields, manage changes to product and processes, and reduce scrap and rework. CIM also helps in utilizing people and equipment more effectively, eliminating production crises, and reducing lead time and product costs.

CIM Plan in Physical Distribution

In physical distribution, where external demands are satisfied with products shipped to the customer, CIM helps in planning requirements; managing the flow of products; improving efficiency of shipment, vehicle, and service scheduling; allocating supplies to distribution centers; and expediting processing of returned goods.

CIM Plan for Business Management

For business management activities such as managing manufacturing, finance, and accounting, and developing enterprise directives and financial plans, CIM offers better product cost tracking, more accuracy in financial projections, and improved cash flow.

CIM Plan for the Enterprise

For the enterprise as a whole, these advantages add up to faster release of new products, shorter delivery times, optimized finished goods inventory, shorter production planning and development cycles, reduced production lead time, improved product quality, reliability and serviceability, increased responsiveness, and greater competitiveness. In effect, CIM replaces an enterprise's short-term technical improvements with a long-term strategic solution.

The advantages of CIM with sensors and control systems are not just limited to the four walls of an enterprise. It can also deliver real productivity gains in the outside world. For example, suppliers will be able to plan production, schedule deliveries, and track shipments more efficiently. Customers will benefit from shorter order-to-delivery times, on-time deliveries, and less expensive, higher-quality products.

MANUFACTURING ENTERPRISE MODEL

The integration and productivity gains made possible by CIM with sensors and control systems are the key to maintaining a competitive edge in today's manufacturing environments. The enterprise model defines an enterprise in terms of its functions. In a traditional enterprise that relies on a complex organization structure, operations and functional management are divided into separate departments, each with its own objectives, responsibilities, resources, and productivity tools.

Yet, for the enterprise to operate profitably, these departments must perform in concert. Sensors and control systems that improve one operation at the expense of another, and tie up the enterprise's resources, are counterproductive. New sensors and control systems in CIM can create a systematic network out of these insulated pockets of productivity. But to understand how, one must examine the elements of an enterprise model and see how its various functional areas work—independently and with each other.

Creating a model of the enterprise can help to expose operations that are redundant, unnecessary, or even missing. It can also help determine which information is critical to a successful implementation once effective sensors and control systems are incorporated.

Obviously, this model is a general description. Many industry-unique variations to it exist. Some enterprises may not require all of the functions described, while others may require more than are listed. Still other enterprises may use the same types of functions, but group them differently.

For example, in the aerospace industry, life-cycle maintenance of products is an essential requirement and may require extensions to the model. In the process industry, real-time monitoring and control of the process must be included in the model.

Computer integrated manufacturing harnesses sensors and control system technology to integrate these manufacturing and business objectives. When implemented properly, CIM can deliver increased productivity, cost-efficiency, and responsiveness throughout the enterprise. CIM can accomplish these objectives by addressing each of the major functional areas of the manufacturing enterprise:

- Marketing
- Engineering research and development of sensors in CIM strategy
- Production planning
- Plant operations incorporating sensors and shop floor control systems
- Physical distribution
- Business management

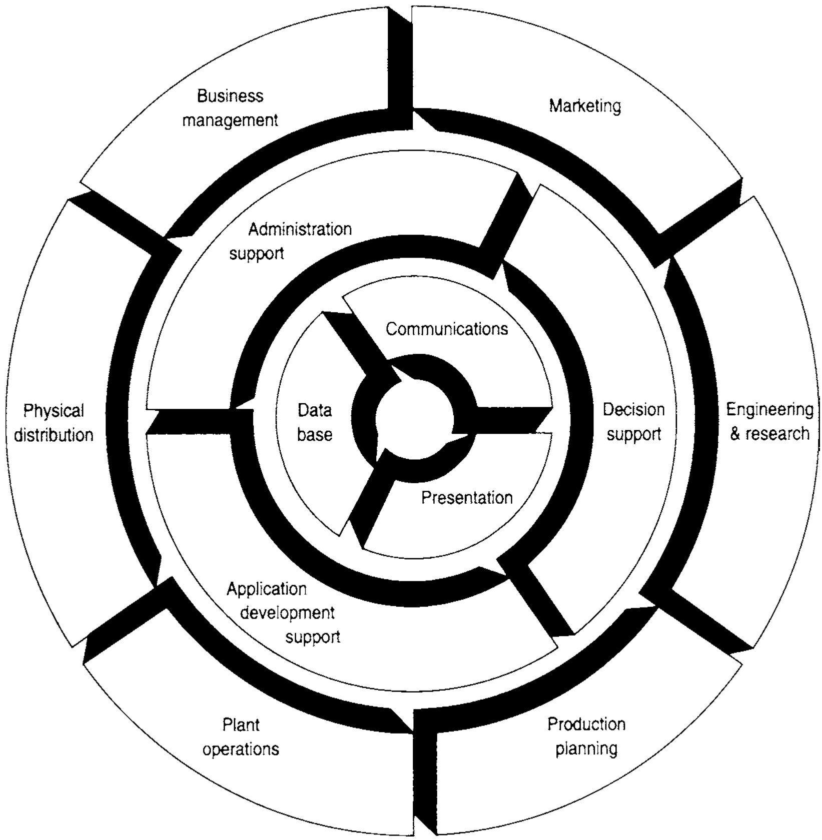


FIGURE 4.1 Major functional areas of manufacturing.

Integrating these functions and their resources requires the ability to share and exchange information about the many events that occur during the various phases of production; manufacturing systems must be able to communicate with the other information systems throughout the enterprise (Fig. 4.1). There must also be the means to capture data close to the source, then distribute the data at the division or corporate level, as well as to external suppliers, subcontractors, and even customers.

To meet these needs, the CIM environment requires a dynamic network of distributed functions. These functions may reside on independent system platforms and require data from various sources. Some may be general-purpose platforms, while others are tailored to specific environments. But the result is an environment that encompasses the total information requirements of the enterprise, from developing its business plans to shipping its products.

With this enterprise-wide purpose, CIM can deliver its benefits to all types of manufacturing operations, from plants that operate one shift per day to processes that must flow continuously, from unit fabrication and assembly to lots with by-products and coproducts. These benefits can also be realized in those enterprises where flexible manufacturing systems are being used to produce diversified products over shorter runs, as well as in those that use CIM with sensors and control systems to maintain an error-free environment.

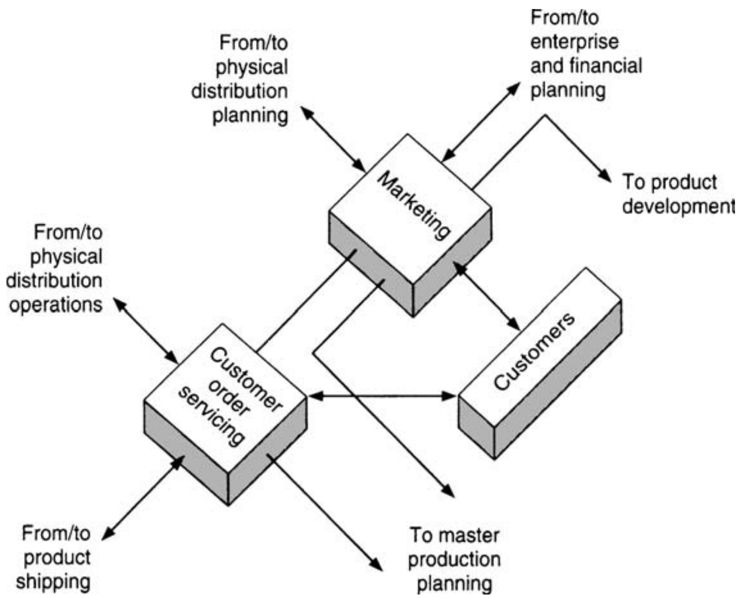


FIGURE 4.2 Marketing relationships.

By creating more efficient, more comprehensive management information systems through sensors and control systems, CIM supports management efforts to meet the challenges of competing effectively in today's world markets.

Marketing

Marketing acts as an enterprise's primary contact with its customer (Fig. 4.2). To help meet the key objectives of increasing product sales, a number of functions are performed within marketing. These include market research; forecasting demand and sales; analyzing sales; tracking performance of products, marketing segments, sales personnel, and advertising campaigns; developing and managing marketing channels; controlling profits and revenues; and managing sales personnel, sales plans, and promotion. Input comes from business management and customers. Output goes to customers, product development, customer order servicing, and master production planning.

The information handling requirements of marketing include monumental texts and graphics as well as queries and analysis of internal and external data. The internal data is gathered through a variety of software routines.

Customer order servicing involves entering and tracking customer orders. This can be for standard products or custom-designed products. Other customer order servicing activities include providing product quotations, checking customer credit, pricing product, allocating order quantities, and selecting shipments from distribution centers.

Input to this area includes order and forecast data from marketing or directly from customers as well as available-to-promise data from production planning. Output can include allocation of all orders, quotations for custom products, communication with production engineering regarding custom products, order consideration, and shipping releases. Customer service will significantly improve through electronic data interchange (EDI).

Engineering and Research

The engineering and research areas of an enterprise can be broken down into separate activities (Fig. 4.3). Each of these has its own special needs, tools, and relationships to other areas.

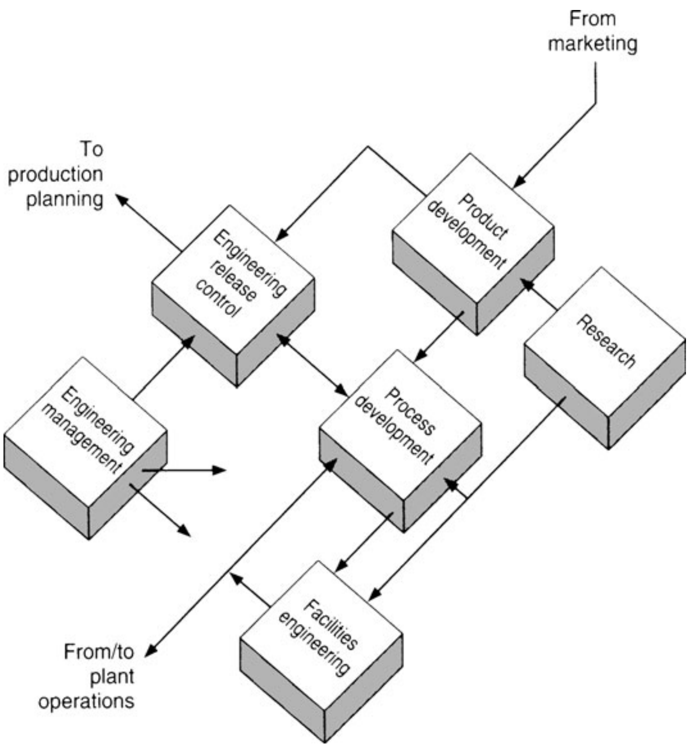


FIGURE 4.3 Engineering and research.

The research activities include investigating and developing new materials, products, and process technology. Information processing needs include complex analyses, extensive texts, imaging, graphics, and videos.

Research input comes from such outside research sources as universities, trade journals, and laboratory reports. Then research must communicate its findings to three other functional areas—product development, process development, and facilities engineering.

Product Development. In today's increasingly competitive manufacturing markets, creation of new material and production technologies is essential for the successful development of products. The product development area uses these materials and production technologies to design, model, simulate, and analyze new products.

Product development activities include preparing product specifications and processing requirements, drawings, materials or parts lists, and bills of material for new products or engineering changes.

In this area, laboratory analysis tools, computer-aided design/computer-aided engineering (CAD/CAE) tools, and group technology (GT) applications are helping reduce product development time, increase productivity, and improve product quality.

Product development comes from marketing, research, and plant operations. Its output—product specifications, manufacturing control requirements, drawings, text, and messages—is directed to process development and engineering release control.

Process Development. This functional area creates process control specifications, manufacturing routings, quality test and statistical quality control specifications, and numerical control (NC) programming. It also validates the manufacturability of product designs. Computer-aided process planning (CAPP) programs and group technology applications for routing similar parts have helped streamline these functions. Expert systems have also been used to supplement traditional product testing and defect-analysis processes. This area is also responsible for the application of new manufacturing technologies such as work cells, conveyer systems, and robotics. Sensors and control systems play a fundamental role within this work-cell structure.

Process development receives input from research and product development as well as statistical process data from plant operations. Output includes providing routings, process control algorithms, and work-cell programming to plant operations by way of engineering release control.

Facilities Engineering. The chief responsibility of facilities engineering is the installation of plant automation incorporating new equipment with sensors and control systems, material flow, inventory staging space, and tools. Tools may include driverless material handling equipment, conveyers, and automated storage and retrieval systems. This area is also responsible for plant layout and implementation of such plant services and utilities as electricity, piping, heat, refrigeration, and light.

Input to facilities engineering is from research and process development. Output, such as plant layouts, changes of schedule, and forecasts of new equipment availability, goes to plant operations.

Engineering Release Control. This function involves the coordination of the release of new products, processes, tools, and engineering changes to manufacturing. A major check point in the product cycle is obtaining assurance that all necessary documentation is available, after which the information is released to manufacturing.

Input is received from product and process development activities. Specific output, including product and tool drawings, process and quality specifications, NC programs, and bills of material, is transferred to production planning and plant operations.

Engineering Management. Among the activities of engineering management are introducing engineering releases, controlling product definition data, estimating cost, entering and controlling process data, and defining production resources.

Input and output for engineering management include exchanging designs and descriptions with engineering design; reviewing product data, routings, and schedules with production planning; and accepting engineering change requests from plant operations.

Production Planning

Production planning can be viewed as five related functional areas (Fig. 4.4).

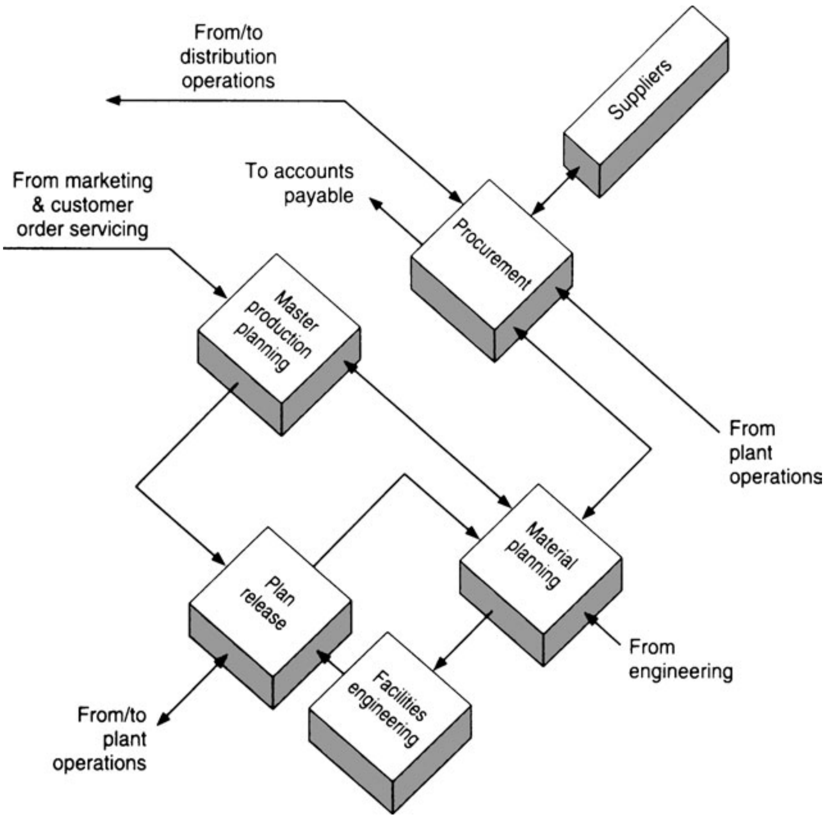


FIGURE 4.4 Production planning.

Master Production Planning. In master production planning, information is consolidated from customer order forecasts, distribution centers, and multiple plans in order to anticipate and satisfy demands for the enterprise's products. Output includes time-phased requirements sent to the material planning function, as well as the final assembly schedule.

Material Planning and Resource Planning. These two areas require timely and accurate information—demand schedules, production commitment, inventory and work-in-progress status, scrap, actual versus planning receipts, shortages, and equipment breakdowns—in order to keep planning up to date with product demands.

Product and process definition data come from the engineering areas. Output is to plant operation and procurement and includes production schedules, order releases, and plans for manufactured and purchased items.

Procurement. Procurement involves selecting suppliers and handling purchase requirements and purchase orders for parts and materials. Among the input is material requirements from material planning and just-in-time delivery requests from plant operations. Other input includes shipping notices, invoices, and freight bills.

Output to suppliers includes contracts, schedules, drawings, purchase orders, acknowledgments, requests for quotations, release of vendor payments, and part and process specifications. In order to streamline this output, as well as support just-in-time concepts, many enterprises rely on sensors for electronic data interchange with vendors.

Plant Release. The functions of this area can vary, depending on the type of manufacturing environment. In continuous-flow environments, for example, this area produces schedules, recipes to optimize use of capacity, specifications, and process routings.

For job-shop fabrication and assembly environments, this area prefers electronic or paperless-shop documents consisting of engineering change levels; part, assembly, setup, and test drawings and specifications; manufacturing routings; order and project control numbers; and bar codes, tags, or order-identification documents.

However, large-volume industries are evolving from typical job-shop operations to continuous-flow operations, even for fabrication and assembly-type processes.

Input—typically an exploded production plan detailing required manufacturing items—comes from material planning. Output—schedules, recipes, or shop packets—is used in plant operations for scheduling.

Plant Operations

Plant operations can be described in terms of nine functions (Fig. 4.5).

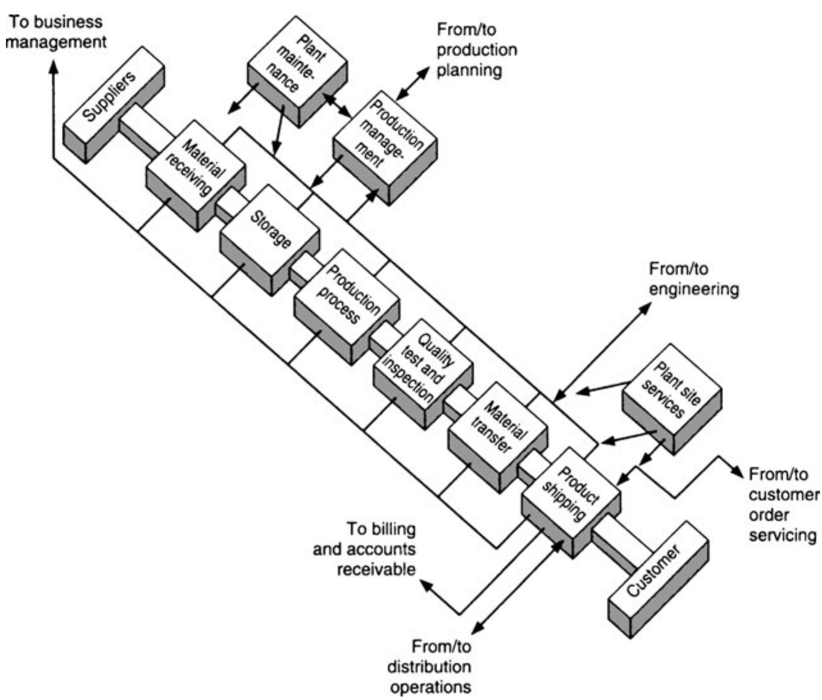


FIGURE 4.5 Plant operations.

Production Management. This area provides dynamic scheduling functions for the plant floor by assigning priorities, personnel, and machines. Other activities include sending material and tool requests for just-in-time delivery.

Input for production management includes new orders and schedules from production planning and real-time feedback from plant operation. Output includes dynamic schedules and priorities that are used to manage operations on the plant floor.

Material Receiving. The material receiving function includes accepting and tracking goods, materials, supplies, and equipment from outside suppliers or other locations within the enterprise. Input to this area includes receiving reports and purchase order notices.

Output includes identifying receipts with appropriate documentation, then routing materials to the proper destination. Data are also sent to accounting, procurement, and production management. Production management can also direct materials based on more immediate needs.

Storage. This represents inventory management—where materials are stored and are accessible to the proper production locations. The materials include finished goods, raw materials, parts, supplies, work-in-progress, and tools, as well as nonproduction materials and equipment.

Storage functions include preparing item identification and storage tags, managing storage locations, processing pick requests, reporting picks and kit activities, and planning physical inventory cycles and reporting counts.

Storage input includes storage and picking requests from production management scheduling functions. Output includes receiving and disbursement reports for use in production management and accounting.

Production Process. Production process functions include managing the production process, processing materials, fabricating parts, grading or reworking components, assembling final products, and packaging for distribution.

One of today's trends in fabrication and assembly processes is toward continuous processing, such as continuous movement with minimal intermediate inventories. This can be described with such terms as continuous-flow manufacturing, flexible manufacturing cells, just-in-time logistics, and group technology. Unfortunately, in many instances, these automation efforts are autonomous, without regard to the other functions of the enterprise.

The information handling needs of the production process can include analog and digital data, text, graphics, geometries, applications programs—even images, video, and voice. Processing this information may require subsecond access and response time.

Input to this area includes shop documents, operator instructions, recipes, and schedules from production management as well as NC programs from process development. Output consists of material and tool requests, machine maintenance requests, material transfer requests, production, and interruption reports for production management, production and labor reports for cost accounting and payroll, and statistical process reports for production management and process development.

Quality Test and Inspection. Testing items and products to assure the conformity of specifications is the main activity in quality test and inspection. This includes analyzing and reporting results quickly by means of metrological sensors and control systems, in order to reduce scrap and rework costs.

Quality test and product specifications are input from engineering. Chief output includes purchased-item inspection results to procurement, manufactured-item inspection and product test results to production process and production management, quality test and inspection activity reports to cost accounting, and rejected part and product dispositions to material handling.

Material Transfer. Material transfer involves the movement of materials, tools, parts, and products among the functional areas of the plant. These activities may be manual. Or they may be semiautomated, using control panels, forklifts, trucks, and conveyers. Or they may be fully automated, relying on programmable logical controllers, distribution control systems, stacker cranes, programmed conveyers, automated guided vehicles, and pipelines.

Input in this area may be manual requests or those generated by the system. Output includes reporting completed moves to production management.

Product Shipping. This area supports the movement of products to customers, distributors, warehouses, and other plants. Among the activities are selecting shipment and routing needs, consolidating products for a customer or carrier order, preparing shipping lists and bills of lading, reporting shipments, and returning goods to vendors.

The primary input is from customer order servicing, and it includes the type of product and the method and date of shipment. Output includes reporting shipment dates to customer, order servicing, billing, and accounts receivable.

Plant Maintenance. Plant maintenance includes those functions that ensure the availability of production equipment and facilities. Maintenance categories include routine, emergency, preventive, and inspection services. In addition, many of today's advanced users of systems are moving toward diagnostic tools based on expert systems, which reduce equipment downtime. Input (maintenance requests) can be initiated by plant personnel, a preventive maintenance and inspection system, or a process and equipment monitoring system. Output includes requests for purchase of maintenance items, schedules for maintenance for use in production management, requests for equipment from facilities engineering, and maintenance work order costs to cost accounting.

Plant Site Services. The final area of plant operation is plant site services. Input received and output provided cover such functions as energy supply and utilities management, security, environmental control, grounds maintenance, and computer and communications installations.

Physical Distribution

Physical distribution can be viewed as having two functional areas (Fig. 4.6).

Physical Distribution Planning. This involves planning and control of the external flow of parts and products through warehouses, distribution centers, other manufacturing locations, and points of sale. These functions may also include allocating demand, planning finished goods inventory, and scheduling vehicles. Some industries require another major set of functions that relate to logistics reports. This includes spare parts, maintenance, training, and technical documentation.

Input and output are exchanged with marketing and physical distribution operations. The information handling requirements of physical distribution planning are usually medium to heavy, especially if multiple distribution centers exist.

Physical Distribution Operations. This area includes receiving, storing, and shipping finished goods at the distribution center or warehouse. Receipts arrive from plants or other suppliers, and shipments are made to customers and dealers. Other functions can include scheduling and dispatching vehicles, processing returned goods, and servicing warranties and repairs.

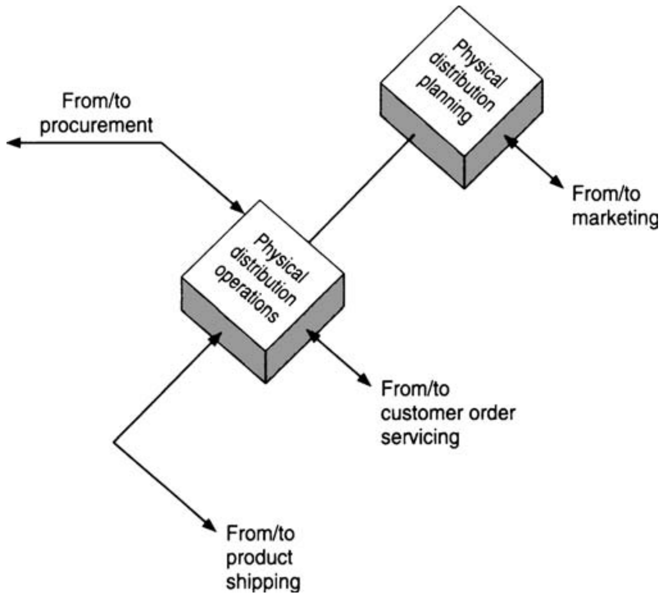


FIGURE 4.6 Physical distribution.

Input is received from plant site product shipping, procurement, physical distribution planning, and customer order servicing. Output includes acknowledgments to plant site product shipping and procurement, as well as data for updates to physical distribution planning and customer order servicing.

Business Management

Within the enterprise model, a business management function may be composed of seven areas (Fig. 4.7).

Financial Planning and Management. In financial planning and management, financial resource plans are developed and enterprise goals are established. Among the functions are planning costs, budgets, revenues, expenses, cash flow, and investments.

Input includes financial goals and objectives established by management as well as summarized financial data received from all areas of the enterprise. Output includes financial reports to stockholders, departmental budgets, and general ledger accounting.

Accounts Payable. Accounts payable primarily involves paying vendors. Input includes vendor invoices and goods-received reports. The output includes discount calculations to vendors and payment checks. This last function lends itself to the use of electronic data interchange to electronically transfer funds to vendors.

Billing and Accounts Receivable. This area prepares invoices to customers and manages customer account collections. Input to this area consists of product shipping data and cash received. Output includes customer invoices, delinquent account reports, and credit ratings. Transferring funds electronically through EDI can simplify this area significantly.

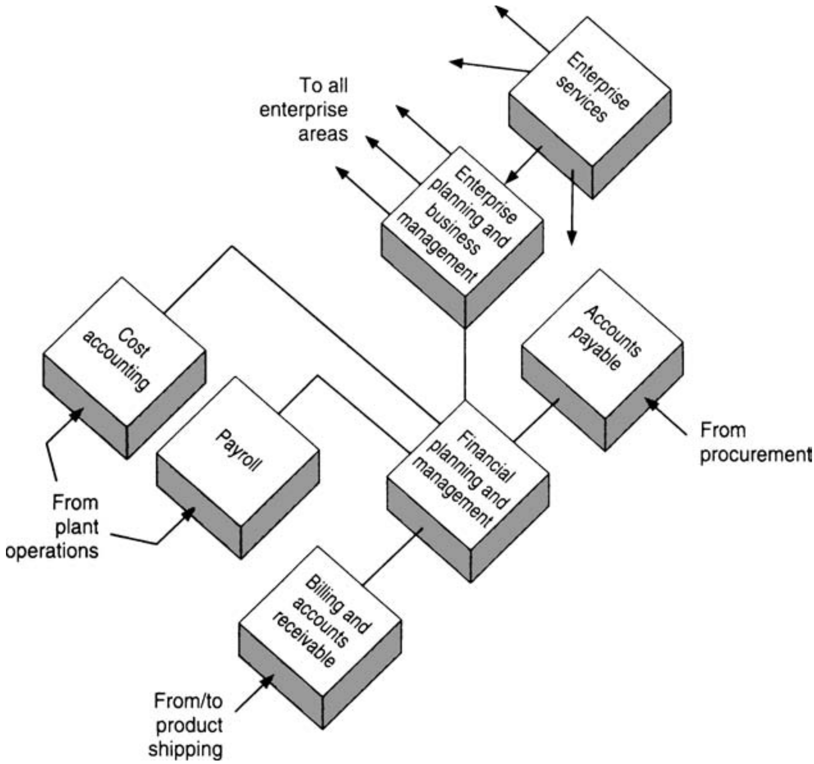


FIGURE 4.7 Business management.

Cost Accounting. Cost accounting activity supports product pricing and financial planning by establishing product costs. These costs can include those of materials, direct and indirect labor, fixed production elements (machinery and equipment), variable production elements (electricity, fuels, or chemicals), and overhead. Other functions can include establishing standard costs, reporting variances to standards, costing job orders, and determining accrued costs.

Cost accounting input is acquired primarily from plan operations. Output is sent to financial management and planning.

Payroll. This area computes payments, taxes, and other deductions for employees. It also includes reporting and paying employee tax withholdings to government agencies. Input includes time and attendance information and production data from plant operations. Output includes payroll checks, labor distribution, and government reports and payments.

Enterprise Planning and Business Management. These functions include establishing goals and strategies for marketing, finance, engineering and research, plant automation, sensors and control systems, and information systems. Input and output are exchanged through sensors and control systems with virtually every other area of the enterprise.

Enterprise Services. Enterprise services include such support services as office personnel, management information services, personnel resources, and public relations. These services require extensive administrative support tools, such as text processing, decision support, and graphic tools. But since input and output will be exchanged throughout the entire enterprise, it is imperative that these tools be integrated with the enterprise's other systems.

DESIGN OF CIM WITH SENSORS AND CONTROL SYSTEMS

With the advent of low-priced computers and sensors and control systems, a number of technological developments related to manufacturing can be used to make production more efficient and competitive. The primary purpose is to develop several computer concepts as related to the overall manufacturing plan of CIM systems.

In order for the manufacturing enterprise to succeed in the future, it is imperative that it adopt a manufacturing strategy that integrates its various functions. CIM systems have the potential to accomplish this task. The implementation of CIM with sensory and control systems on the shop floor represents a formidable, albeit obtainable, objective. To accomplish this goal, enterprises must have access to information on what is available in CIM. A secondary purpose of obtaining access to information is to provide a framework that can aid in the search for information. Once the information is obtained, it becomes necessary to look at the current system objectively and decide how to approach the problem of implementing CIM with sensors and control systems.

While many of the ideas associated with CIM are new and untried, progressive enterprises, with the realization that old methods are ineffective, are doing their part in implementing this new technology. Some of the concepts currently being implemented are flexible manufacturing systems, decision support systems (DSS), artificial intelligence (AI), just-in-time (JIT) inventory management, and group technology. While all of these concepts are intended to improve efficiency, each one alone can only accomplish so much. For example, an FMS may reduce work-in-process (WIP) inventory while little is accomplished in the area of decision support systems and artificial intelligence to relate all aspects of manufacturing management and technology to each other for FMS. The advent of inexpensive sensors and control systems enables the concept of CIM to be implemented with greater confidence.

Recent advances in computer technology and sensors in terms of speed, memory, and physical space have enabled small, powerful, personal computers to revolutionize the manufacturing sector and become an essential part of design, engineering, and manufacturing, through, for example, database management systems (DBMSs) and local area networks (LANs). The coordination of the various aspects of a manufacturing environment means that complex systems inherently interact with one another. Due to a lack of standards and poor communication between departments, many components and databases are currently incompatible.

Table 4.1 describes some benefits of CIM.

TABLE 4.1 CIM Benefits

Application	Improvement with CIM, %
Manufacturing productivity	120
Product quality	140
Lead time (design to sale)	60
Lead time (order to shipment)	45
Increase in capital equipment utilization	340
Decrease in WIP inventory	75

The potential for CIM, according to Table 4.1, is overwhelming, but the main issue is how to analyze and design CIM that incorporates sensors, control systems, and decision support so it is utilized effectively.

Manufacturing problems are inherently multiobjective. For example, improving quality usually increases cost and/or reduces productivity. Furthermore, one cannot maximize quality and productivity simultaneously; there is a tradeoff among these objectives. These conflicting objectives are treated differently by different levels and/or units of production and management. Obviously, without a clear understanding of objectives and their interdependence at different levels, one cannot successfully achieve CIM with sensors and control systems.

Components of CIM with Sensors and Control Systems

The decision making in design of CIM with effective sensors and control systems can be classified into three stages.

1. *Strategic level.* The strategic level concerns those decisions typically made by the chief executive officer (CEO) and the board of directors. Upper management decisions of this type are characterized by a relatively long planning horizon, lasting anywhere from one to ten years. Implementing CIM with sensors and control systems has to begin at this level. Even though small enterprises may not have as many resources at their disposal, they have the added advantage of fewer levels of management to work through while constructing CIM.
2. *Tactical level.* At the tactical level, decisions are made that specify how and when to perform particular manufacturing activities. The planning horizon for these decisions typically spans a period from 1 to 24 months. Activities at this level include such intermediate functions as purchasing and inventory control. They affect the amount of material on the shop floor but do not control the use of the material within the manufacturing process.
3. *Operational level.* Day-to-day tasks, such as scheduling, are performed at the operational level. The primary responsibility at this level is the effective utilization of the resources made available through the decisions made on the strategic and tactical levels. Because of the variability in demand or machine downtime, the planning horizon at this level must be relatively short, normally 1 to 15 days.

While each of these levels has certain responsibilities in a manufacturing plant, the objectives are often conflicting. This can be attributed to inherent differences between departments—for example, sales and marketing may require a large variety of products to serve every customer's needs, while the production department finds its job easier if there is little product variation. One of the main causes of conflicting decisions is a lack of communication due to ineffective sensors and control systems between levels and departments. CIM with adequate sensors and control systems provides the ability to link together technological advances, eliminate much of the communication gap between levels, and bring all elements into a coherent production system.

CIM with Sensors and Control Systems at the Plant Level

Some of the important emerging concepts related to CIM with effective sensors and control systems are flexible manufacturing systems, material handling systems, automated storage and retrieval systems (AS/RS), computer-aided design (CAD), computer-aided engineering (CAE), computer-aided manufacturing (CAM), and microcomputers. These components of CIM can be classified into three major groups (Fig. 4.8).

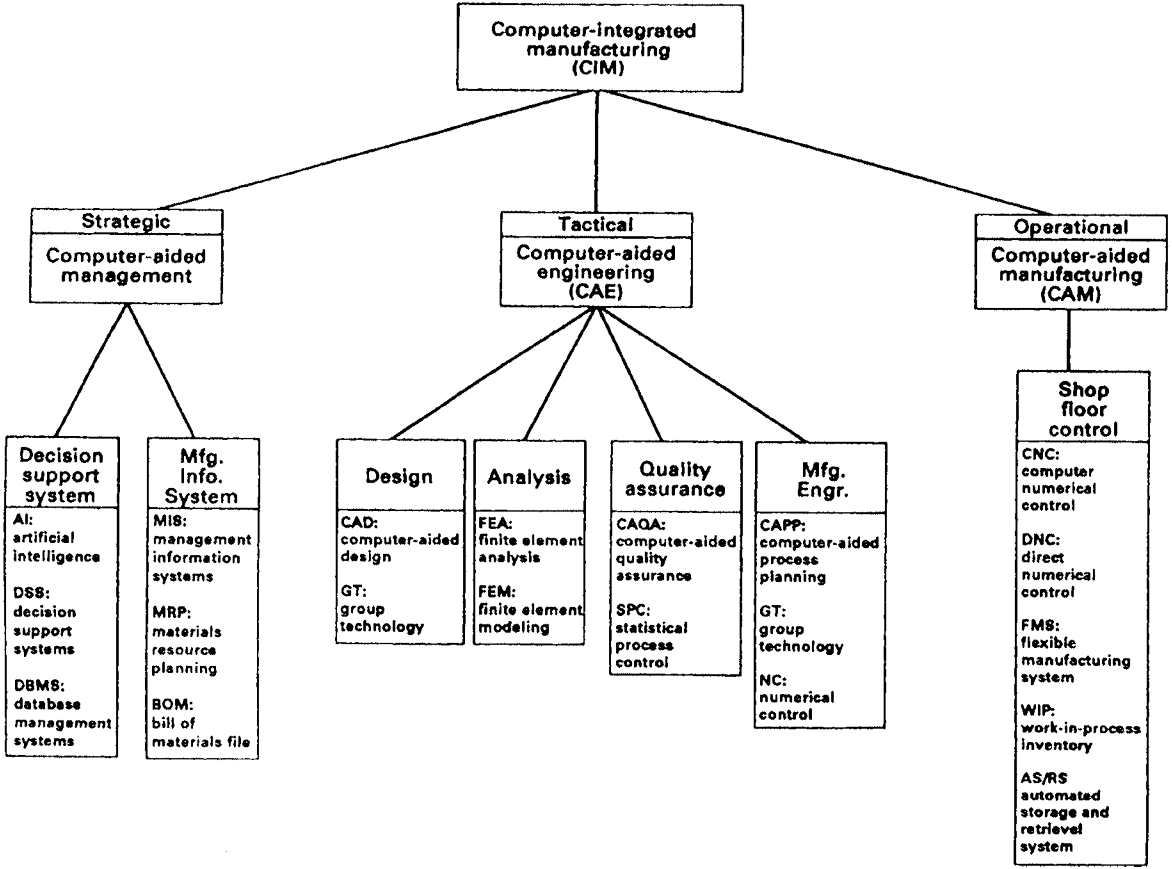


FIGURE 4.8 Components within CIM.

Flexible Manufacturing Systems Incorporating Sensors and Control Systems. An FMS can link several elements on the shop floor through sensors in order to coordinate those elements. While CIM can be applied to any manufacturing industry, FMSs find their niche in the realm of discrete production systems such as job shops.

The most important FMS elements are numerical control machines and an automated material handling network to transport the product from raw material inventory, through the NC operations, and finally to the finished goods inventory.

Numerical control technology has made major advances with the advent of computer numerical control and direct numerical control (CNC/DNC). Microprocessors and sensors located on the machine itself can now provide the codes necessary for the parts to be operated on.

Material Handling. Material handling is the means of loading, unloading, and transporting workpieces among different machines and departments. Material handling can be accomplished in several ways:

- *Transfer lines* consist of fixed automation machinery such as conveyer belts. Their advantages are high speed and low cost. Their major disadvantage is their lack of flexibility. Dedicated transfer lines can handle only a limited number of parts and cannot be easily changed once in place, thus defeating the goals of an FMS.
- *Robots* provide another alternative for moving workpieces. Generally, robots can be made very flexible because of their programmability, but they are limited to their region of access.
- *Automated guided vehicles* can move workpieces a great distance, but they lack the speed found in both robot and transfer lines. Yet, because of their ability to be programmed to different routes, they are more flexible than transfer lines.

Automated Storage and Retrieval Systems. By means of AGVs, raw materials can be taken from the loading dock and placed in a designated place in inventory. By means of an AS/RS, inventory can be tracked throughout the manufacturing process and optimized for strategically locating items in storage. Because the process is computerized, data on what exactly is in inventory assist planners in determining order and production schedules.

Inventories consist of raw materials, work in process, and finished goods. Inventories should be controlled by keeping track of inventory locations and amounts. An AS/RS can accomplish this task.

Computer-Aided Engineering/Design/Manufacturing (CAE/CAD/CAM). Computer-aided design helps the engineer in many ways during the design stage. Simply drawing the part on a computer increases the productivity of designers, but CAD is more than just automated drafting. CAD can facilitate group technology and the construction of a bill of materials (BOM) file.

Computer-aided engineering consists of the many computerized facets of engineering that go into a particular product. When a part has been designed, the CAE subgroup is responsible for generating the NC code that can be used by the NC machines on the floor.

By using GT, similar parts can be classified by similar attributes and placed in part families. By grouping parts this way, much redundancy in design and manufacturing is eliminated.

In *computer-aided manufacturing* (*CAM*), computers are used to plan and conduct production. They can correct and process data, control the manufacturing process, and provide information that can be used in decision making. CAM can involve distributed quality control and product testing and inspection, which are built into the manufacturing process to support the larger functional relationships.

Microcomputers for CIM. The integration of different elements of CIM and sensor systems can be accomplished only with the aid of a computer. Because of the magnitude of the CIM and sensor database, a mainframe (or a minicomputer for a small enterprise) is necessary for storage and retrieval of information. The power of microcomputers has increased so dramatically, however, that they are suitable for most other applications, such as:

- *Programmable logic controllers (PLCs).* On the factory floor, microcomputers are subject to harsh conditions. Dust, high temperature, and high humidity can quickly destroy an ordinary personal computer. Traditionally, the shop floor belongs to the programmable logic controller. The components can operate in extreme conditions—for example, 55°C and 90 percent humidity. In addition, PLCs are geared to real-time control of factory operations. Through advances in microprocessor technology, functions once controlled by relays and mechanical switching mechanisms can now be performed by PLCs. The ladder diagrams used to notate logic circuits can now be programmed directly into the memory of the programmable controller. Because of microelectronic circuitry, PLCs can process control information quickly and shut down automatically in case of emergencies. Whereas PLCs have become commonplace on the floor of discrete product operations, process control computers have become indispensable in the control of process plants where conditions must be monitored constantly. They are also used in the control of office and factory environments. For example, a PLC can turn furnaces and air-conditioning units ON and OFF to provide suitable working conditions while optimizing energy use.
- *Industrial personal computers (PCs).* Until recently, PCs were commonly found in the protected environments of offices. Now, manufacturers have introduced rugged versions of popular personal computers. For example, IBM has developed the 5531 and 7531/32, industrialized versions of PC/XT and PC/AT, respectively, to withstand the environment of the factory floor. They have the advantage of being able to run any PC-DOS–based software, but are unable to perform real-time control. This problem has been remedied with the advent of the industrial computer disk operating system (IC-DOS), a real-time operating system that is compatible with other IBM softwares. This allows a shop floor computer to provide real-time control while using software packages previously found only on office models.
- *Microsupercomputers.* The hardware of microsupercomputers has increased computing power significantly. Offering high performance, transportability, and low price, the new microsupercomputers compare favorably to mainframes for many applications.

DECISION SUPPORT SYSTEM FOR CIM WITH SENSORS AND CONTROL SYSTEMS

With an increase in production volume and efficiency comes a need to have a more effective method of scheduling and controlling resources. Herein lies a connection between CAE and computer-aided management. The long-range plans of a company must include forecasts of what the demand will be for various products in the future. Through these forecasts, the enterprise determines what strategy it will take to ensure survival and growth.

For the enterprise to make intelligent decisions, reliable information must be available. In regard to the three levels of decision making, it is also important that the information be consistent throughout each level. The best way to assure this availability and consistency is to make the same database available to all individuals involved in the production process. Because of lack of good communication between levels and sometimes the reluctance of upper-level managers to commit themselves to CIM, constructing a centralized database represents one of the most difficult problems in the implementation of CIM.

Computer-Integrated Manufacturing Database (CIM DB)

The creation of a CIM DB is at the heart of the effective functioning of CIM. Most manufacturers have separate databases set up for nearly every application. Since data from one segment of an enterprise may not be structured for access by other segments' software and hardware, a serious problem for meeting the CIM goal of having readily available data for all levels occurs. Another problem with multiple databases is in the redundancy of data. Both the strategic and tactical decision makers, for example, may need information concerning a bill of material file. Even with the assumption that the databases contain consistent data (i.e., the same information in each), maintaining them both represents inefficient use of computer time and storage and labor. To install CIM, these databases must be consolidated. Unfortunately, bringing multiple databases into one CIM DB that remains available to everyone and consistent in all levels presents a significant obstacle because of the large investment needed in both time and computer hardware and software.

Structure of Multiobjective Support Decision Systems

The success of CIM also depends largely on the ability to incorporate sensor technology with a database. The database is utilized in making decisions on all levels—decisions that are used to update to the database. Decision support systems can provide a framework for efficient database utilization by allowing storage and retrieval of information and problem solving through easy communications.

Decision making problems in manufacturing can be grouped into two general classes:

- *Structured decisions* are those constrained by physical or practical limitations and can be made almost automatically with the correct input. An example is generating a group technology part code given the geometry of the part.
- *Unstructured decisions* typically are those that contain a large degree of uncertainty. Decisions considered by strategic planners are almost always unstructured. Deciding whether or not to expand a certain product line, for example, may be based on demand forecasting and on the expected growth of competitors.

Due to the predictive nature of these decisions, they inherently contain more uncertainty than structured decisions. Long-range planning consists primarily of unstructured information.

Decision support systems mainly consist of three separate parts:

- *Language systems.* The function of a language system (LS) is to provide a means for the user to communicate with the DSS. Some considerations for the choice of a language are that the formulation should be easily understood, implementable, and modifiable. Moreover, processing the language should be possible on a separate level or on the problem processing system (PPS) level. An obvious choice for a language would be the spoken language of the user. This would require little or no training for the user to interact with a computer, but the complexity of sentences and the use of words that have multiple meanings present difficult problems that, when solved, would introduce unwanted inefficiency into the language system. An alternative would be to use a more formal language based on logic (e.g., PROLOG). The advantage here is that a language can be used at all levels of the DSS. In the design and use of an LS for the user interface, one can consider objectives such as ease of communication, the level of complexity that can be presented by the LS, and the time needed for the user to learn it.
- *Knowledge systems.* The basic function of a knowledge system (KS) is the representation and organization of the "knowledge" in the system. Two possible approaches are

storing the information in database form or representing the data as a base for artificial intelligence using methods from, for example, predicate calculus. The objective of KS is to ease accessibility of data for the DSS. The KS should be able to organize and classify databases and problem domains according to objectives that are sensible and convenient for the user. Some of the objectives in the design of the KS are to reduce the amount of computer memory required, increase the speed with which the data can be retrieved or stored, and increase the number of classifications of data and problem domains possible.

- *Problem processing systems.* The problem processing system of a DSS provides an interface between the LS and the KS. The primary function is to receive the problem from the user via the LS and use the knowledge and data from the KS to determine a solution. Once a solution is found, the PPS sends it through the KS to be translated into a form the user can recognize. More importantly, in the model formulation, analysis, and solution procedure of PPS, the conflicting objectives of stated problems must be considered. The PPS should provide methodology that can optimize all conflicting objectives and generate a compromise solution acceptable to the user. Some of the objectives in the development of such multiobjective approaches are to reduce the amount of time the user must spend to solve the problem, increase the degree of interactiveness (e.g., how many questions the user should answer), reduce the difficulty of questions posed to the user, and increase the robustness of the underlying assumptions and procedures.

ANALYSIS AND DESIGN OF CIM WITH SENSORS AND CONTROL SYSTEMS

Many manufacturing systems are complex, and finding a place to begin a system description is often difficult. Breaking down each function of the system into its lowest possible level and specifying objectives for each level and their interactions will be an effective step. The objectives and decision variables, as related to elements or units for all possible levels, are outlined in the following sections.

Structured Analysis and Design Technique (SADT)

The structured analysis and design technique is a structured methodology. It combines the graphical diagramming language of structured analysis (SA) with the formal thought discipline of a design technique (DT). The advantage of SADT is that it contains a formalized notation and procedure for defining system functions.

Psychological studies have shown that the human mind has difficulty grasping more than five to seven concepts at one time. Based on this observation, SADT follows the structured analysis maxim: "Everything worth saying must be expressed in six or fewer pieces." Limiting each part to six elements ensures that individual parts are not too difficult to understand. Even complex manufacturing systems can be subjected to top-down decomposition without becoming overwhelming.

The basic unit for top-down decomposition is the structural analysis box (Fig. 4.9). Each of the four sides represents a specific action for the SA box. The four actions implement:

- *Input*, measured in terms of different decision variables.
- *Control*, to represent constraints and limitations.
- *Output*, measured in the form of a set of objectives.
- *Mechanism for translation*, which performs the translations (for mapping) of input into output as constrained by the control action.

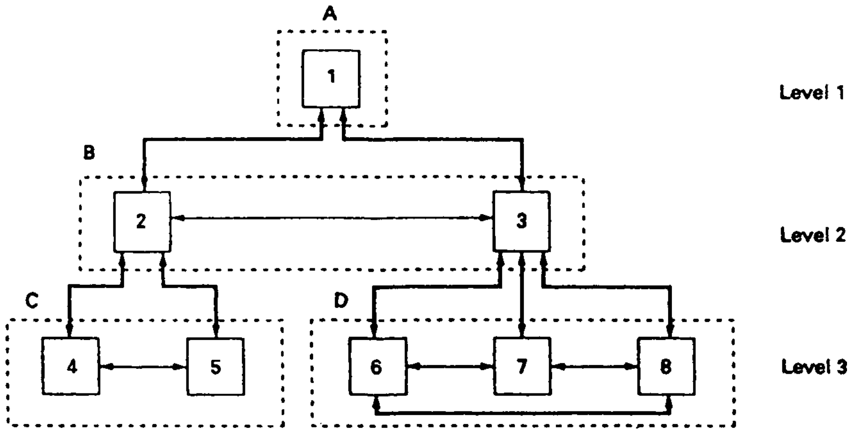


FIGURE 4.9 Structural analysis box.

Since each box represents an idea in the system, each box contains a detailed diagram. A parent-child relationship exists between the box being detailed and the boxes under it. The same relationship holds for the parent diagram as the diagram directly preceding the child diagrams.

The interfaces linking boxes through their neighbors at the same level are the input, output, and control actions. In graphical terms, these interfaces are designated by arrows. Methodically proceeding through a given network, the entire system can be modeled in terms of boxes and arrows.

While SADT provides a realistic approach to modeling any system, it cannot provide the solution to any problem. The integrated computer-aided manufacturing definition method comes one step closer to the realization of a functional CIM system. It utilizes teamwork and demands that all correspondence and analysis be in written form so others can obtain a grasp of the situation and so errors can be more readily located. Because the written word is required during the implementation phases, the documentation usually done at the end of most major projects can be nearly eliminated. Keeping accurate records also plays an important role in debugging the system in the future.

A Multiobjective Approach for Selection of Sensors in Manufacturing

The evaluation of sensors in manufacturing has six criteria:

- *Cost* is simply the price for the sensor and its integrated circuitry if it should be purchased.
- *Integrability* is the degree to which the sensor can be used in conjunction with the manufacturing system it serves. This can usually be measured in terms of compatibility with existing hardware control circuits and software.
- *Reliability* is the quality of the sensors as indicated by the mean time between failures (MTBFs), and can be measured by performing a simple stress test on the sensor under severe limits of operation. If the sensor operates under a certain high temperature for a certain period of time, it will assure the user that the system will perform satisfactorily under normal operating conditions. It will also indicate that the electronic control circuits are reliable, according to the *burn-in philosophy*.

- *Maintenance* involves the total cost to update and maintain the sensor and how often the sensor needs to be serviced.
- *Expandability* is how readily the sensor can be modified or expanded as new needs arise because of a changing environment.
- *User friendliness* indicates the ease of using and understanding the unit. It may include the quality of documentation in terms of simplicity, completeness, and step-by-step descriptions of procedures.

DATA ACQUISITION FOR SENSORS AND CONTROL SYSTEMS IN CIM ENVIRONMENT

The input signals generated by sensors can be fed into an interface board, called an *I/O board*. This board can be placed inside a PC-based system. As personal computers for CIM have become more affordable, and I/O boards have become increasingly reliable and readily available, PC-based CIM data acquisition has been widely implemented in laboratory automation, industrial monitoring and control, and automatic test and measurement.

To create a data acquisition system for sensors and control systems that really meets the engineering requirements, some knowledge of electrical and computer engineering is required. The following key areas are fundamental in understanding the concept of data acquisition for sensors and control systems:

- Real-world phenomena
- Sensors and actuators
- Signal conditioning
- Data acquisition for sensors and control hardware
- Computer systems
- Communication interfaces
- Software

Real-World Phenomena

Data acquisition and process control systems measure real-world phenomena, such as temperature, pressure, and flow rate. These phenomena are sensed by sensors, and are then converted into analog signals that are eventually sent to the computer as digital signals.

Some real-world events, such as contact monitoring and event counting, can be detected and transmitted as digital signals directly. The computer then records and analyzes this digital data to interpret real-world phenomena as useful information.

The real world can also be controlled by devices or equipment operated by analog or digital signals generated by the computer (Fig. 4.10).

Sensors and Actuators

A sensor converts a physical phenomenon such as temperature, pressure, level, length, position, or presence or absence, into a voltage, current, frequency, pulses, and so on.

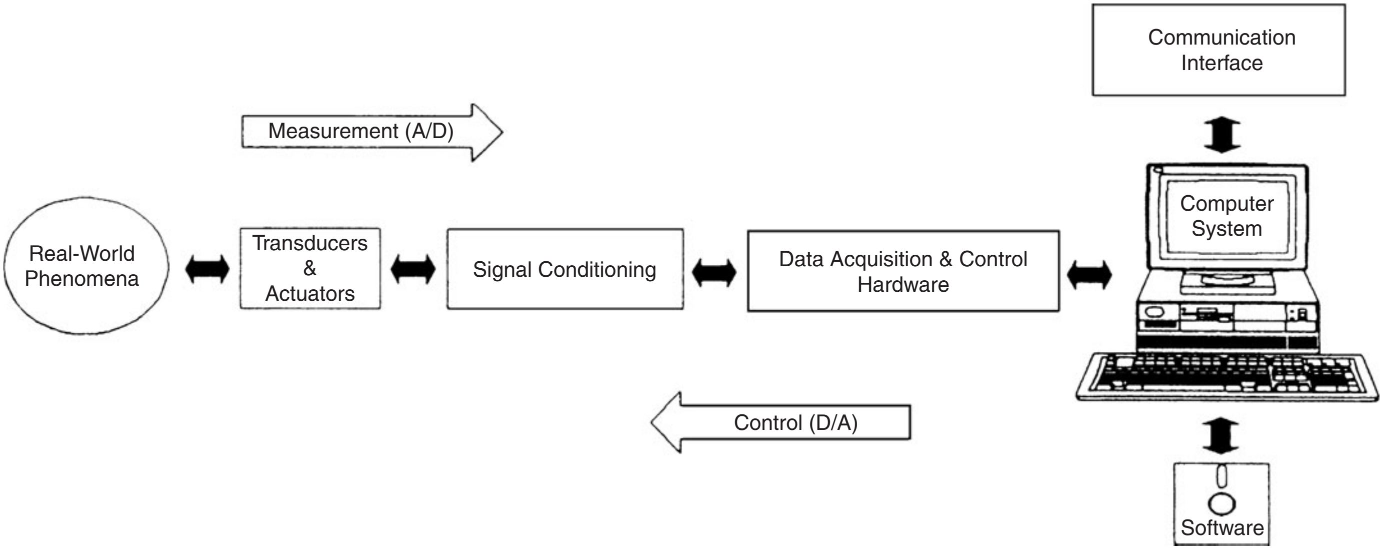


FIGURE 4.10 Integration of computer-controlled devices.

For temperature measurements, some of the most common sensors include thermocouples, thermistors, and resistance temperature detectors (RTDs). Other types of sensors include flow sensors, pressure sensors, strain gauges, load cells, and optical sensors.

An actuator is a device that activates process control equipment by using pneumatic, hydraulic, electromechanical, or electronic signals. For example, a valve actuator is used to control fluid rate for opening and closing a valve.

Signal Conditioning

A signal conditioner is a circuit module specifically intended to provide signal scaling, amplification, linearization, cold junction compensation, filtering, attenuation, excitation, common mode rejection, and so on. Signal conditioning improves the quality of the sensor signals that will be converted into digital signals by the PC's data acquisition hardware.

One of the most common functions of signal conditioning is amplification. Amplifying a sensor signal provides an analog-to-digital (A/D) converter with a much stronger signal and thereby increases resolution. To acquire the highest resolution during A/D conversion, the amplified signal should be equal to approximately the maximum input range of the A/D converter.

Data Acquisition for Sensors and Control Hardware

In general, data acquisition for sensors and control hardware performs one or more of the following functions:

- Analog input
- Analog output
- Digital input
- Digital output
- Counter/timer

Analog Input (A/D). An analog-to-digital converter produces digital output directly proportional to an analog signal input so it can be digitally read by the computer. This conversion is imperative for CIM (Fig. 4.11).

The most significant aspects of selecting A/D hardware are:

- Number of input channels
- Single-ended or differential input
- Sampling rate (in samples per second)

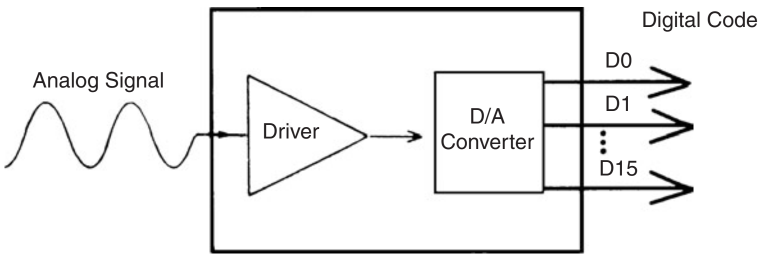


FIGURE 4.11 Analog-to-digital converter.

- Resolution (in bits)
- Input range (specified as full-scale volts)
- Noise and nonlinearity

Analog Output (D/A). A digital-to-analog (D/A) converter changes digital information into a corresponding analog voltage or current. This conversion allows the computer to control real-world events.

Analog output may directly control equipment in a process that is then measured as an analog input. It is possible to perform a closed loop or proportional integral-differential (PID) control with this function. Analog output can also generate waveforms in a function generator (Fig. 4.12).

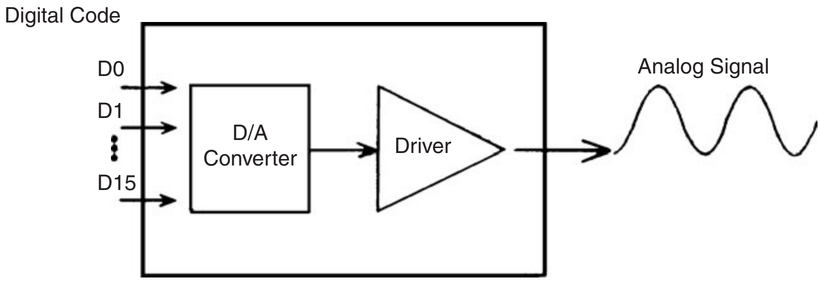


FIGURE 4.12 Analog output.

Digital Input and Output. Digital input and output are useful in many applications, such as contact closure and switch status monitoring, industrial ON/OFF control, and digital communication (Fig. 4.13).

Counter/Timer. A counter/timer can be used to perform event counting, flow meter monitoring, frequency counting, pulse width and time period measurement, and so on.

Most data acquisition and control hardware is designed with the multiplicity of functions previously described on a single card for maximum performance and flexibility. Multifunction data acquisition for high-performance hardware can be obtained through PC boards specially designed by various manufacturers for data acquisition systems.

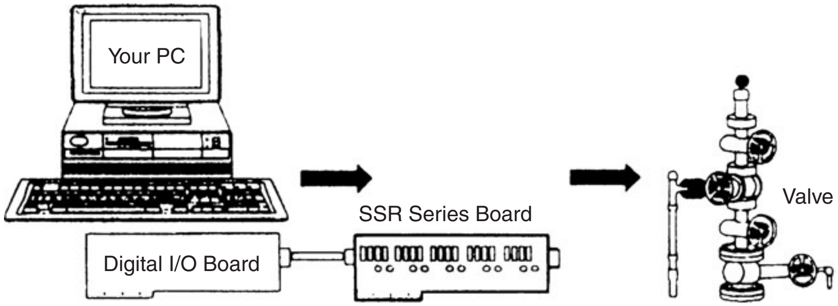


FIGURE 4.13 Application of digital input/output.

Computer System

Today's rapidly growing PC market offers a great selection of PC hardware and software in a wide price range. Thus, a CIM strategy can be economically implemented.

Hardware Considerations. Different applications require different system performance levels. Currently, there are 286, 386, and 486 CPUs, which allow a PC to run at benchmark speeds from 20 up to 150 MHz. Measurements and process control applications usually require 80286 systems. But for applications that require high-speed real-time data analysis, an 80386 or 80486 system will be much more suitable.

Industrial PCs. An *industrial PC* (IPC) is designed specifically to protect the system hardware in harsh operating environments. IPCs are rugged chasses that protect system hardware against excessive heat, dust, moisture, shock, and vibration. Some IPCs are even equipped with power supplies that can withstand temperatures from –20 to +85°C for added reliability in harsh environments.

Passive Backplane and CPU Card. More and more industrial data acquisition for sensors and control systems are using passive backplane and CPU card configurations. The advantages of these configurations are reduced mean time to repair (MTTR), ease of upgrading the system, and increased PC-bus expansion slot capacity.

A passive backplane allows the user to plug in and unplug a CPU card without the effort of removing an entire motherboard in case of damage or repair.

Communication Interfaces

The most common types of communication interfaces used in PC-based data acquisition for sensor and control system applications are RS-232, RS-422/485, and the IEEE-488 general-purpose interface bus (GPIB).

The RS-232 interface is the most widely used interface in data acquisition for sensors and control systems. However, it is not always suitable for distances longer than 50 m or for multidrop network interfaces. The RS-422 protocol has been designed for long distances (up to 1200 m) and high-speed (usually up to 56,000 bit/s) serial data communication. The RS-485 interface can support multidrop data communication networks.

Software

The driving force behind any data acquisition for sensors and control systems is its software control. Programming the data acquisition for sensors and control systems can be accomplished by one of three methods:

- *Hardware-level programming* is used to directly program the data acquisition hardware's data registers. In order to achieve this, the control code values must determine what will be written to the hardware's registers. This requires that the programmer use a language that can write or read data from the data acquisition hardware connected to the PC. Hardware-level programming is complex, and requires significant time—time that might be prohibitive to spend. This is the reason that most manufacturers of data acquisition hardware supply their customers with either driver-level or package-level programs.
- *Driver-level programming* uses function calls with popular programming languages such as C, PASCAL, and BASIC, thereby simplifying data register programming.

- *Package-level programming* is the most convenient technique of programming the entire data acquisition system. It integrates data analysis, presentation, and instrument control capabilities into a single software package. These programs offer a multitude of features, such as pull-down menus and icons, data logging and analysis, and real-time graphic displays.

DEVELOPING CIM STRATEGY WITH EMPHASIS ON SENSORS' ROLE IN MANUFACTURING

To develop a comprehensive CIM strategy incorporating sensors and control systems, an enterprise must begin with a solid foundation, such as a CIM architecture. A CIM architecture is an information system structure that enables the industrial enterprise to integrate information and business processes. It accomplishes this by (1) establishing the direction that integration will take and (2) defining the interfaces between the users and the providers of this integration function.

Figure 4.14 shows how CIM architecture answers the enterprise's integration means. A CIM architecture provides a core of common services. These services support every other

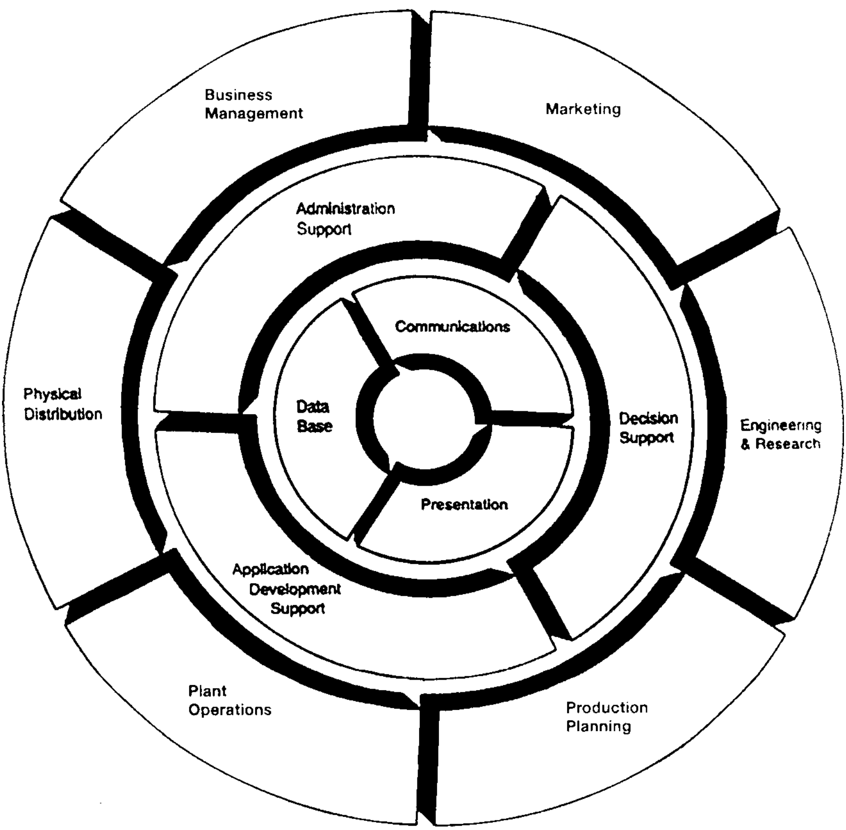


FIGURE 4.14 CIM architecture.

area in the enterprise, from its common support function to its highly specialized business processes.

CIM and Building Blocks

The information environment of an industrial enterprise is subject to frequent changes in system configuration and technologies. A CIM architecture incorporating sensors and control systems can offer a flexible structure that enables it to react to the changes. This structure relies on a number of modular elements that allow systems to change easily to grow with the enterprise's needs.

Figure 4.17 shows a modular structure that gives CIM flexibility. It is based on three key building blocks:

- *Communications.* The communication and distribution of data
- *Data management.* The definition, storage, and use of data
- *Presentation.* The presentation of data to people and devices throughout the enterprise

Utilizing the building blocks, CIM can provide a base for integrating the enterprise's products, processes, and business data. It can define the structure of the hardware, software, and services required to support the enterprise's complex requirements. It can also translate information into a form that can be used by the enterprise's people, devices, and applications.

CIM Communications

Communications—the delivery of enterprise data to people, systems, and devices—is a critical aspect of CIM architecture. This is because today's industrial environment brings together a wide range of computer systems, data acquisition systems, technologies, system architectures, operating systems, and applications. This range makes it increasingly difficult for people and machines to communicate with each other, especially when they describe and format data differently.

Various enterprises, in particular IBM, have long recognized the need to communicate data across multiple environments. IBM's response was to develop *systems network architecture* (SNA) in the 1970s. SNA supports communication among different IBM systems, and over the years it has become the standard for host communications in many industrial companies.

However, the CIM environment with sensor communications must be even more integrated. It must expand beyond individual areas, throughout the entire enterprise, and beyond—to customers, vendors, and subcontractors.

Communications in the CIM environment involves a wide range of data transfer, from large batches of engineering or planning data to single-bit messages from a plant floor device. Many connectivity types and protocols must be supported so the enterprise's people, systems, and devices can communicate. This is especially true in cases where response time is critical, such as during process alerts.

Plant Floor Communications

Plant floor communications can be the most challenging aspect of factory control. This is due to the wide range of manufacturing and computer equipment that have been used in production tasks over the decades.

IBM's solution for communicating across the systems is the IBM plant floor series, a set of software products. One of these products, *Distributed Automation Edition* (DAE), is a systems enabler designed to provide communication functions that can be utilized by plant floor applications. These functions include:

- Defining and managing networks
- Making logical device assignments
- Managing a program library for queuing and routing messages
- Establishing alert procedures
- Monitoring work-cell status

With these functions, Distributed Automation Edition can (1) help manufacturing engineers select or develop application programs to control work-cell operations and (2) provide communication capabilities between area- and plant-level systems (Fig. 4.15).

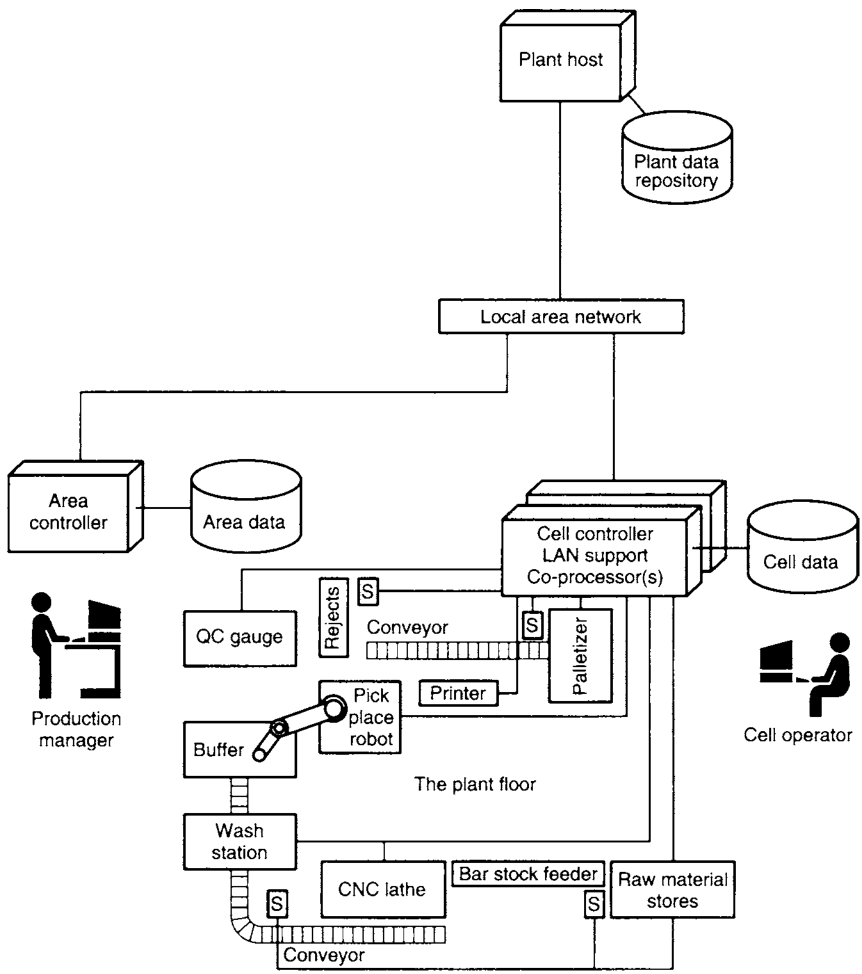


FIGURE 4.15 Distributed Automation Edition (DAE).

DAE supports several communication protocols to meet the needs of a variety of enterprises. For example, it supports SNA for connections to plant host systems and the IBM PC network, as well as the IBM token-ring protocol and manufacturing automated protocol (MAP) for plant floor communications. MAP is the evolving plant floor communication industry standard, adopted by the International Standards Organization for communications among systems provided by different vendors.

Managing Data in the CIM Environment

The second building block of a CIM architecture incorporating sensors and control technology is data management. This includes how data are defined, how different data elements are related, where data are stored, and who has access to that data. Data management is particularly critical in today's industrial environment since a multitude of different databases, formats, and storage and access techniques exist.

Standards are evolving. For example, *Structured Query Language* (SQL) provides a medium for *relational database* applications and for users to access a database. Unfortunately, a significant amount of data exists today in other database technologies that are not accessible by current standards.

Data management defines and records the location of the data created and used by the enterprise's business functions. Data management also means enabling users to obtain the data needed without having to know where the data are located.

Relationships among several data elements must be known if data are to be shared by users and applications. In addition, other data attributes are important in sharing data. These include the type of data (text, graphic, image), their status (working, review, completed), and their source (person, application, or machine).

In CIM with sensory architecture, data management can be accomplished through three individual storage functions: (1) the data repository, (2) the enterprise data storage, and (3) the local data files.

Some of the key data management functions—the repository, for example—are already being implemented by the *consolidated design file* (CDF) established through the IBM *Data Communication Service* (DCS).

The consolidated design file operates on a relational database and is built on SQL. One example of its use is as an engineering database to integrate CAD/CAM applications with the business needs of the engineering management function. This environment, IBM's DCS/CDF, provides the following repository functions:

- Transforming data to a user-selected format
- Storing CAD/CAM data
- Adding attributes to CAD/CAM data
- Enabling users to query data and attributes

DCS/CDF also provides communications functions to transfer data between the repository and CAD/CAM applications (Fig. 4.16).

CIM Environment Presentation

Presentation in the CIM environment means providing data to and accepting data from people and devices. Obviously, this data must assume appropriate data definitions and screen formats to be usable.

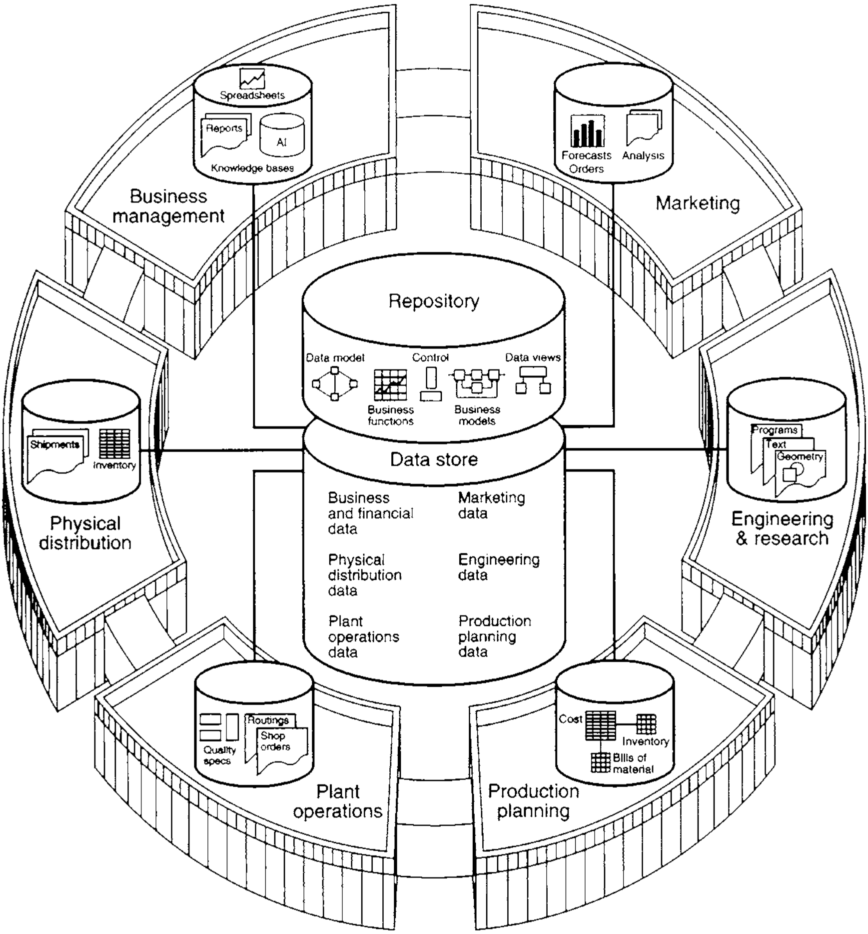


FIGURE 4.16 Data Communication Service/consolidated design file (DCS/CDF).

Because today's industrial enterprise contains such a wide array of devices and information needs, it must have a consistent way to distribute and present information to people, terminals, workstations, machine tools, robots, sensors, bar-code readers, automated guided vehicles, and part storage and retrieval systems. The range of this information covers everything from simple messages between people to large data arrays for engineering design and applications (Fig. 4.17). It may originate from a CIM user in one functional area of the enterprise and be delivered to a CIM user or device in another area.

In today's environments, presentation occurs on displays that utilize various technologies. Some are nonprogrammable terminals, some are programmable workstations, and some are uniquely implemented for each application. As a result, the same information is often treated differently by individual applications.

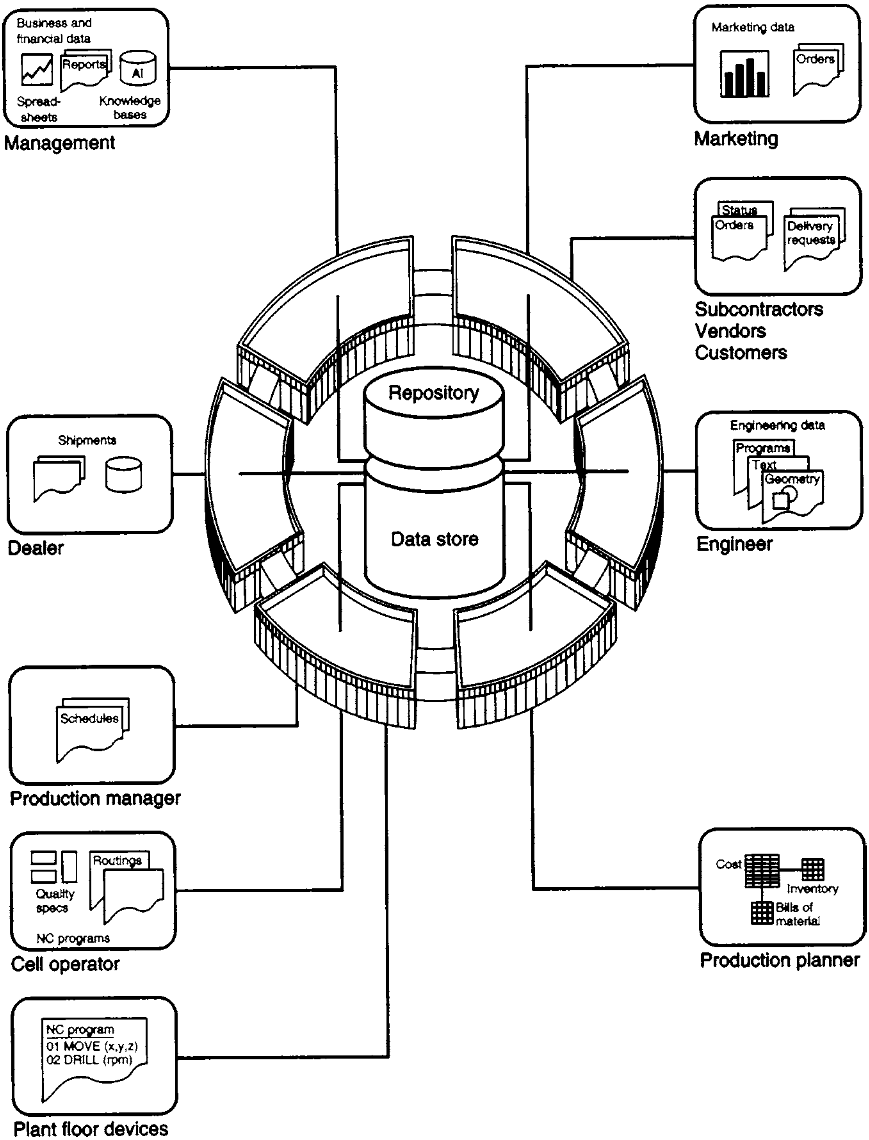


FIGURE 4.17 Presentation of data.

For example, the same manufactured part may be referred to as a part number in a bill of material in production planning, as a drawing in engineering's CAD application, and as a routing in a paperless shop order from plant operations.

As data are shared across the enterprise, they must be transformed into definitions and formats that support the need of individual users and applications. Applications must be able to access shared data, collect the required information, and then format that information for delivery.

The Requirement for Integration

Communication, data management, and presentation each have their own set of technical requirements. In addition, before these three building blocks can be integrated, a CIM architecture must also address a number of enterprise-wide constraints. For example, a CIM architecture should be able to:

- Utilize standard platforms
- Integrate data
- Protect the installed base investment
- Work with heterogeneous systems
- Utilize industry-standard operator interfaces
- Reduce application support cost
- Provide a customizable solution
- Offer phased implementation
- Deliver selectable functions
- Improve the business process

Standard Platforms. Utilizing standard computing platforms is one step industrial enterprises can take toward integration. Today, many products are available that utilize standard platforms. These include processors, operating systems, and enablers for communications, database management, and presentation. In addition, platforms such as IBM's *systems application architecture* (SAA) and *advance interactive executive* (AIX) help make application consistency a reality across strategic IBM and UNIX operating system environments.

SAA, for example, is a comprehensive IBM blueprint for consistency and compatibility of software products. SAA begins the process by establishing definitions for four key application aspects: common user access, common programming interface, common communication support, and common applications. Through these definitions, SAA will support the development of new applications across major operating environments.

AIX is IBM's version of the UNIX operating system and combines consistent user and application interfaces to aid in the development of an integrated application across UNIX environments. AIX consists of six related system enablers:

- Base systems
- Programming
- Interface
- User interface
- Communication support
- Distributed processing and applications

Data Integration. Integration requirements are often met by creating bridges between individual applications. Bridges usually copy a collection of data between two applications. A bridge between engineering and production planning allows these two functions to share a bill of material. Another bridge permits an engineering CAD/CAM application to download an NC program to a plant floor personal computer. Or a bridge between production planning and plant operations may be used to provide a copy of the production schedule to the plant floor system.

However, a problem with bridges is that changes made to the original set of data are not immediately incorporated into the copy of the data. This results in out-of-date information. Another problem is that bridges become difficult to maintain when more than two applications must work together.

As enterprises begin to integrate their operations, it will be imperative that the latest information is shared among multiple applications and across business functions. For example, engineering, marketing, cost accounting, production planning, and plant operations may all need access to inventory status information. At other times, the enterprise's various business functions may need information about product specifications, order status, operating cost, and more.

A CIM architecture must be able to simplify and accelerate this integration. It must provide the facilities to integrate data across the various applications of the business functions—facilities such as data query, data communication, controlled access and editing, and consistent data definitions.

Installed Base Investment. Today's industrial enterprises have made considerable investments in their installed bases, including systems, data, and even training. In the United States alone, manufacturers spend billions of dollars per annum on information systems hardware, software, and integration services for production planning, engineering, and plant operations. CIM with sensory technology must help protect this investment by permitting the integration of existing systems, applications, and data.

Heterogeneous Systems. In today's heterogeneous environment, data are located on different systems and in different formats. Applications have different needs, which are answered by processors, communications, and displays utilizing different technologies and architectures.

In an enterprise model, production planning may automate its operations on a single mainframe using an interactive database. Engineering may store drawings in an engineering database, then design and analyze products on a network of graphics workstations. Plant operations and sensors and control systems may be automated with personal computers and specialized machine controllers connected by both standard and proprietary networks. The data needed to operate the enterprise are scattered across all these diverse systems.

The heterogeneous environment is also characterized by an installed system base provided by multiple computer system suppliers, software vendors, and systems integrators. A CIM architecture must allow the integration of these varied system solutions and operating platforms.

Industry Standards and Open Interfaces. As integration technologies mature, there will be the need to support an expanding set of industry standards. Today these standards include communication protocols such as MAP, token ring, and Ethernet; data exchange formats such as the *initial graphics exchange specifications* (IGES) for engineering drawings; data access methods like SQL; and programming interfaces such as *programmer's hierarchical interactive graphics standard* (PHIGS). A CIM architecture must be able to accommodate these and other evolving standards. One framework for accomplishing this has already been established, the open systems architecture for CIM (CIM-OSA). CIM-OSA is being defined in the *Esprit* program by a consortium of European manufacturers, universities, and information system suppliers, including IBM. Data exchange formats are also being extended to accommodate product definition in the *product definition exchange specification* (PDES). In addition, a CIM architecture must be able to support well-established solutions, such as IBM's SNA, which have become de facto standards.

In this competitive marketplace, manufacturers must also be able to extend operations as needed and support these new technologies and standards as they become available. These needs may include adding storage to a mainframe system, replacing engineering workstations, installing a new machine tool, upgrading an operating system, and utilizing new software development tools. A CIM architecture with open interfaces will allow enterprises to extend the integration implementation over time to meet changing business needs.

Reduced Application Support Cost. A CIM architecture incorporating sensors must also yield application solutions at a lower cost than traditional stand-alone computing. This includes reducing the time and labor required to develop integrated applications and data. It also means reducing the time and effort required to keep applications up to speed with the changes in the enterprise's systems environment, technology, and integration needs.

Customizable Solutions. Every enterprise has its own business objectives, shared data, system resources, and applications. For example, one enterprise may choose to address CIM requirements by reducing the cycle time of product development. It does this by focusing on the data shared between the product development and process development functions in accelerating the product release process.

Another enterprise's aim may be to reduce the cycle time required to deliver an order to a customer. So it addresses this by exchanging data between production planning and operations functions and automating the order servicing and production processes. It must also be able to customize operations to individual and changing needs over time.

Phased Implementation. Implementing enterprise-wide integration will take place in many small steps instead of through a single installation. This is because integration technology is still evolving, implementation priorities are different, installed bases mature at different rates, new users must be trained, and lessons will be learned in pilot installations.

As enterprises begin their implementation efforts in phases, they will be integrating past, present, and future systems and applications. A CIM architecture must be able to support the integration of this diverse installed base.

Selectable Functions. Most enterprises will want to weigh the benefits of integration against the impact this change will bring to each application and set of users. For example, an emphasis on product quality may require that production management gain greater insight into the quality of individual plant operations activities by implementing advanced sensors and control systems developed for particular applications. When an existing shop floor control application adequately manages schedules and support shop personnel with operating instructions, some additional information, such as that on quality, may be added, but rewriting the existing application may not be justified.

However, the plant manager may plan to develop a new production monitoring application to operate at each workstation. This application will make use of various sensors, share data with shop floor control, and utilize software building blocks for communications, database management, and presentation.

As is evident, the existing application requires only a data sharing capability, while the new application can benefit from both data sharing and the architecture building blocks. A CIM architecture with selectable functions will provide more options that can support the variety of needs within an enterprise.

Improved Business Process. Obviously, an enterprise will not implement integration on the basis of its technical merits alone. A CIM architecture must provide the necessary business benefits to justify change and investment.

The integration of information systems must support interaction between business functions and the automation of business processes. This is a key function if corporate goals, such as improved responsiveness to customer demands and reduced operating cost, are to be met. A CIM architecture must provide the means by which an entire enterprise can reduce the cycle times of business processes required for order processing, custom offerings, and new products. It must also reduce the impact of changes in business objectives and those business processes.

REFERENCES

1. Bucker, D. W., "10 Principles to JIT Advancement," *Manufacturing Systems* (March) 55 (1988).
2. Campbell, J., *The RS-232 Solution*, Sybex, Inc., Alameda, Ca., 1984.
3. Clark, K. E., "Cell Control, The Missing Link to Factory Integration," International Industry Conference Proceedings, Toronto (May) 641–646 (1989).
4. Datapro Research Corporation, "How U.S. Manufacturing Can Thrive," *Management and Planning Industry Briefs* (March) 39–51 (1987).
5. Groover, M. P., and E. W. Zimmer, Jr., *CAD/CAM: Computer Aided Design and Manufacturing*, Prentice Hall, Englewood Cliffs, N.J., 1984.
6. IBM Corp., *Introducing Advanced Manufacturing Applications*, IBM, Atlanta, Ga., 1985.
7. "APICS: Tweaks and Distributed Systems Are the Main Focus—MRP No Longer Looks to the Future for Finite Capacity Scheduling," *Managing Automation* (January) 31–36 (1992).
8. Manufacturing Studies Board, National Research Council, *Toward a New Era in U.S. Manufacturing: The Need for a National Vision*, National Academy Press, Washington, D.C., 1986.
9. Orlicky, J., *Material Requirements Planning*, McGraw-Hill, New York, 1985.
10. Schonberger, R. J., "Frugal Manufacturing," *Harvard Business Review* (September–October) 95–100 (1987).
11. Skinner, W., *Manufacturing: The Formidable Competitive Weapon*, Wiley, New York, 1985.
12. "Support for the Manufacturing Floor," *Manufact. Eng.* (March) 29–30 (1989).
13. Vollmann, T. E., W. L. Berry, and D. C. Whyback, *Manufacturing Planning and Control Systems*, Richard D. Irwin, Homewood, Ill., 1984.

CHAPTER 5

ADVANCED SENSOR TECHNOLOGY IN PRECISION MANUFACTURING APPLICATIONS

IDENTIFICATION OF MANUFACTURED COMPONENTS

In an automated manufacturing operation, one should be able to monitor the identification of moving parts. The most common means of automatic identification is *bar-code technology*. However, other approaches offer advantages under certain conditions.

Bar-Code Identification Systems

The *universal product code* (UPC) used in retail stores is a standard 12-digit code. Five of the digits represent the manufacturer and five the item being scanned. The first digit identifies the type of number system being decoded (a standard supermarket item, for example) and the second is a parity digit to determine the correctness of the reading. The first six digits are represented by code in an alternating pattern of light and dark bars. Figure 5.1 shows two encodings of the binary string 100111000. In both cases, the minimum printed width is the same. The delta code requires nine such widths (the number of bits), while the width code requires 13 such widths (if a wide element is twice the width of the narrow element). Different bar widths allow for many character combinations. The remaining six digits are formed by dark alternating with light bars reversing the sequence of the first six digits. This allows backward scanning detection (Fig. 5.2).

A bar-code reader can handle several different bar-code standards, decoding the stripes without knowing in advance the particular standard. The military standard, code 1189, specifies the type of coding to be used by the Department of Defense, which is a modification of code 39. Code 39 consists of 44 characters, including the letters A through Z. Because of its alphanumeric capabilities, code 39 is very effective for manufacturing applications. Code 39 is structured as follows: three of nine bars (light and dark) form wide characters; the rest are narrow.

Bar-code labels are simple to produce. Code 39, for example, can be generated by a personal computer. Such labels are ideal for inventory identifications and other types

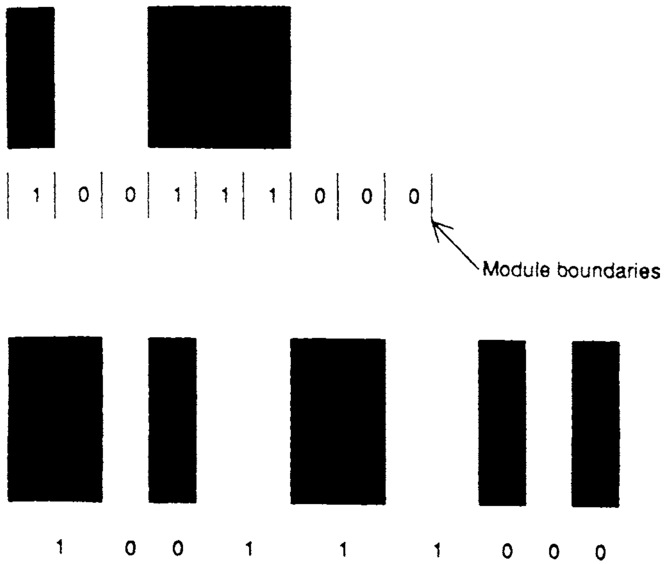


FIGURE 5.1 Encoding of the binary string 10011100 by delta code (top) and width code (bottom).

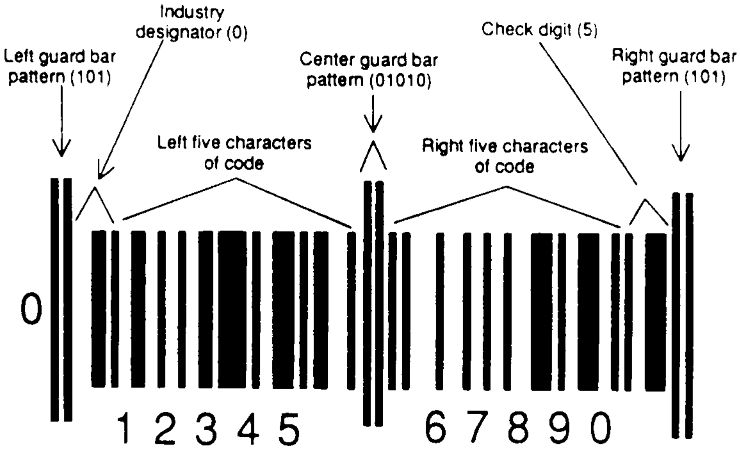


FIGURE 5.2 Specifications of the UPC symbol. The readable characters are normally printed in the OCR-B font.

of fixed-information gathering. Bar codes are not necessarily placed on labels. Tools, for example, have had the code etched on their surfaces to allow for tool tracking. Techniques have been developed for molding bar codes onto rubber tires. Holographic scanners allow reading around corners so that parts need not be oriented perpendicular to the reader as they feed down a processing line.

A difficulty with bar coding has been the fact that it cannot be read if the bars become obscured by dirt, grease, or other substances. Infrared scanners are used to read codes that are coated with black substances to prevent secrecy violations through reproduction of the codes. One way to generally offset the problem of a dirty environment is to use magnetic-stripe-encoded information.

Transponders

While bar-code labels and magnetic stripes are very effective on the shop floor, shop circumstances may require more information to be gathered about a product than can be realistically handled with encoded media. For instance, with automobiles being assembled to order in many plants, significant amounts of information are necessary to indicate the options for a particular assembly. Radio-frequency (RF) devices are used in many cases. An RF device, often called a transponder, is fixed to the chassis of a car during assembly. It contains a chip that can store a great amount of information. A radio signal at specific assembly stations causes the transponder to emit information that can be understood by a local receiver. The transponder can be coated with grease and still function. Its potential in any assembly operation is readily apparent. Several advanced transponders have read/write capability, thus supporting local decision making.

Electromagnetic Identification of Manufactured Components

Many other possible electronic schemes can identify manufactured parts in motion. Information can be coded on a magnetic stripe in much the same way that bars represent information on a bar-code label, since the light and dark bars are just a form of binary coding.

Operator identification data are often coded on magnetic stripes that are imprinted on the operators' badges. Magnetic stripe information can be fed into a computer. Such information might include the following: (1) the task is complete, (2) x number of units have been produced, (3) the unit part numbers, (4) the operator's identification number, and so on. This same scanning station can also be set up using bar-code information; however, with magnetic striping, the information can be read even if the stripe becomes coated with dirt or grease. A disadvantage of magnetic striping is that the reader has to contact the stripe in order to recall the information.

Surface Acoustic Waves

A process similar to RF identification is *surface acoustic waves* (SAW). With this process, part identification is triggered by a radar-type signal that can be transmitted over greater distances than in RF systems.

Optical Character Recognition

Another form of automatic identification is *optical character recognition* (OCR). Alphanumeric characters form the information, which the OCR reader can "read." In mail processing centers, high-speed sorting by the U.S. Postal Service is accomplished using OCR. The potential application to manufacturing information determination is obvious.

Many other means exist for part identification, such as vision systems and voice recognition systems. *Vision systems* utilize TV cameras to read alphanumeric data and transmit the information to a digital converter. OCR data can be read with such devices, as can conventionally typed characters. *Voice recognition systems* have potential where an individual's arms and hands are utilized in some function that is not conducive to reporting information. Such an application might be the inspection of parts by an operator who has to make physical measurements on the same parts.

In *laser scanning applications*, a laser beam scans and identifies objects at a constant speed. The object being scanned interrupts the beam for a time proportional to its diameter or thickness. Resolutions of less than 1 mm are possible.

In *linear array applications*, parallel light beams are emitted from one side of the object to be measured to a photo-optical diode array on the opposite side. Diameters are measured by the number of array elements blocked. Resolutions of 5 mm or greater are possible.

In *TV camera applications*, a TV camera is used in the digitizing of the image of an object and the result is compared to the stored image. Dimensions can be measured, part orientation can be determined, and feature presence can be checked. Some exploratory work is being accomplished with cameras that can fit in a tool changer mechanism. The camera can be brought to the part like a tool and verify part characteristics.

DIGITAL ENCODER SENSORS

Digital encoder sensors provide directly an output in a digital form and thus require only simple signal conditioning. They are also less susceptible to electromagnetic interference, and are therefore useful for information processing and display in measurement and control systems. Their ability to rapidly scan a series of patterns provides additional manufacturing automation opportunities when light and dark patterns are placed in concentric rings in a disk. Figure 5.3 illustrates a portion of such a disk that can be rigidly attached to a shaft or an object and housed in an assembly containing optical sensors for each ring (Fig. 5.4). The assembly, called an *optical encoder*, automatically detects the rotation of a shaft or an object. The shaft rotation information can be fed back into

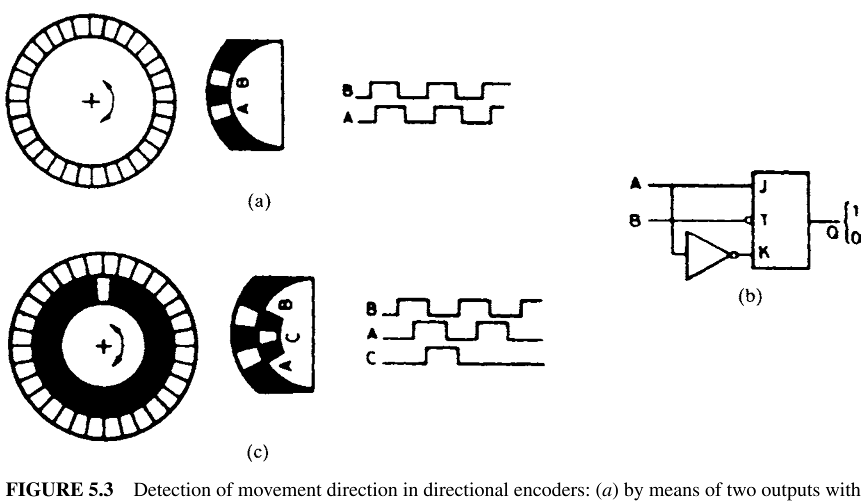


FIGURE 5.3 Detection of movement direction in directional encoders: (*a*) by means of two outputs with 90° phase shift; (*b*) output electronic circuit; (*c*) additional marker for absolute positioning.

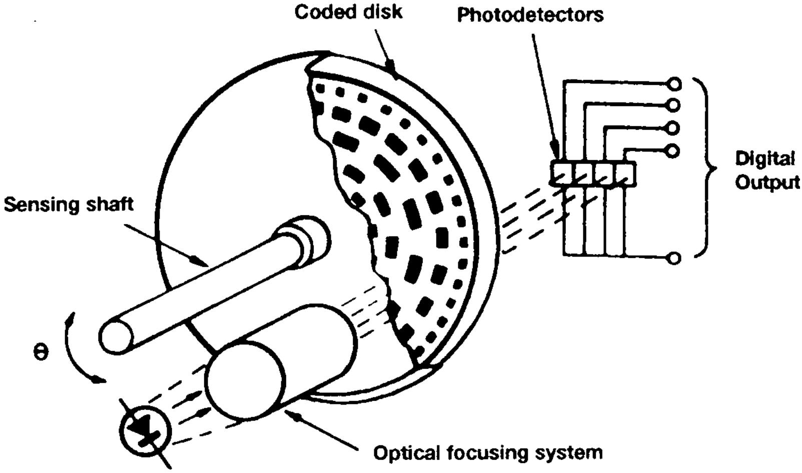


FIGURE 5.4 Principle of absolute position encoders for linear and rotary movements.

a computer or controller mechanism for controlling the velocity or position of the shaft. Such a device has application in robots and numerical control machine tools and for precision measurements of strip advancement to generate a closed-loop feedback actuation for displacement compensation.

The two classes of digital encoder sensors include:

- Digital encoder sensors yielding at the output a digital data although an analog input signal was applied. This class of encoder sensors includes position encoders.
- Digital encoder sensors that rely on some physical oscillatory phenomenon transduced by a conventional modulating sensor. This class of sensors may require an electronic circuit acting as a digital counter in order to yield a desired digital output signal.

There are no sensors where the transduction process directly yields a digital output. The usual process is to convert an analog input quantity into a digital signal by means of a sensor without the requirement to convert an analog voltage into its digital equivalent.

Position Encoder Sensors in Manufacturing

Position encoder sensors can be categorized as linear and angular position encoder sensors. The optical encoder sensor can be either incremental or absolute. The incremental types transmit a series of voltages proportional to the angle of rotation of the shaft or object. The control computer must know the previous position of the shaft or object in order to calculate the new position. Absolute encoders transmit a pattern of voltages that describes the position of the shaft at any given time. The innermost ring reaches from dark to light every 180°, the next ring every 90°, the next 45°, and so on, depending on the number of rings on the disk. The resulting bit pattern output by the encoder reveals the exact angular position of the shaft or object. For an absolute optical encoder disk that has eight rings and eight LED sensors, and in turn provides 8-bit outputs, the result

TABLE 5.1 Absolute Optical Encoder

Encoder ring	Angular displacement, degrees	Observed pattern	Computed value, degrees
1 (innermost)	180	1	180
2	90	0	
3	45	0	
4	22.5	1	22.5
5	11.25	0	
6	5.625	1	5.625
7	2.8125	1	2.8125
8	1.40625	0	
			210.94

is 10010110. Table 5.1 shows how the angular position of the shaft or object can be determined.

The incremental position encoder sensor suffers three major weaknesses:

- The information about the position is lost whenever the electric supply fails or the system is disconnected, and when there are strong perturbations.
- The digital output, to be compatible with the input/output peripherals of a computer, requires an up/down counter.
- The incremental position encoder does not detect the movement direction unless elements are added to the system (Fig. 5.5).

Physical properties used to define the disk pattern can be magnetic, electrical, or optical. The basic output generated by the physical property is a pulse train. By differentiating the signal, an impulse is obtained for each rising or falling edge, increasing by two the number of counts obtained for a given displacement (Fig. 5.6).

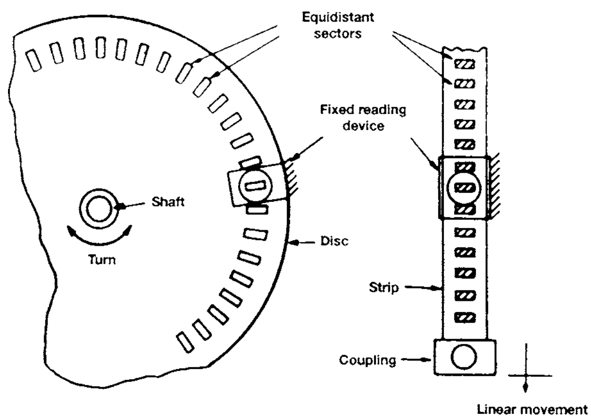


FIGURE 5.5 Principle of linear and rotary incremental position encoders.

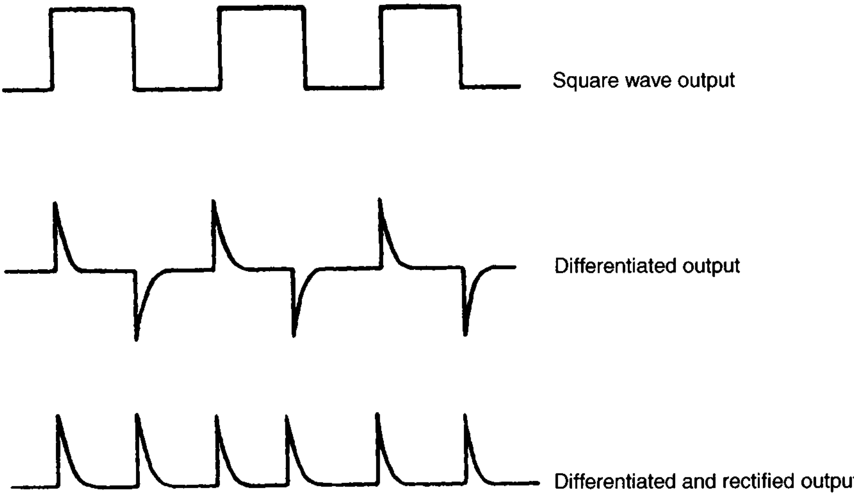


FIGURE 5.6 Improving output resolution of an incremental encoder by differentiation and rectification.

FUZZY LOGIC FOR OPTOELECTRONIC COLOR SENSORS IN MANUFACTURING

Fuzzy logic will most likely be the wave of the future in practical and economical solutions to control problems in manufacturing. Fuzzy logic is simply a technique that mimics human reasoning. This technology is now being explored throughout various industries. Fuzzy logic color sensors can relay information to microprocessors to determine color variance within an acceptable range of colors. A conventional sensor could not perform this function because it could choose only a specific color and reject all other shades—it uses a very precise set of rules to eliminate environmental interference.

The research and development activities for fuzzy logic technology began in mid-1990. The research has lead to the creation of a fuzzy logic color sensor that can learn a desired color and compare it with observed colors. The sensor can distinguish between acceptable and unacceptable colors for objects on a conveyer belt. The development of new light source technology allows the color sensor to produce more accurate color measurement (Fig. 5.7). Also, the integration of the fuzzy logic sensor with a microprocessor enables the data to be collected and interpreted accurately.

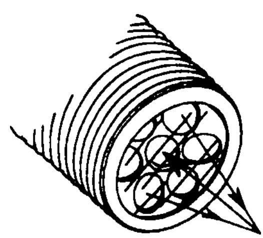


FIGURE 5.7 Integration of a fuzzy logic sensor in a sensing module.

Sensing Principle

The sensor is designed with a broad-spectrum solid-state light source utilizing a light-emitting diode cluster. The LED-based light source provides stable, long-lasting, high-speed target illumination capabilities. The LED cluster is made up of three representative hues of *red*, *green*, and *blue*, which provide a triple bell-shaped spectral power distribution for the light

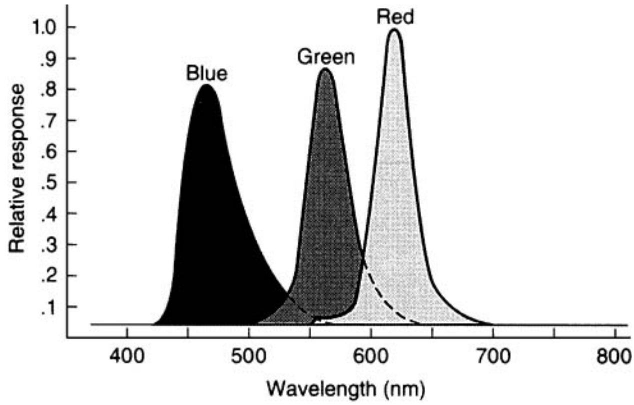


FIGURE 5.8 Spectra of red, blue, and green light.

source (Fig. 5.8). The light incident on the target is reflected with varying intensities, depending on the particular target color under analysis.

The reflected light is received by a semiconductor receiver in the center of the LED cluster. The amount of light reflected back onto the receiver is transduced to a voltage and converted to digital format immediately via an analog-to-digital converter. The internal processing of the converted red, green, and blue (RGB) values offers variable sample size and averaging to compensate for signal noise (Fig. 5.8).

Ambient light is sampled between every component pulse and immediately subtracted from the sampled signal so that the effects of factory ambient light are suppressed. Thus, hooding the sensor is not totally necessary. In an area of a very bright or high-frequency lighting, it may be beneficial to provide some degree of hooding to at least limit the brightness. The sensor's electronics also employs temperature compensation circuitry to stabilize readings over temperature ranges.

Color Theory

Color science defines color in a space, with coordinates of *hue*, *saturation*, and *intensity* (HSI). These three general components uniquely define any color within HSI color space. Hue is related to the reflected wavelength of a color when a white light is shined on it. Intensity (lightness) measures the degree of whiteness, or gray scale, of a given color. Saturation is a measure of the vividness of a given hue. The term *chromaticity* primarily includes elements of the hue and saturation components. Researchers depict color in space using hue as the angle of a vector, saturation as the length of it, and intensity as a plus or minus height from a center point (Fig. 5.9).

The concepts of hue, saturation, and intensity can be further clarified by a simplified pictorial presentation. Consider Fig. 5.10, where a color is depicted at a molecular level. Color is created when light interacts with pigment molecules. Color is generated by the way pigment molecules return (bend) incoming light. For example, a red pigment causes a measurable hue component of the color. The relative density of the pigment molecules leads to the formation of the saturation component. Some molecules are present that return almost all wavelengths, and appear white as a result, leading to the intensity (lightness) component.

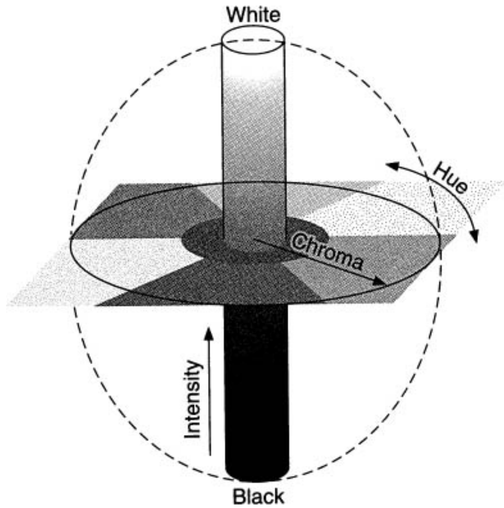


FIGURE 5.9 Coordinates of the hue, saturation, and intensity of color in space.

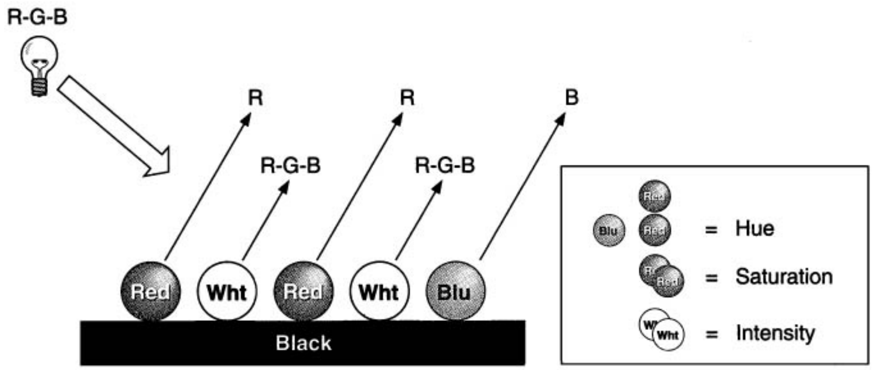


FIGURE 5.10 A model of color interpretation.

Units of Color Measurement

If color description depends on measuring the interaction of a target color with a given white light source, it is clear that in order to have the system of measurement standardized, both the light source and means of detection must be well-defined. One very popular set of standardization rules has been set up by the Commission International de l'Eclairage (CIE), a color standardization organization. From color theory, it is known that the response to a color stimulus (its determination), depends on the spectral power distribution of the light source (illuminant), times the spectral reflectance of the target (color) surface, times the spectral response of the detector (observer) (Fig. 5.11).

With this principle in mind, the CIE presented a detailed description of the standard light source and a standard observer (photodetector). The result of the study was the popular CIE diagram, which creates a two-dimensional mapping of a color space (Fig. 5.12).

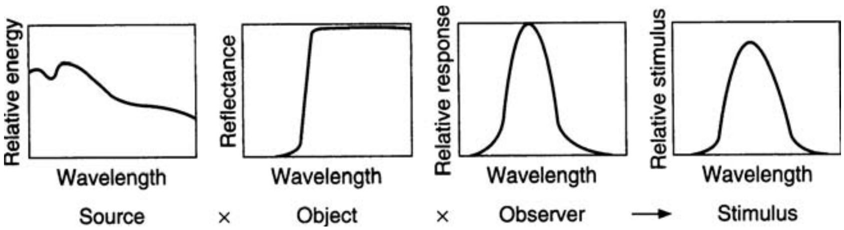


FIGURE 5.11 Stimulus response to a color (detector/determination) = illuminant × target × observer.

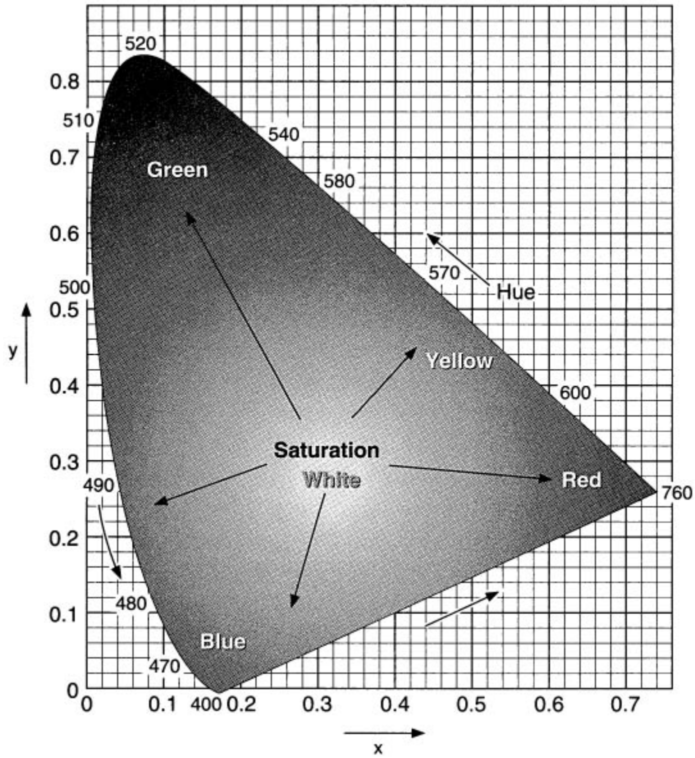


FIGURE 5.12 Two-dimensional mapping of color in a space.

Further manipulation of the CIE observations has lead to another color coordinate system, the so-called L.a.b numbers for describing a color. The L.a.b. numbering system is fairly prevalent in industrial applications. The machines that measure color according to this theory are referred to as *color spectrophotometers* or *color meters*. These machines are typically expensive, bulky, and not well-suited for distributed on-line color sensing.

The fuzzy color sensor does not offer CIE-based color measurement; however, it is a very high resolution color comparator. The sensor learns a color with its own standard light source (trio-stimulus LEDs) and its own observer (semiconductor photoreceiver). It thereby sets up its own unique color space with the three dimensional coordinates being the red, blue, and green readings (Fig. 5.12).

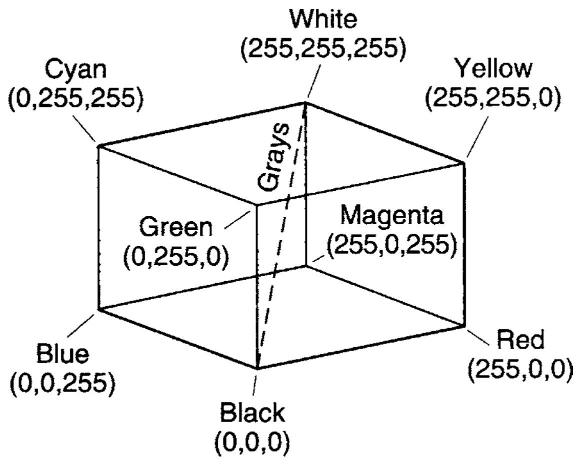


FIGURE 5.13 Three-dimensional coordinates of color in a space.

Theoretically there would be 256^3, or 16,777,216, unique positions in its color space for defining color (Fig. 5.13). In reality, the actual number of colors a sensor can reliably distinguish is much less because of optical noise and practical limitations of the design.

The device compares the RGB colors it observes to the internally learned standard. Each time a standard color is relearned, it essentially recalibrates the sensor.

Color Comparators and True Color Measuring Instruments

If the learned color is initially defined by a color spectrometer, the fuzzy logic color sensor (comparator) can be installed in conjunction with it. The color sensor can learn the same area that the spectrometer has read, thereby equating its learned standard to the absolute color. By using the color sensor in this manner, a temporary correlation to an absolute color standard can be established (Fig. 5.14). This permits the relative drift from the standard to be monitored. The advantage is that more color sensing can be economically distributed across

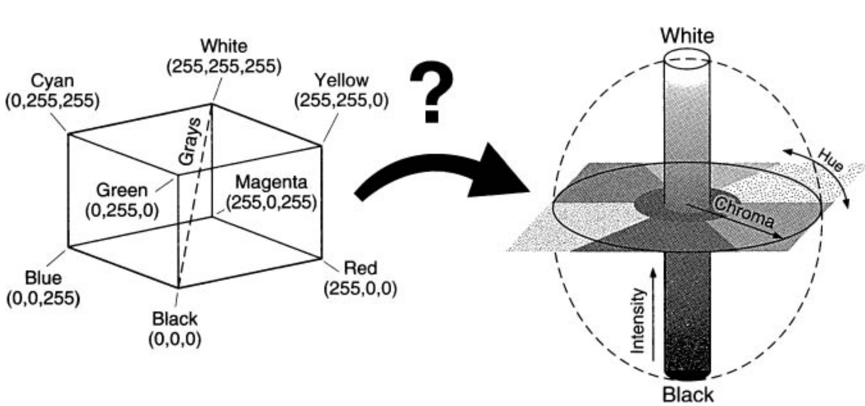


FIGURE 5.14 Correlation to an absolute color standard.

the target area. When a significant relative deviation is detected, an alert signal can flag the absolute color sensor to take a reading in the suspect area. If enough storage space is available in a central processing computer, lookup tables can be constructed to relate the serially communicated color sensor readings to a standard color coordinate system like the CIE system.

Color Sensor Algorithms

Two internal software algorithms, or sets of rules, are used for analyzing fuzzy logic color sensor data:

- The absolute algorithm: compares color on the basis of absolute voltages
- The relative algorithm: compares color on the basis of relative percentages of each RGB component voltage

The choice of algorithm depends on the sensing distance variation and the type of color distinction one needs. If the outputs vary excessively with distance when the absolute algorithm is used, a relative (ratio) algorithm must be considered. While a relative algorithm does not retain the lightness information, it greatly reduces unwanted distance-related variations. The relative algorithm shows changes in chromaticity (hue and chroma) that exist in most color differences. If the sensing distances can be held constant, the absolute algorithm works well at detecting subtle changes in density (shades) of a single color.

Design Considerations in Fuzzy Logic Color Sensors

The design of fuzzy logic color sensors aims to achieve maximum color sensing ability, while maintaining the expected simplicity of operation and durability of typical discrete industrial sensors. Several other key goals must be considered in system design such as high speed, small size, configurability, high repeatability, and long light source life.

The choice of a solid-state light source satisfies the majority of the criteria for a good industrialized color sensor design. The fuzzy logic color sensor utilizes two sets of three different LED photodiodes as its illumination source. The three LED colors, red, green, and blue, were chosen essentially for their coverage of the visible light spectrum (Fig. 5.15).

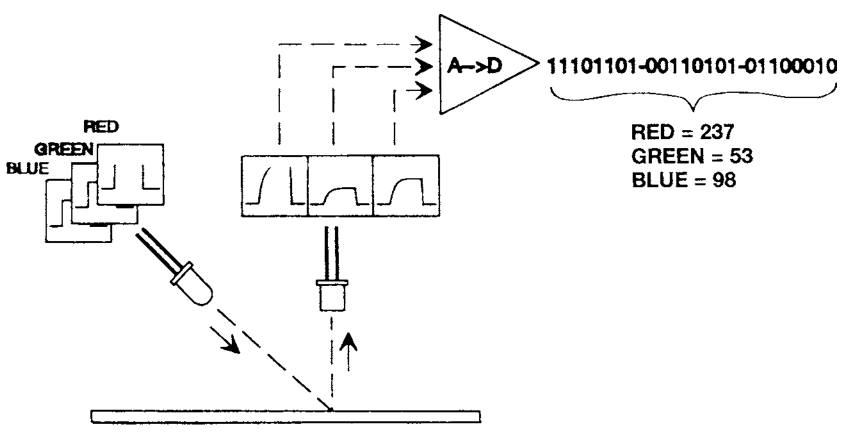


FIGURE 5.15 Conversion of red, green, and blue LED outputs from analog to digital.

The light from each LED is sequentially pulsed onto the target and its reflected energy is collected by a silicon photoreceiver chip in the LED cluster. Ambient light compensation circuitry is continually refreshed between each LED pulse, so the reported signals are almost entirely due to the LED light pulses. The LED sources offer a very fast (microsecond response) and stable (low spectral drift, steady power) source of a given wavelength band, without resorting to filters. Emerging blue LEDs, in combination with the more common red and green LEDs, have made it possible to use three solid-state

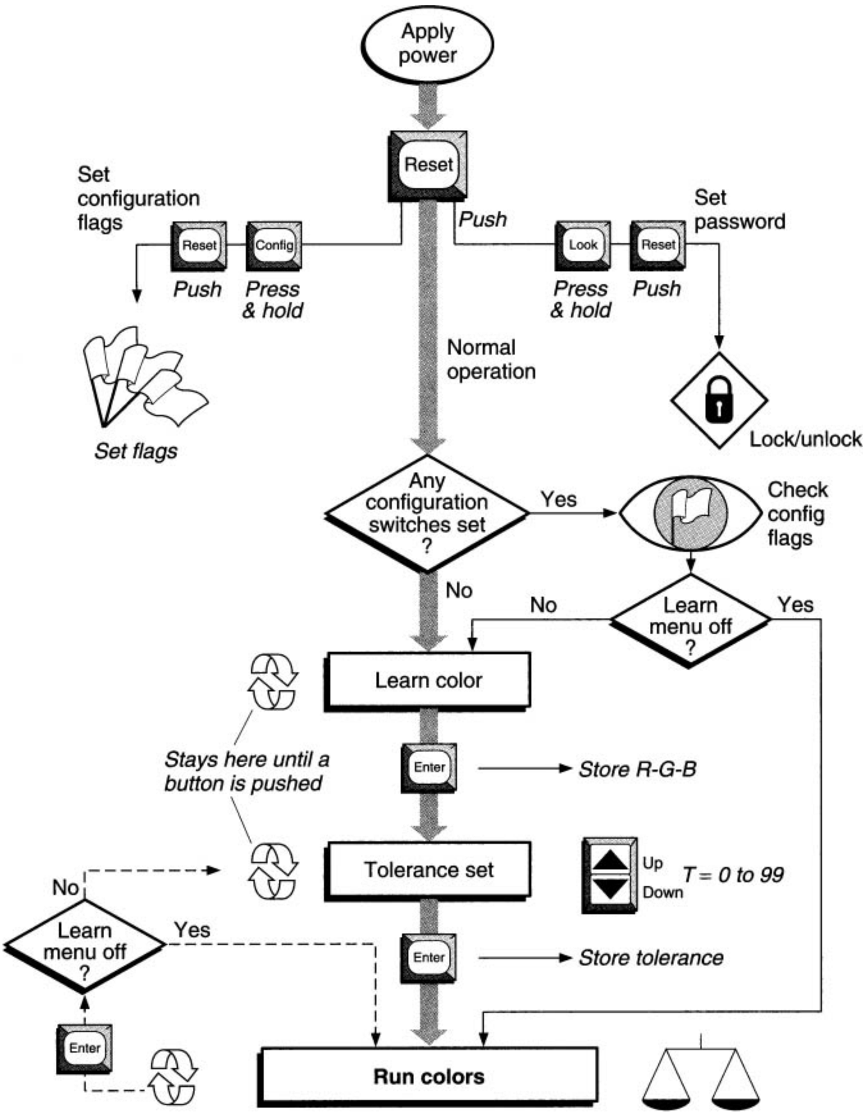


FIGURE 5.16 A fuzzy-logic controller flowchart.

spectra to define a hue. The choice of the specific LED bandwidth (or spectral distribution) is made so as to obtain the best color distinction through broader coverage of the illumination spectrum.

Fuzzy Logic Controller Flowchart

The entire sensor operation is illustrated in Fig. 5.16. An internal microcontroller governs the device operations. It directs signals in and out of the sensor head, to maintain local and remote communications, and provides color discrimination algorithms to produce the appropriate signal output at the control pin. As the device proceeds out of reset, it checks the locally (or remotely) set configuration dip-switches, which in turn define the path, or operating menu, to proceed through. There is permanent storage of learned or remotely loaded values of RGB readings, tolerance, number of readings to average, and the white-card calibration value, so these settings can be available at reset or at power-up.

One or more of three optional menus can be selected before entering into the main operation of learning and running colors. The three alternative menus are (1) white-card gain set, (2) number of reads to average set, and (3) the observed stored reading menus.

If none of the alternative menus is activated, the sensor will proceed directly to the primary modes, which are the learn, tolerance set, and run modes (Fig. 5.17). By pressing and holding the appropriate buttons while pushing the reset button, two other programming menus can be entered. These are the set configuration flags (dip-switch) menu and the set password menu.

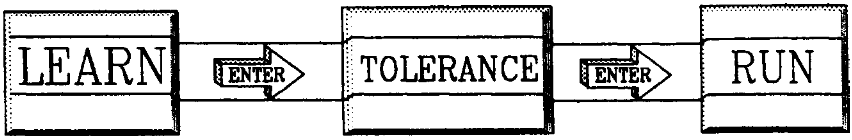


FIGURE 5.17 Simple path of operation.

SENSORS DETECTING FAULTS IN DYNAMIC MACHINE PARTS (BEARINGS)

A system consisting of analog and digital signal processing equipment, computers, and computer programs would detect faults in ball bearings in turbomachines and predict the remaining operating time until failure. The system would operate in real time, extracting the diagnostic and prognostic information from vibrations sensed by accelerometers, strain gauges, acoustical sensors, and from the speed of the machine as measured by a tachometer.

The vibrations that one seeks to identify are those caused by impacts that occur when pits in balls make contact with races and pits in races make contact with balls. These vibrations have patterns that are unique to bearings and repeat at known rates related to ball-rotation, ball-pass, and cage-rotation frequencies. These vibrations have a wide spectrum that extends up to hundreds of kilohertz, where the noise component is relatively low.

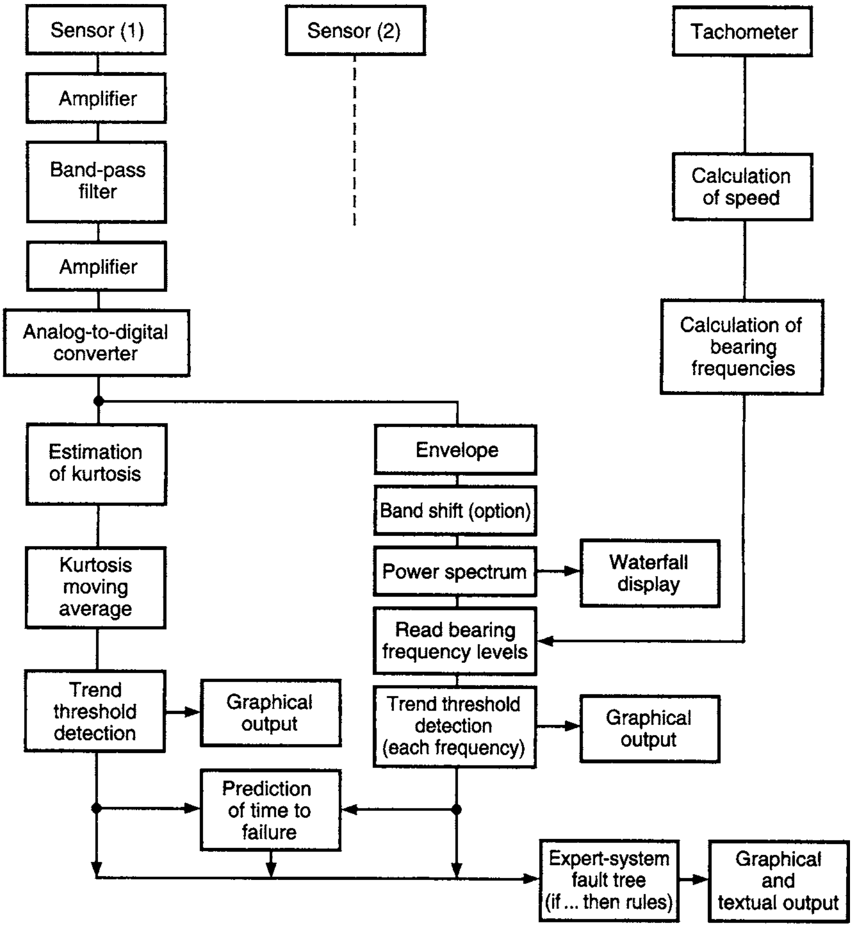


FIGURE 5.18 Data flow for an automatic bearing fault detection system.

The system in Fig. 5.18 would accept input from one of two sensors. Each input signal would be amplified, bandpass-filtered, and digitized. The digitized signal would be processed in two channels: one to compute the keratosis of the distribution of the amplitudes, the other to calculate the frequency content of the envelope of the signal. The *keratosis* is the fourth statistical moment and is known, from theory and experiment, to be indicative of vibrations caused by impact on faults. The keratosis would be calculated as a moving average for each consecutive digitized sample of the signal by using a number of samples specified by the technician. The trend of a keratosis moving average would be computed several times per second, and the changes in the keratosis value deemed to be statistically significant would be reported.

In the other signal processing channel, the amplitude envelope of the filtered digitized signal would be calculated by squaring the signal. Optionally, the high-frequency sample data would be shifted to a lower frequency band to simplify processing by use of a Fourier transformation. This transformation would then be applied to compute the power spectrum.

The output of the tachometer would be processed in parallel with the spectral calculations so the frequency bins of the power spectrum could be normalized on the basis of the speed of rotation of the machine. The power spectrum would be averaged with a selected number of previous spectra and presented graphically as a "waterfall" display; this is similar to a sonar display with which technicians can detect a discrete frequency before an automatic system can.

The bearing frequencies would be calculated from the measured speed and the known parameters of the bearings, with allowances for slip. The power spectrum levels would be read for each bearing frequency; a moving average of the amplitude at each bearing frequency and harmonic would be maintained, and trends representing statistically significant increases would be identified by threshold detection and indicated graphically.

By using algorithms based partly on analyses of data from prior tests, the results of both keratosis and power spectrum calculations would be processed onto predictions of the remaining operating time until failure. All the results would then be processed by an expert system. The final output would be a graphical display and text that would describe the condition of the bearings.

SENSORS FOR VIBRATION MEASUREMENT OF A STRUCTURE

An advanced sensor was developed to gauge structure excitations and measurements that yield data for design of robust stabilizing control systems (Fig. 5.19).

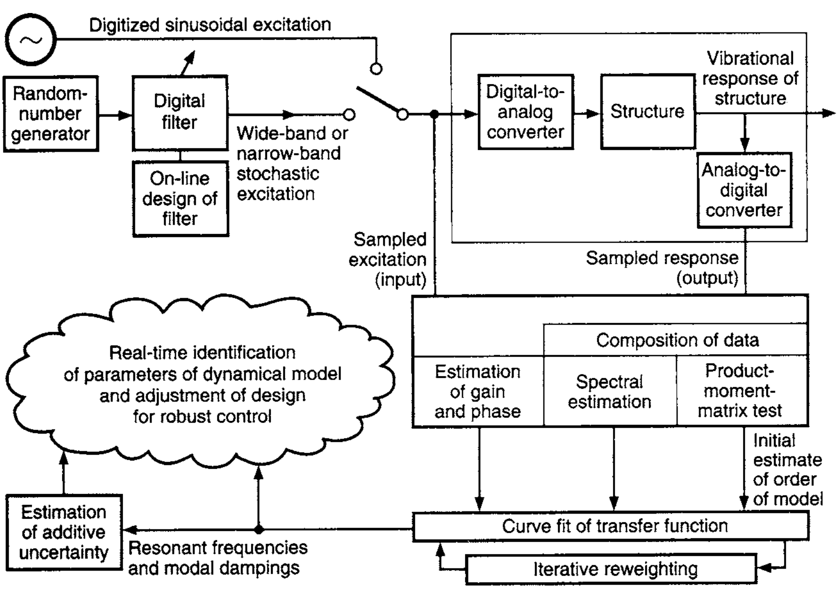


FIGURE 5.19 Automated characterization of vibrations of a structure.

An automated method for characterizing the dynamic properties of a large flexible structure estimates model parameters that can be used by a robust control system to stabilize the structure and minimize undesired motions. Although it was developed for the control of large flexible structures in outer space, the method is also applicable to terrestrial structures in which vibrations are important—especially aircraft, buildings, bridges, cranes, and drill rigs.

The method was developed for use under the following practical constraints:

- The structure cannot be characterized in advance with enough accuracy for purposes of control.
- The dynamics of the structure can change in service.
- The numbers, types, placements, and frequency responses of sensors that measure the motions and actuators that control them are limited.
- Time available during service for characterization of the dynamics is limited.
- The dynamics are dominated by a resonant mode at low frequency.
- In-service measurements of the dynamics are supervised by a digital computer and are taken at a low rate of sampling, consistent with the low characteristic frequencies of the control system.
- The system must operate under little or no human supervision.

The method is based on extracting the desired model and control-design data from the response of the structure to known vibrational excitations (Fig. 5.19). Initially, wideband stochastic excitations are used to obtain the general characteristics of the structure. Narrow-band stochastic and piece-wise-constant (consistent with sample-and-hold discretizations) approximations to sinusoidal excitations are used to investigate specific frequency bands in more detail.

The relationships between the responses and excitations are first computed non-parametrically—by spectral estimation in the case of stochastic excitations and by estimation of gains and phases in the case of approximately sinusoidal excitations. In anticipation of the parametric curve fitting to follow, the order of a mathematical model of the dynamics of the structure is estimated by use of a *product moment matrix* (PMM). Next, the parameters of this model are identified by a least-squares fit of transfer-function coefficients to the nonparametric data. The fit is performed by an iterative reweighting technique to remove high-frequency emphasis and assure minimum-variance estimation of the transfer-function coefficient. The order of the model starts at the PMM estimate and is determined more precisely thereafter by successively adjusting a number of modes in the fit at each iteration until an adequately small output-error profile is observed.

In the analysis of the output error, the additive uncertainty is estimated to characterize the quality of the parametric estimate of the transfer function and for later use in the analysis and design of robust control. It can be shown that if the additive uncertainty is smaller than a certain calculable quantity, then a conceptual control system could stabilize the model structure and could also stabilize the real structure. This criterion can be incorporated into an iterative design procedure. In this procedure, each controller in a sequence of controllers for the model structure would be designed to perform better than the previous one did, until the condition for robust capability was violated. Once the violation occurred, one could accept the penultimate design (if its performances were satisfactory) or continue the design process by increasing a robustness weighting (if available). In principle, convergence of this iterative process guarantees a control design that provides high performance for the model structure while guaranteeing robustness of stability to all perturbations of the structure within the additive uncertainty.

OPTOELECTRONIC SENSOR TRACKING TARGETS ON A STRUCTURE

The location and exact position of a target can be accurately sensed through optoelectronic sensors for tracking a retroreflective target on a structure. An optoelectronic system simultaneously measures the positions of as many as 50 retroreflective targets within 35° of view with an accuracy of 0.1 mm. The system repeats the measurement ten times per second. The system provides an unambiguous indication of the distance to each target that is not more than 75 m away from its sensor module. The system is called a *spatial high-accuracy position-encoding sensor* (SHAPES).

SHAPES fills current needs in the areas of system identification and control of large flexible structures, such as large space- and ground-based antennas and elements of earth-orbiting observational platforms. It is also well-suited to applications in rendezvous and docking systems. Ground-based applications include boresight determination and precise pointing of 70-m deep-space-network antennas.

SHAPES illuminates the retroreflective targets by means of a set of lasers in its sensor module. In a typical application (Fig. 5.20) a laser diode illuminates each target with 30-ps

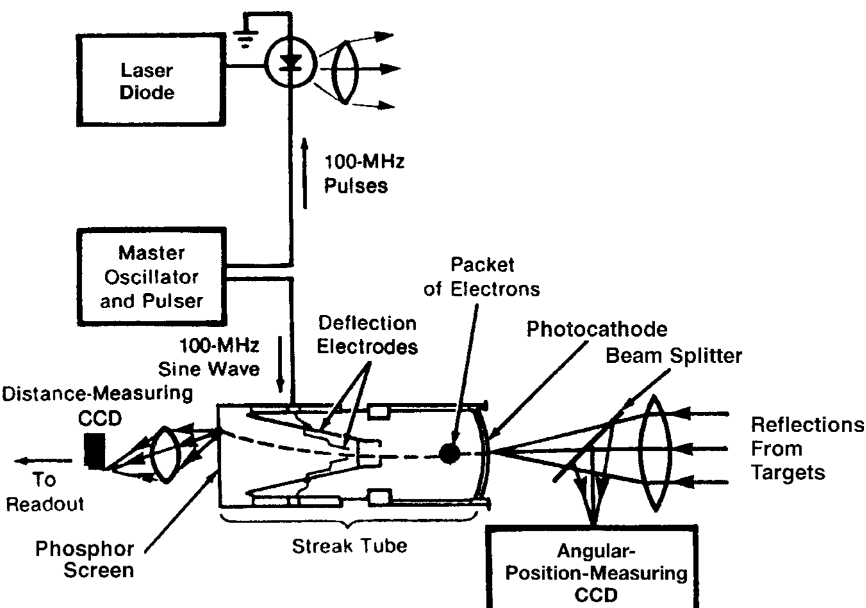


FIGURE 5.20 Beam splitter diverts reflections from a continuous-wave laser into a CCD camera for measurement of angles of reflection.

pulses at a repetition rate of 100 MHz. Light reflected from the target is focused by a lens and passed through a beam splitter to form images on a charge-coupled device (CCD) and on the photocathode of a streak tube. The angular position of the target is determined simply from the position of its reflection on the charge-coupled device.

The measurement of the distance to the target is based on the round-trip time of the optical pulses. The round-trip distance can be measured in terms of the difference between the phase of the train of return pulses incident on the photocathode and the phase of a reference sine wave that drives the deflection plate of the streak tube. This difference, in turn, manifests itself as a displacement between the swept and unswept positions, at the output end of the streak tube, of the spot of light that represents the reflection from the target. The output of the streak tube is focused on a CCD for measurement and processing of the position of this spot. Three microprocessors control the operation of SHAPES and convert the raw data required from the angular-position and distance-measuring CCDs into position of the target in three dimensions.

OPTOELECTRONIC FEEDBACK SIGNALS FOR SERVOMOTORS THROUGH FIBER OPTICS

In what is believed to be among its first uses to close a digital motor-control loop, fiber-optics transmission provides immunity to noise and rapid transmission of data. An optoelectronic system effects closed-loop control of the shaft angle of four servomotors and could be expanded to control as many as 16. The system includes a full-duplex fiber-optic link (Fig. 5.21) that carries feedforward and feedback digital signals over a distance of many meters, between commercial digital motor-control circuits that execute a PID control algorithm with programmable gain (one such control circuit dedicated to each servomotor) and modules that contain the motor-power switching circuits, digital-to-analog buffer circuits for the feedforward control signals, and analog-to-digital buffer circuits for the feedback signals from the shaft-angle encoders (one such module located near, and dedicated to, each servomotor).

Besides being immune to noise, optical fibers are compact and flexible. These features are particularly advantageous in robots, which must often function in electromagnetically noisy environments and in which it would otherwise be necessary to use many stiff bulky wires (which could interfere with movement) to accommodate the required data rates.

Figure 5.21 shows schematically the fiber-optic link and major subsystems of the control loop of one servomotor. Each digital motor-control circuit is connected to a central control computer, which programs the controller gains and provides the high-level position commands. The other inputs to the motor-control circuit include the sign of the commanded motor current and pulse-width modulation representing the magnitude of the command motor current.

The fiber-optic link includes two optical fibers—one for feedforward, one for feedback. The ends of the fibers are connected to identical bidirectional interface circuit boards, each containing a transmitter and a receiver. The fiber-optic link has a throughput rate of 175 MHz; at this high rate, it functions as though it were a 32-bit parallel link (8 bits for each motor control loop), even though the data are multiplexed into a serial bit stream for transmission. In the receiver, the bit stream is decoded to reconstruct the 8-bit pattern and a programmable logic sequencer expands the 8-bit pattern to 32 bits and checks for errors by using synchronizing bits.

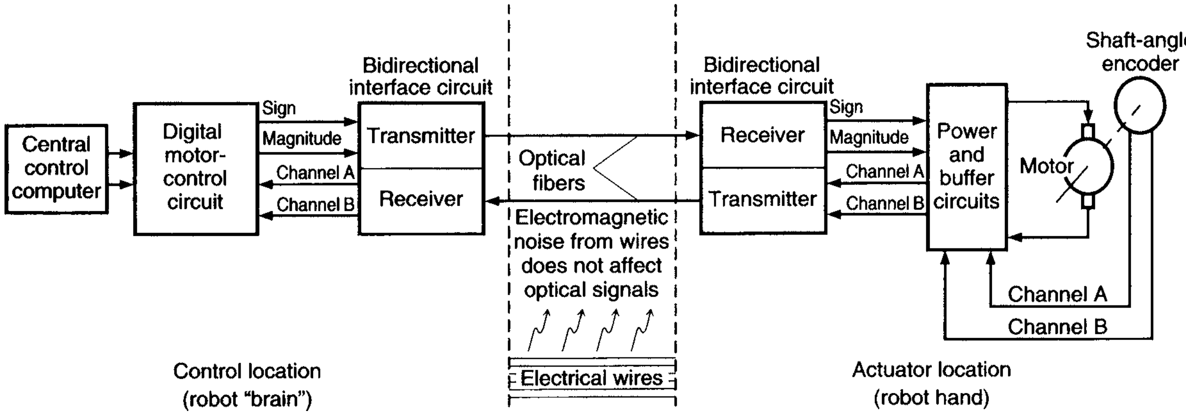


FIGURE 5.21 Full-duplex fiber-optic transmission link.

ACOUSTOOPTICAL/ELECTRONIC SENSOR FOR SYNTHETIC-APERTURE RADAR UTILIZING VISION TECHNOLOGY

An acoustooptical sensor operates in conjunction with analog and digital electronic circuits to process frequency-modulated *synthetic-aperture radar* (SAR) return signals in real time. The acoustooptical SAR processor will provide real-time SAR imagery aboard moving aircraft or space SAR platforms. The acoustooptical SAR processor has the potential to replace the present all-electronic SAR processors that are currently so large and heavy and consume so much power that they are restricted to use on the ground in the postprocessing of the SAR in-flight data recorder.

The acoustooptical SAR processor uses the range delay to resolve the range coordinates of a target. The history of the phase of the train of radar pulses as the radar platform flies past a target is used to obtain the azimuth (cross-range) coordinate by processing it coherently over several returns. The range-compression signal processing involves integration in space, while the azimuth-compression signal processing involves integration in time.

Figure 5.22 shows the optical and electronic subsystems that perform the space and time integrations. The radar return signal is heterodyned to the middle frequency of an

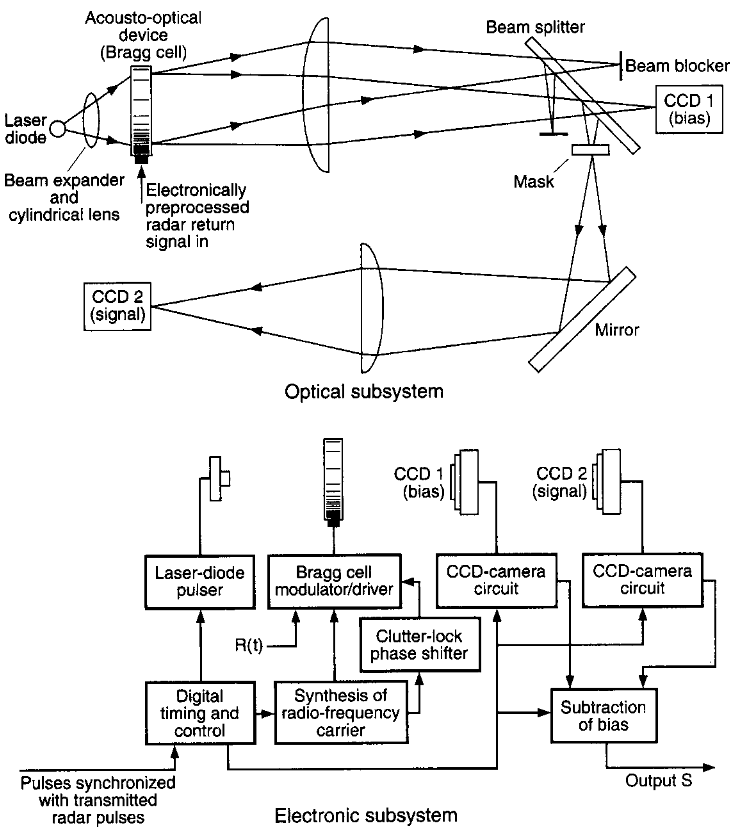


FIGURE 5.22 Acoustooptical synthetic-aperture radar.

acoustooptical sensor and added electronically to a reference sinusoid to capture the history of the phase of the return signal interferometrically for compression in azimuth. The resulting signal is applied to the acoustooptical sensor via a piezoelectric transducer. The acoustooptical sensor thus becomes a cell that encodes the evolving SAR return.

Meanwhile, pulses of light a few tens of nanoseconds long are generated by a laser diode in synchronism with the transmitted pulses and are used to sample and process the return signal. Lenses shape that laser light into a plane wave incident upon the acoustooptical sensor. The integration in space is effected at the moment of sampling by the focusing action. The position of the focal point in the cell depends on the range delay of the corresponding target, and light is brought to focus on two CCD imaging arrays at positions that depend on the range.

The sinusoidal reference signal component of the cell interacts with laser radiation to generate a plane wave of light that interferes with the light focused by the cell. This produces interference fringes that encode the phase information in the range-compressed optical signal. These fringes are correlated with a mask that has a predetermined spatial distribution of density and that is placed in front of, or on, one of the CCD arrays. This CCD array is operated in a delay-and-integrate mode to obtain the desired correlation and integration in time for the azimuth compression. The output image is continuously taken from the bottom picture element of the CCD array.

Two CCDs are used to alleviate a large undesired bias of the image that occurs at the output as a result of optical processing. CCD_1 is used to compute this bias, which is then subtracted from the image of CCD_2 to obtain a better image.

THE USE OF OPTOELECTRONIC/VISION ASSOCIATIVE MEMORY FOR HIGH-PRECISION IMAGE DISPLAY AND MEASUREMENT

Storing an image of an object often requires large memory capacity and a high-speed interactive controller. Figure 5.23 shows schematically an optoelectronic associative memory that responds to an input image by displaying one of *M* remembered images. The decision about which if any of the remembered images to display is made by an optoelectronic analog computation of an inner-product-like measure of resemblance between the input image and each of the remembered images. Unlike associative memories implemented as

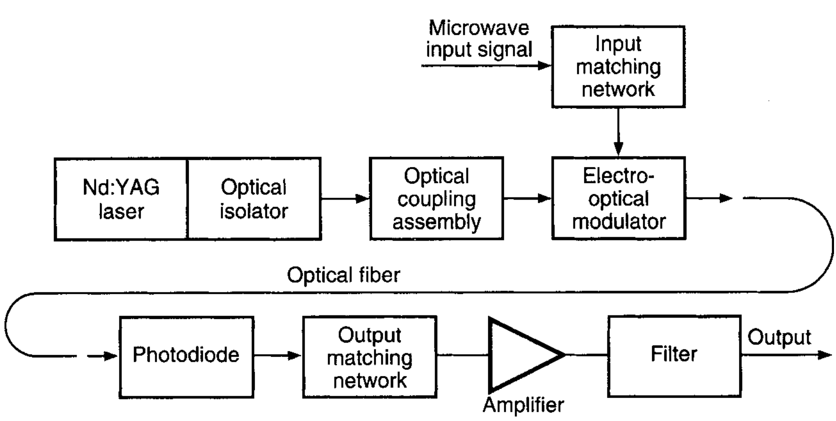


FIGURE 5.23 A developmental optoelectronic associative memory.

all-electronic neural networks, this memory does not rely on the precomputation and storage of an outer-product synapse matrix. Instead, the optoelectronic equivalent of this matrix is realized by storing remembered images in two separate spatial light modulators placed in tandem. This scheme reduces the required size of the memory by an order of magnitude.

A partial input image is binarized and displayed on a liquid-crystal light valve spatial modulator that reprocesses the image in real time by operating in an edge-enhancement mode. This preprocessing increases the orthogonality (with respect to the inner product) between the input image and each of the remembered images, thereby increasing the ability of the memory to discriminate among different images.

The light from the input image is passed through a polarizing beam splitter, a lens, a binary diffraction grating, and another lens, to focus an array of *M* replicas of the input image on one face of a liquid-crystal-television spatial light modulator that is displaying the *M* remembered images. The position of each replica of the input image coincides with that of one of the remembered images. Light from the array of pairs of overlapping input and remembered images is focused by a corresponding array of lenslets onto a corresponding array of photodetectors. The intensity of light falling on each photodetector is proportional to the inner product between the input image and the corresponding remembered image.

The outputs of the photodetectors are processed through operational amplifiers that respond nonlinearly to inner-product level (in effect executing analog threshold functions). The outputs of the amplifiers drive point sources of white light, and an array of lenslets concentrates the light from each source onto the spot occupied by one of *M* remembered images displayed on another liquid-crystal-television spatial light modulator. The light that passes through this array is reflected by a pivoted ray of mirrors through a lens, which focuses the output image onto a CCD television camera. The output image consists of superpositioned remembered images, the brightest of which are those that represent the greatest inner products (the greatest resemblance to the input image). The television camera feeds the output image to a control computer, which performs a threshold computation, then feeds the images through a cathode-ray tube back to the input liquid-crystal light valve. This completes the associative recall loop. The loop operates iteratively until one (if any) of the remembered images is the sole output image.

SENSORS FOR HAND-EYE COORDINATION OF MICROROBOTIC MOTION UTILIZING VISION TECHNOLOGY

The micro motion of a robotic manipulator can be controlled with the help of dual feedback by a new method that reduces position errors by an order of magnitude. The errors—typically on the order of centimeters—are differences between real positions on the one hand and measured and computed positions on the other; these errors arise from several sources in the robotic actuators and sensors and in the kinematic model used in control computations. In comparison with current manufacturing methods of controlling the motion of a robot with visual feedback (the robotic equivalent of hand-eye coordination), the novel method requires neither calibration over the entire work space nor the use of an absolute reference coordinate frame for computing transformations between field of view and robot joint coordinates.

The robotic vision subsystem includes five cameras: three stationary ones that provide wide-angle views of the work space and two mounted on the wrist of an auxiliary robot arm to provide stereoscopic close-up views of the work space near the manipulator (Fig. 5.24). The vision subsystem is assumed to be able to recognize the objects to be avoided and manipulated and to generate data on the coordinates of the objects from sent positions in the field-of-view reference frame.

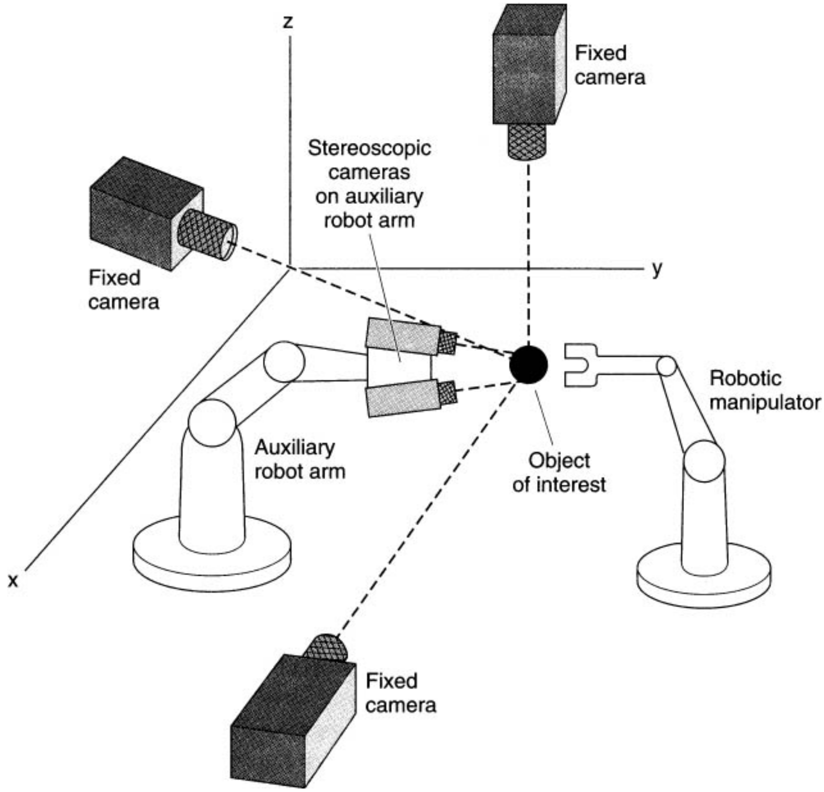


FIGURE 5.24 Stereoscopic cameras on an auxiliary robot arm.

The new method can be implemented in two steps:

1. The close-up stereoscopic cameras are set initially to view a small region that contains an object of interest. The end effector is commanded to move to a nominal position near the object and within the field of view. Typically, the manipulator stops at a slightly different position, which is measured by the cameras. Then, the measured error in position is used to compute a small corrective motion. This procedure is designed to exploit the fact that small errors in relative position can be measured accurately and small relative motions can be commanded accurately.
2. The approximate direct mapping between the visual coordinates and the manipulator joint-angle coordinates can be designed without intermediate transformation to and from absolute coordinates. This is, in effect, a calibration, but it requires fewer points than does a conventional calibration in an absolute reference frame over the entire work space. The calibration is performed by measuring the position of a target (in field-of-view coordinates) when the target is held rigidly by the manipulator at various commanded positions (in manipulator joint-angle coordinates) and when the cameras are placed at various commanded positions. Interpolations and extrapolations to positions near the calibration points are thereafter performed by use of the nonlinear kinematic transformations.

FORCE AND OPTICAL SENSORS CONTROLLING ROBOTIC GRIPPER FOR AGRICULTURE AND MANUFACTURING APPLICATIONS

A robotic gripper operates in several modes to locate, measure, recognize (in a primitive sense), and manipulate objects in an assembly subsystem of a robotic cell that is intended to handle geranium cuttings in a commercial greenhouse. The basic concept and design of the gripper could be modified for handling other objects—for example, rods or nuts—including sorting the objects according to size. The concept is also applicable to real-time measurement of the size of an expanding or contracting part gripped by a constant force and to measurement of the size of a compliant part as a function of the applied gripping force.

The gripper is mounted on an industrial robot. The robot positions the gripper at a fixed distance above the cutting to be processed. A vision system locates the cutting in the *x-y* plane lying on a conveyer belt (Fig. 5.25).

The robot uses fiber-optic sensors in the fingertip of the gripper to locate the cutting along the axis. The gripper grasps the cutting under closed-loop digital servo force control. The size (that is, the diameter of the stem) of the cutting is determined from the finger position feedback, while the cutting is being grasped under force control. The robot transports the

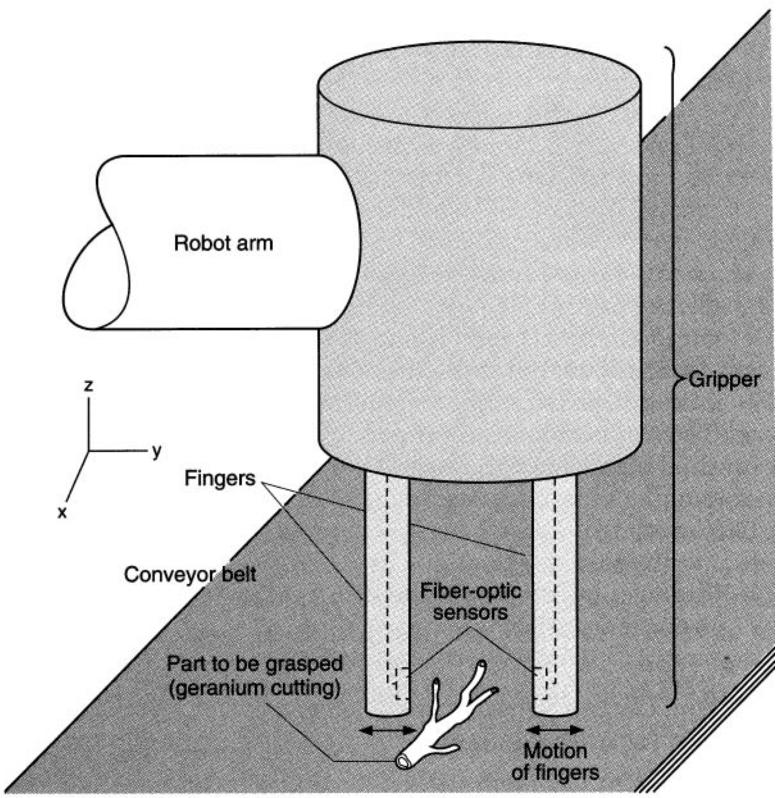


FIGURE 5.25 Robotic gripper for geranium cutting.

cutting to a scale for weighing, to a trimming station, and finally to a potting station. In this manner, cuttings are sorted according to weight, length, and diameter.

The control subsystem includes a 32-bit minicomputer that processes the vision information and collects system grating data. The minicomputer communicates with a robot controller and the gripper node control. The gripper node control communicates with the scale. The robot controller communicates with the gripper node controller via discrete input/output triggering of each of the gripper functions.

The gripper subsystem includes a PC/AT-compatible industrial computer; a gripper mechanism, actuated by two DC servomotors, with an integrated load cell; discrete input/output components; and two fiber-optic analog-output distance sensors. The computer includes a discrete input/output circuit card, an 8-channel A/D converter circuit card, a motor-control circuit card with A/D components, two serial ports, and a 286 processor with coprocessor.

A geranium cutting comprises a main stem and several petioles (tiny stem leaves). The individual outputs from the fiber-optic sensors can be processed into an indication of whether a stem or a petiole is coming into view as the gripper encounters the cutting. Consequently, the gripper can be commanded to grasp a stem but not a petiole. The axial centerline of a stem can also be recognized from the outputs of the five optic sensors. Upon recognition of a centerline, the gripper signals the robot, and the robot commands the gripper to close.

The motor-controller circuit card supplies the command signals to the amplifier that drives the gripper motors. This card can be operated as a position control with digital position feedback or as a force control with analog force feedback from the load cell mounted in the gripper. A microprocessor is located on the motor control card. Buffered command programs are downloaded from the computer to this card for independent execution by the card.

Prior to a controlled force closure, the motor-control card controls the gripper in position-servo mode until a specified force threshold is sensed, indicating contact with the cutting. Thereafter, the position-servo loop is opened, and the command signal to the amplifier is calculated as the difference between the force set point and the force feedback from the load cell. This distance is multiplied by a programmable gain value, then pulse-width-modulated with a programmable duty cycle of typically 200 percent. This technique provides integral stability to the force-control loop. The force-control loop is bidirectional in the sense that, if the cutting expands between the fingertips, the fingers are made to separate, and if the cutting contracts, the fingers are made to approach each other.

ULTRASONIC STRESS SENSOR MEASURING DYNAMIC CHANGES IN MATERIALS

An *ultrasonic dynamic vector stress sensor* (UDVSS) has recently been developed to measure the changes in dynamic directional stress that occur in materials or structures at the location touched by the device when the material or structure is subjected to cyclic load. A strain gauge device previously used for the measurement of such a stress measured strain in itself, not in the part being stressed, and thus provided a secondary measurement. Other techniques, such as those that involve thermoelasticity and shearography, have been expensive and placed demands on the measured material. The optical measurement of stress required the application of a phase coat to the object under test. The laser diffraction method required notching or sharp marking of the specimen.

A UDVSS is the first simple portable device able to determine stress directly in the specimen itself rather than in a bonded gauge attached to the specimen. As illustrated in Fig. 5.26, a typical material testing machine applies cyclic stress to a specimen. The UDVSS includes a probe, which is placed in contact with the specimen; an electronic system connected to the probe; and a source of a reference signal. The probe assembly includes a probe

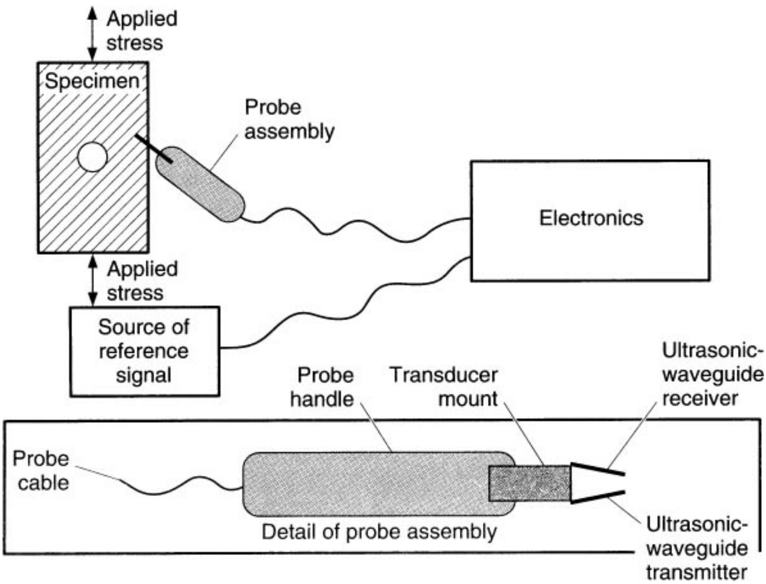


FIGURE 5.26 Ultrasonic dynamic stress sensor.

handle that holds the probe, a transducer mount that contains an active ultrasonic driver and receiver, and an ultrasonic waveguide transmitter and ultrasonic waveguide receiver that convert the electrical signals to mechanical motion and the inverse, and a cable that connects the probe of the electronics. When in contact with the specimen, the ultrasonic waveguide transmitter causes acoustic waves to travel across the specimen to the ultrasonic waveguide receiver, wherein the wave is converted to an electrical signal.

The operation of the UDVSS is based on the physical phenomenon that the propagation of sound in the specimen changes when the stress in the specimen changes. A pulse phase-locked loop reacts to a change in propagation of sound and therefore in stress by changing its operational frequency. The component of that signal represents that change in voltage needed to keep the system at quadrature to follow the system change in stress. That signal provides the information on changing stress.

The UDVSS can be moved around on the specimen to map out the stress field, and by rotating the probe, one can determine the direction of a stress. In addition, the probe is easily calibrated. The UDVSS should find wide acceptance among manufacturers of aerospace and automotive structures for stress testing and evaluation of designs.

PREDICTIVE MONITORING SENSORS SERVING CIM STRATEGY

Computer-integrated manufacturing technology can be well-served by a predictive monitoring system that would prevent a large number of sensors from overwhelming the electronic data monitoring system or a human operator. The essence of the method is to select only a few of the many sensors in the system for monitoring at a given time and to set alarm levels of the selected sensor outputs to reflect the limit of expected normal operation at the given time. The method is intended for use in a highly instrumented system that includes many interfacing components and subsystems—for example, an advanced aircraft, an environmental chamber, a chemical processing plant, or a machining work cell.

Several considerations motivate the expanding effort in implementing the concept of predictive monitoring. Typically, the timely detection of anomalous behavior of a system and the ability of the operator or electronic monitor to react quickly are necessary for the continuous safe operation of the system.

In the absence of a sensor-planning method, an operator may be overwhelmed with alarm data resulting from interactions among sensors rather than data directly resulting from anomalous behavior of the system. In addition, much raw sensor data presented to the operator may by irrelevant to an anomalous condition. The operator is thus presented with a great deal of unfocused sensor information, from which it may be impossible to form a global picture of events and conditions in the system. The predictive monitoring method would be implemented in a computer system running artificial intelligence software, tentatively named *PREMON*. The predictive monitoring system would include three modules: (1) a causal simulator, (2) a sensor planner, and (3) a sensor interpreter (Fig. 5.27).

The word *event* in Fig. 5.27 denotes a discontinuous change in the value of a given quantity (sensor output) at a given time. The inputs to the causal simulator would include a causal mathematical model of the system to be monitored, a set of events that describe the initial state of the system, and perhaps some future scheduled events. The outputs of the causal simulator would include a set of predicted events and a graph of causal dependency among events.

The sensor planner would use the causal dependency graph generated by the causal simulator to determine which few of all the predicted events are important enough to be verified. In many cases, the most important events would be taken to be those that either caused, or are caused by, the greatest number of other events. This notion of causal importance would serve as the basis for the election of those sensors, the outputs of which should be used to verify the expected behavior of the system.

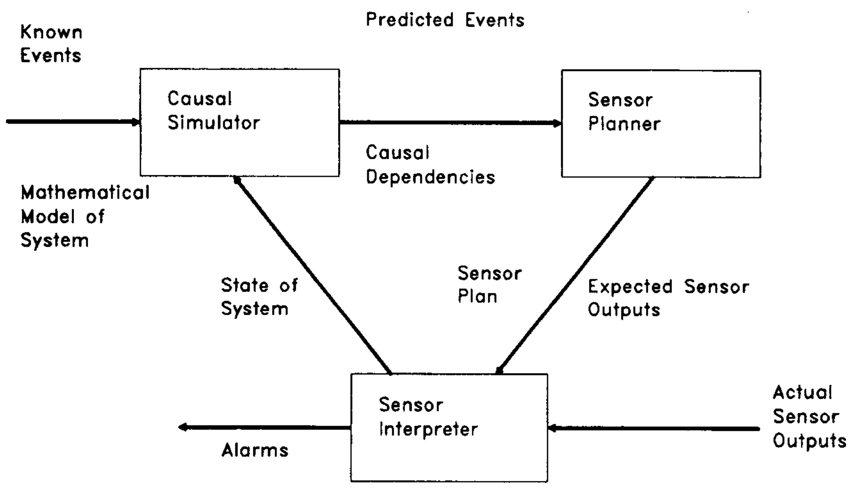


FIGURE 5.27 Predictive monitoring system would concentrate on the few sensor outputs that are most important at the moment.

The sensor interpreter would compare the actual outputs of the selected sensors with the values of those outputs predicted by the causal simulator. Alarms would be raised where the discrepancies between predicted and actual values were significant.

REFLECTIVE STRIP IMAGING CAMERA SENSOR—MEASURING A 180°-WIDE ANGLE

A proposed camera sensor would image a thin striplike portion of a field of view 180° wide. For example, it could be oriented to look at the horizon in an easterly direction from north to south or it could be rotated about a horizontal north/south axis to make a "pushbroom" scan of the entire sky proceeding from the easterly toward the westerly half of the horizon. Potential uses for the camera sensor include surveillance of clouds, coarse mapping of terrain, measurements of the bidirectional reflectance distribution functions of aerosols, imaging spectrometry, oceanography, and exploration of the planets.

The imaging optics would be a segment of concave hemispherical reflecting surfaces placed slightly off center (Fig. 5.28). Like other reflecting optics, it would be achromatic. The unique optical configuration would practically eliminate geometric distortion of the image. The optical structure could be fabricated and athermalized fairly easily in that it could be machined out of one or a few pieces of metal, and the spherical reflecting surface could be finished by diamond turning. In comparison, a camera sensor with a fish-eye lens, which provides a nearly hemispherical field of view, exhibits distortion, chromatism,

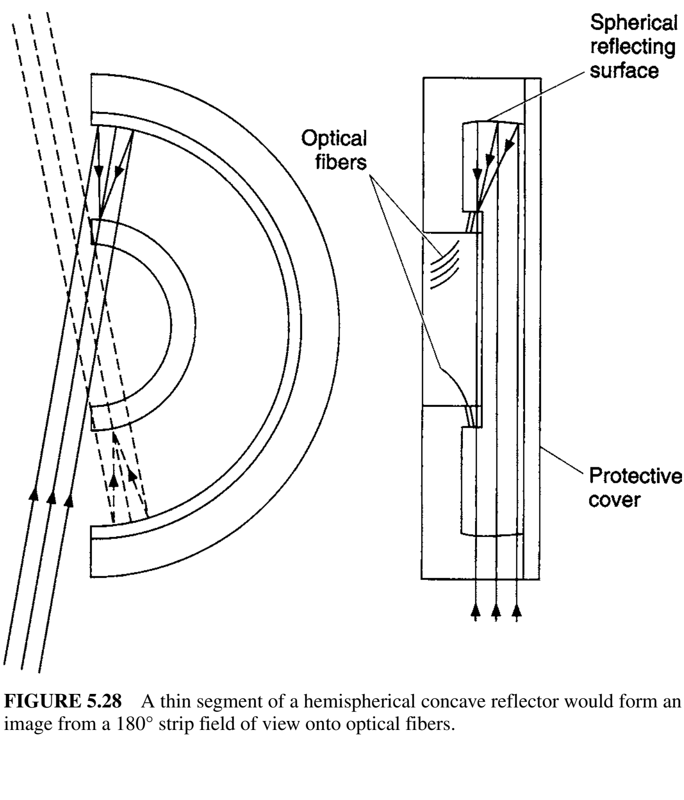


FIGURE 5.28 A thin segment of a hemispherical concave reflector would form an image from a 180° strip field of view onto optical fibers.

and poor athermalization. The image would be formed on a thin semicircular strip at half the radius of the sphere. A coherent bundle of optical fibers would collect the light from this strip and transfer the image to a linear or rectangular array of photodetectors or to the entrance slit of an image spectrograph. Provided that the input ends of the fibers were properly aimed, the cones of acceptance of the fibers would act as aperture stops; typically, the resulting width of the effective aperture of the camera sensor would be about one-third the focal length (*f*/3).

The camera sensor would operate at wavelengths from 500 to 1100 nm. The angular resolution would be about 0.5°. In the case of an effective aperture of *f*/3, the camera would provide an unvignetted view over the middle 161° of the strip, with up to 50 percent vignetting in the outermost 9.5° on each end.

The decentration of the spherical reflecting surface is necessary to make room for the optical fibers and the structure that would support them. On the other hand, the decentration distance must not exceed the amount beyond which the coma that results from decentration would become unacceptably large. In the case of an effective aperture of *f*/3, the coma would be only slightly in excess of the spherical aberration if the decentration were limited to about *f*/6. This would be enough to accommodate the fibers and supporting structure.

OPTICAL SENSOR QUANTIFYING ACIDITY OF SOLUTIONS

With environmental concerns increasing, a method for taking effective measurements of acidity will minimize waste and reduce both the cost and the environmental impact of processing chemicals. Scientists at Los Alamos National Laboratory (LANL) have developed an optical sensor that measures acid concentration at a higher level than does any current device. The optical high-acidity sensor, reduces the wear generated from acidity measurements and makes possible the economic recycling of acid waste.

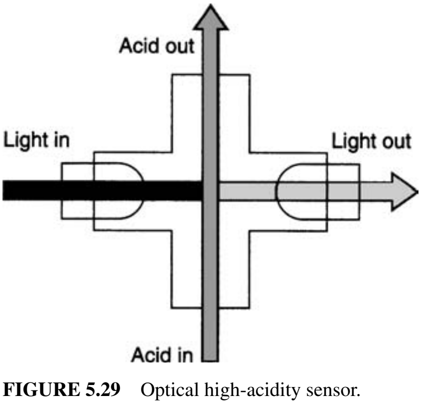


FIGURE 5.29 Optical high-acidity sensor.

The high-acidity sensor (Fig. 5.29) consists of a flow cell (about 87.5 mm across) in which two fused silica lenses are tightly mounted across from one another. Fiber-optic cables connect the two lenses to a spectrophotometer. One lens is coated with a sensing material consisting of a polymer that is chemically bound to the lens and an indicator that is physically entrapped with the polymer. Acidic solutions flow vertically up through the flow cell. The light from a fixed source in the spectrophotometer is collimated by the coated lens, passes through the acidic solution, and then is focused by the second lens.

The amount of light absorbed by the indicator depends on the acidity of the solution. The absorption spectrum of the indicator reflects the concentration of hydrogen ions in the solution, which is how acidity is measured.

The LANL sensor's sensitivity range—acid concentrations of 4 to 12 molar—is unmatched by any other acidity measurement technique in current manufacturing. Present techniques are unsuitable for measuring solutions with a negative pH or for measuring chemical processes. With the LANL sensor, high acid concentrations can be measured in 3 min or less, which is 50 times faster than measuring with manual sampling and titration. The sensing material can function in highly acidic solutions for four to six months with calibration less than once a week, and needs to be replaced only once or twice a year. A sensor using similar principles has recently been developed from measuring lower acidity concentrations.

The sensor is selective in its measurements because the small pores of the polymer allow only hydrogen ions to pass through the indicator. Metal ions do not form a complex with the indicator under acidic solutions. The sensor is reusable—that is, chemically reversible. In short, no comparable product for measuring high acidity is reversible or as sensitive, fast, accurate, and selective as the LANL optical high-acidity sensor.

The prime use for the sensor is monitoring highly acidic chemical processes and waste solutions. High-acidity processes are common in preparing metals from ore-mining operations, treating fuel elements from nuclear power plants, manufacturing bulk acid, and metal finishing including passivation and electroplating.

For highly acidic processor applications, the LANL sensor will save companies thousands of dollars by improving efficiency and decreasing the time devoted to acid measurements. The sensor can be used on manufacturing lines, allowing control and waste management adjustment to be made before acidity fluctuations become a problem. The sensor will improve efficiency at least 25 percent by eliminating the need to reprocess material processed incorrectly on the first try because its true acidity was not known and controlled. Higher efficiency will mean lower cost and minimal waste products generated; the sensor itself generates no caustic waste. Finally, the sensor's waste monitoring capabilities will help ensure that any discharged waste is environmentally benign.

For the 500 acidity measurements done at LANL in 1991, the sensor saved $99,500, a 99.5 percent savings in labor costs in the first year. And, because the sensor generates no waste, 20 L a year of caustic waste was avoided, a 100 percent reduction.

SENSORS FOR BIOMEDICAL TECHNOLOGY

In recent years, advanced imaging and other computer-related technology have greatly expanded the horizons of basic biological and biochemical research. Currently, such drivers as the growing needs for environmental information and increased understanding of genetic systems have provided impetus to biotechnology development. This is a relatively new specialty in the marketplace; nevertheless, the intensity of worldwide competition is escalating. Collaborative research and development projects within the U.S. government, industry, and academia constitute a major thrust for rapid deployment of research and development. The results can place the nation in a position of world leadership in biotechnology.

Sensor for Detecting Minute Quantities of Biological Materials

A new device based on laser-excited fluorescence provides unparalleled detection of biological materials for which only minuscule samples may be available. This device, invented at the Ames Laboratories, received a 1991 R&D 100 award.

The Ames Microfluor detector was developed to meet a need for an improved detection technique, driven by important new studies of the human genome, abuse substances, toxins, DNA adduct formation, and amino acids, all of which may be available only in minute amounts. Although powerful and efficient methods have been developed for separating biological mixtures in small volume (i.e., capillary electrophoresis), equally powerful techniques for subsequent detection and identification of these mixtures have been lacking.

The Microfluor detector combines very high sensitivities with the ability to analyze very small volumes. The instrument design is based on the principle that many important biomaterials are fluorescent, while many other biomaterials, such as peptides and oligonucleotides, can be made to fluoresce by adding a fluorescent tag.

When a sample-filled capillary tube is inserted into the Microfluor detector and is irradiated by a laser beam, the sample will fluoresce. The detector detects, monitors, and quantifies the contents by sensing the intensity of the fluorescent light emitted. The signal is proportional to the concentration of the materials. The proportionality constant is characteristic of the material itself.

Analyses can be performed with sample sizes 50 times smaller than those required by other methods, and concentrations as low as 10^{-11} molar (1 part per trillion) can be measured. Often, the critical components in a sample are present at these minute concentrations. These two features make the Microfluor detector uniquely compatible with capillary electrophoresis. In addition, the Ames-developed detector is distinct from other laser-excited detectors in that it is not seriously affected by stray light from the laser itself; it also allows simple alignment and operation in full room light. The Microfluor detector has already been used to determine the extent of adduct formation and base modification in DNA so that the effects of carcinogens on living cells can be studied. Future uses of the sensor will include DNA sequencing and protein sequencing. With direct or indirect fluorescence detection, researchers are using this technique to study chemical contents of individual living cells. This capability may allow pharmaceutical products to be tested on single cells rather than on whole organisms, with improved speed and safety.

Sensors for Early Detection and Treatment of Lung Tumors

A quick accurate method for early detection of lung cancer would raise chances of patient survival from less than 50 percent to 80 or 90 percent. Until now, small cancer cells deep in the lung have been impossible to detect before they form tumors large enough to show up in x-rays. Researchers at Los Alamos National Laboratory (LANL) in collaboration with other institutions, including the Johns Hopkins school of medicine and St. Mary's Hospital in Grand Junction, Colorado, are developing methods for finding and treating lung cancer in its earliest stages.

A detection sensor involves a porphyrin, one of an unusual family of chemicals found naturally in the body that concentrates in cancer cells. The chemical is added to a sample of sputum coughed up from the lung. When exposed to ultraviolet or laser light, cells in the porphyrin-treated sputum glow a bright red. When the sample is viewed under the microscope, the amount and intensity of fluorescence in the cells determines the presence of cancer.

The first clinical test of a detection technique using porphyrin was done by LANL and St. Mary's Hospital in 1988. Four different porphyrins were tested on sputum samples from two former miners, one known to have lung cancer and one with no detectable cancer. One of the porphyrins was concentrated in certain cells only in the sputum of the miner with lung cancer. Other tests concluded that these were cancer cells. Later, a blind study of sputum samples from 12 patients, eight of whom had lung cancer in various stages of development, identified all the cancer patients as well as a ninth originally thought to be free of cancer. Further tests showed that this patient also had lung cancer.

Identifying the ninth patient prompted a new study, in which the procedure was evaluated for its ability to detect precancerous cells in the lung. In this study, historical sputum samples obtained from Johns Hopkins were treated with the porphyrin. Precancerous conditions that chest x-rays had not detected were identified in samples from patients who later developed the disease. Although further testing is needed, researchers now feel confident that the technique and the detection sensor can detect precancerous conditions three to four years before onset of the disease.

Working in collaboration with industry, researchers expect to develop an instrument in several years for rapid automated computerized screening of larger populations of smokers, miners, and other people at risk of developing lung cancer. Because lung cancer is the leading cancer killer in both men and women, a successful screening procedure would dramatically lower the nation's mortality rate from the disease. A similar procedure has potential for the analysis of Pap smears, now done by technicians who screen slides one at a time and can misread them.

In addition to a successful screening program, LANL hopes to develop an effective treatment for early lung cancer. The National Cancer Institute will help investigate and implement both diagnostic and therapeutic programs. Researchers are performing basic cell studies on animals to investigate the effect of porphyrin on lung cancer cells. They are also exploring the use of the LANL-designed porphyrin attached to radioactive copper to kill early cancer cells in the lung in a single search-and-destroy mission. The porphyrin not only seeks out cancer cells but also has a molecular structure, similar to a hollow golf ball, that can take certain metals into its center. A small amount of radioactive copper 67 placed inside the porphyrin should destroy tumors the size of a pinhead, as well as function as a tracer.

Ultrasensitive Sensor for Single-Molecule Detection

Enhancing the sensitivity of research instruments has long been a goal of physicists, biologists, and chemists. With the single-molecule detector, researchers have achieved the ultimate in this pursuit: detection of a single fluorescent molecule in a liquid. The new instrument—a thousand times more sensitive than existing commercial detectors—brings new capabilities to areas of science and technology that affect lives in many ways, from DNA sequencing, biochemical analysis, and virus studies to identifying environmental pollutants.

For some time, scientists have observed individual molecules trapped in a vacuum, where they are isolated and relatively easy to find. However, many important biological and chemical processes occur in a liquid environment, where many billions of other molecules surround the molecules of interest. Developing a single-molecule detector that operates in such an environment presented a difficult challenge.

The observation technique involves attaching fluorescent dye molecules to the molecules of interest and then exciting the dye molecule by passing them in solution to a rapidly pulsed laser beam and detecting the subsequent faint light, or photons, they emit. The fluorescent lifetimes of the molecules are much shorter than the time the molecules spend in the laser beam; therefore, each molecule is reexcited many times and yields many fluorescent photons. The signature of the passing molecule is the burst of photons that occurs when the molecule is passing the laser beam (Fig. 5.30).

A lens, or microscope objective, and a slit are arranged to image the photons from a small region around the laser beam waist onto a microchannel plate photomultiplier (MCPP) that counts individual photons. The intense excitation light from the laser is blocked from reaching the MCPP by a bandpass spectral filter, which is centered near the peak fluorescent wavelength. The excitation light consists of exceptionally short pulses, each pulse is about 70 trillionths of a second. The dye molecule does not emit light until a billionth of a second after excitation, so the flash of laser light fades before the feeble molecular glow occurs. For reliable identification of individual molecules, the technique maximizes the number of the detected photons and minimizes the number of background photons. Although some background remains, the technique registers over 85 percent of the fluorescent molecules.

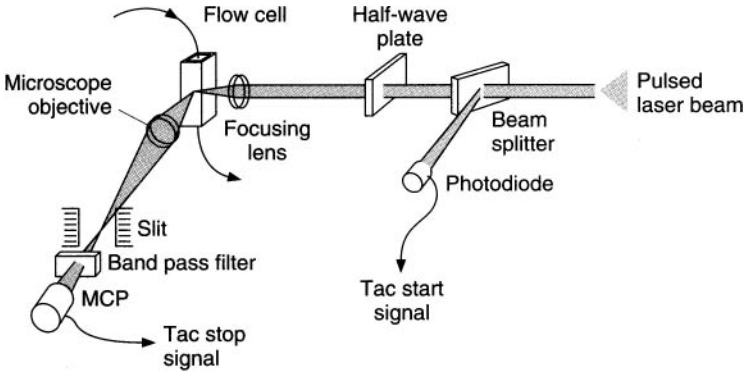


FIGURE 5.30 Single-molecule detector for individual fluorescent molecules in solution.

Developed to aid in sequencing chromosomes for the Human Genome Project, the sensor detector's high activity would allow DNA sequencing rates hundreds of times faster than those obtainable with present techniques. It would also eliminate the need for radioactive materials, gels, and electrophoresis solutions, which often create disposal problems, and it is expected to help make DNA sequencing a routine diagnostic and clinical tool. One eventual benefit of the DNA research may be rapid screening for any of 3500 known genetic diseases such as diabetes, cystic fibrosis, and Alzheimer's disease.

The ultrasensitive detector can be used to find and quantify minute amounts of chemicals, enzymes, and viruses in the blood and monitor the dispersal of extremely low concentrations of environmental pollutants. The device may also be useful in studying the interaction of viruses and their binding sites. It may be possible to develop a procedure for rapidly evaluating the efficiency of vaccines; such a procedure could quickly expedite the search for an effective vaccine against the AIDS virus.

REFERENCES

1. Barks, R. E., "Optical High Acidity Sensor," Los Alamos National Laboratory (June) (1991).
2. Bicknel, T. J., and W. H. Farr, "Acousto Optical/Electronic Processor for SAR," *NASA Tech Briefs*, 16 (May) (1992).
3. Bonnert, R., "Design of High Performance Digital Tachometer with Digital Microcontroller," *IEEE Trans. Instrum.*, 38, 1104–1108 (1989).
4. Buser, R. A., and N. F. Rooij, "Resonant Silicon Structures," *Sensors and Actuators*, 17, 145–154 (1989).
5. D'Amico A., and E. Verona, "SAW Sensors," *Sensors and Actuators*, 17, 66–66 (1989).
6. Fleming Dias, J., "Physics Sensors Using SAW Devices," *Hewlett-Packard J.* (December) 18–20 (1981).
7. Gast, T., "Sensors with Oscillating Elements," *J. Phys. E. Sci. Instrum.*, 18, 783–789 (1985).
8. Hayman, J. S., and M. Frogatt, "Ultrasonic Dynamic Vector Stress Sensor (UDVSS)," Langley Research Center, 1992.
9. Hewlett-Packard, "Design and Operational Considerations for the HELDS-5000 Incremental Shaft Encoder," *Application Note 1011,* Palo Alto, Ca., 1981.

10. Higbie, B. N., “Automatic Detection of Faults in Turbomachinery Bearings,” *NASA Tech Briefs*, 16 (May) (1992).
11. Higbie, N., *Technical Report,* Technology Integration and Development Group Inc. (May) (1992).
12. Huner, B., P. Klinkhachorn, and E. B. Everton, “Hybrid Clock Oscillator Modules as Deposition Monitors,” *Rev. Sci. Instrum.*, 59, 983–986 (1988).
13. Ito, H., “Balanced Absorption Quartz Hygrometer,” *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, 34, 136–141 (1987).
14. Lokshin, A. M., “Hand/Eye Coordination for Microrobotic Motion—Utilizing Vision Technology,” Caltech, 1991.
15. Los Alamos National Laboratory, Research Team, “Single Molecule Detection,” (January) (1991).
16. Montgomery, J. L., “Force and Optical Sensors Controlling Robotic Gripper for Agriculture,” Martin Marietta Corp., 1992.
17. Noble, M. N., D. N. Mark, and R. Blue, “Tracking Retroreflective Targets on a Structure,” *NASA Tech Briefs*, 16 (May) (1992).
18. “Optical Sensor to Quantify Highly Acidic Solution—Sensing High Acidity without Generating Waste,” *Technology ’91*, LANL, 57–58 (1991).
19. Tien-Hsin Chao, “Experimental Optoelectronic Associative Memory,” *NASA Tech Briefs*, 16 (May) (1992).
20. “Ultrasensitive Detection for Medical and Environmental Analysis—Single Molecule Detector,” *Technology ’91*, LANL, 80–81 (1991).
21. Vaughan, A. H., “Reflective Strip Imaging Camera Sensor—Measuring a 180° Wide Angle,” Caltech, NASA’s Jet Propulsion Laboratory, 1992.
22. Williams, D. E., “Laser Sensor Detecting Microfluor,” Ames Laboratory, July 1991.

This page intentionally left blank

CHAPTER 6
INDUSTRIAL SENSORS AND CONTROL

INTRODUCTION

Current manufacturing strategy defines manufacturing systems in terms of sensors, actuators, effectors, controllers, and control loops. Sensors provide a means for gathering information on manufacturing operations and processes being performed. In many instances, sensors are used to transform a physical stimulus into an electrical signal that may be analyzed by the manufacturing system and used for making decisions about the operations being conducted. Actuators convert an electrical signal into a mechanical motion. An actuator acts on the product and equipment through an effector. Effectors serve as the “hand” that achieves the desired mechanical action. Controllers are computers of some type that receive information from sensors and from internal programming, and use this information to operate the manufacturing equipment (to the extent available, depending on the degree of automation and control). Controllers provide electronic commands that convert an electrical signal into a mechanical action. Sensors, actuators, effectors, and controllers are linked to a control loop.

In limited-capability control loops, little information is gathered, little decision making can take place, and limited action results. In other settings, “smart” manufacturing equipment with a wide range of sensor types can apply numerous actuators and effectors to achieve a wide range of automated actions.

The purpose of sensors is to inspect work in progress, to monitor the work-in-progress interface with the manufacturing equipment, and to allow self-monitoring of manufacturing by the manufacturing system’s own computer. The purpose of the actuator and effector is to transform the work in progress according to the defined processes of the manufacturing system. The function of the controller is to allow for varying degrees of manual, semiautomated, or fully automated control over the processes. In a fully automated case, such as in computer-integrated manufacturing, the controller is completely adaptive and functions in a closed-loop manner to produce automatic system operation. In other cases, human activity is involved in the control loop.

In order to understand the ways in which the physical properties of a manufacturing system affect the functional parameters associated with the manufacturing system, and so as to determine the types of physical manufacturing system properties necessary to implement the various desired functional parameters, it is necessary to understand the technologies available for manufacturing systems that use automation and integration to varying degrees.

The least automated equipment makes use of detailed operator control over all equipment functions. Further, each action performed by the equipment is individually directed by the operator. Manual equipment thus makes the maximum use of human capability and adaptability. Visual observations can be enhanced by the use of microscopes and cameras, and the actions undertaken can be improved through the use of simple effectors. The linkages between the sensory information (from microscopes or through cameras) and the resulting actions are obtained by placing the operator in the loop.

This type of system is clearly limited by the kinds of sensors used and their relationship to the human operator, the types of effectors that can be employed in conjunction with the human operator, and the capabilities of the operator. The manufacturing equipment that is designed for a manual strategy must be matched to human capabilities. The human-manufacturing equipment interface is extremely important in many manufacturing applications. Unfortunately, equipment design is often not optimized as a sensor-operator-actuator/effector control loop.

A manufacturing system may be semiautomated, with some portion of the control loop replaced by a computer. This approach will serve the new demands on manufacturing system design requirements. Specifically, sensors now must provide continuous input data for both the operator and computer. The appropriate types of data must be provided in a timely manner to each of these control loops. Semiautomated manufacturing systems must have the capability for a limited degree of self-monitoring and control associated with the computer portion of the decision-making loop. An obvious difficulty in designing such equipment is to manage the computer- and operator-controlled activities in an optimum manner. The computer must be able to recognize when it needs operator support, and the operator must be able to recognize which functions may appropriately be left to computer control. A continuing machine-operator interaction is part of normal operations.

Another manufacturing concept involves fully automated manufacturing systems. The processing within the manufacturing system itself is fully computer-controlled. Closed-loop operations must exist between sensors and actuators/effectors in the manufacturing system. The manufacturing system must be able to monitor its own performance and decision making for all required operations. For effective automated operation, the mean time between operator interventions must be large when compared with the times between manufacturing setups.

$$\text{MTOI} = \left(\sum_{i}^{n} \tau_1 + \tau_2 + \tau_3 + \tau_4 + \cdots + \tau_n\right) \Big/ n \tag{6.1}$$

where τ = setup time
i = initial setup
n = number of setups

The processes in use must rarely fail; the operator will intervene only when such failures occur. In such a setting, the operator's function is to ensure the adequate flow of work in progress and respond to system failure.

Several types of work cells are designed according to the concept of total manufacturing integration. The most sophisticated cell design involves fully automated processing and materials handling. Computers control the feeding of work in progress, the performance of the manufacturing process, and the removal of the work in progress. Manufacturing systems of this type provide the opportunity for the most advanced automated and integrated operations. The manufacturing system must be modified to achieve closed-loop operations for all of these functions.

Most manufacturing systems in use today are not very resourceful. They do not make use of external sensors that enable them to monitor their own performance. Rather, they depend on internal conditioning sensors to feed back (to the control system) information regarding manipulator positions and actions. To be effective, this type of manufacturing system must have a rigid structure and be able to determine its own position based on internal data (largely independent of the load that is applied). This leads to large, heavy, and rigid structures.

The more intelligent manufacturing systems use sensors that enable them to observe work in progress and a control loop that allows corrective action to be taken. Thus, such manufacturing systems do not have to be as rigid because they can adapt.

The evolution toward more intelligent and adaptive manufacturing systems has been slow, partly because the required technologies have evolved only in recent years and partly because it is difficult to design work cells that effectively use the adaptive capabilities. Enterprises are not sure whether such features are cost-effective and wonder how to integrate smart manufacturing systems into the overall strategy.

The emphasis must be on the building-block elements necessary for many types of processing. If the most advanced sensors are combined with the most advanced manufacturing systems, concepts, and state-of-the-art controllers and control loops, very sophisticated manufacturing systems can result. On the other hand, much more rudimentary sensors, effectors, and controllers can produce simple types of actions.

In many instances today, sensors are analog (they involve a continuously changing output property), and control loops make use of digital computers. Therefore, an analog-to-digital converter between the preprocessor and the digital control loop is often required.

The sensor may operate either passively or actively. In the passive case, the physical stimulus is available in the environment and does not have to be provided. For an active case, the particular physical stimulus must be provided. Machine vision and color identification sensors are an active means of sensing, because visible light must be used to illuminate the object before a physical stimulus can be received by the sensor. Laser sensors are also active-type sensors. Passive sensors include infrared devices (the physical stimulus being generated from infrared radiation that is associated with the temperature of a body) and sensors to measure pressure, flow, temperature, displacement, proximity, humidity, and other physical parameters.

SENSORS IN MANUFACTURING

Many types of sensors have been developed during the past several years, especially those for industrial process control, military uses, medicine, automotive applications, and avionics. Several types of sensors are already being manufactured by commercial companies.

Process control sensors in manufacturing will play a significant role in improving productivity, qualitatively and quantitatively, throughout the coming decades. The main parameters to be measured and controlled in industrial plants are temperature, displacement, force, pressure, fluid level, and flow. In addition, detectors for leakage of explosives or combustible gases and oils are important for accident prevention.

Optical-fiber sensors may be conveniently divided into two groups: (1) intrinsic sensors and (2) extrinsic sensors.

Although intrinsic sensors have, in many cases, an advantage of higher sensitivity, almost all sensors used in process control at present belong to the extrinsic type. Extrinsic-type sensors employ light sources such as LEDs, which have higher reliability, longer life, and lower cost than semiconductor lasers. They also are compatible with multimode fibers, which provide higher efficiency when coupled to light sources and are less sensitive to external mechanical and thermal disturbances.

As described in Chap. 1, objects can be detected by interrupting the sensor beam. Optical-fiber interrupters are sensors for which the principal function is the detection of moving objects. They may be classified into two types: reflection and transmission.

In the reflection-type sensor, the light beam emitted from the fiber is reflected back into the same fiber if the object is situated in front of the sensor.

In the transmission-type sensor, the emitted light from the input fiber is interrupted by the object, resulting in no received light in the output fiber located at the opposite side. Typical obstacle interrupters employ low-cost large-core plastic fibers because of the short transmission distance. The minimum detectable size of the object is typically limited to 1 mm by the fiber core diameter and the optical beam. The operating temperature range of commercially available sensors is typically –40 to +70°C. Optical-fiber sensors have been utilized in industry in many ways, such as:

- Detection of lot number and expiration dates (for example, in the pharmaceutical and food industries)
- Color difference recognition (for example, colored objects on a conveyer)
- Defect detection (for example, missing wire leads in electronic components)
- Counting discrete components (for example, bottles or cans)
- Detecting the absence or presence of labels (for example, packaging in the pharmaceutical and food industries)

Fiber-optic sensors for monitoring process variables such as temperature, pressure, flow, and liquid level are also classified into two types: (1) the normally OFF type in which the shutter is inserted between the fibers in the unactivated state—thus, this type of sensor provides high and low levels as the light output corresponds to ON and OFF states, respectively; and (2) the normally ON type, where the shutter is retracted from the gap in the unactivated state.

In both types, the shutter is adjusted so it does not intercept the light beam completely but allows a small amount of light to be transmitted, even when fully closed. This transmitted light is used to monitor the cable (fiber) continuity for faults and provides an intermediate state. Commercially available sensors employ fibers of 200-μm core diameter. The typical differential attenuation that determines the ON-OFF contrast ratio is about 20 dB. According to manufacturers' specifications, these sensors operate well over the temperature range –40 to +80°C with a 2-dB variation in light output.

TEMPERATURE SENSORS IN PROCESS CONTROL

Temperature is one of the most important parameters to be controlled in almost all industrial plants since it directly affects material properties and thus product quality. During the past few years, several temperature sensors have been developed for use in electrically or chemically hostile environments. Among these, the practical temperature sensors, which are now commercially available, are classified into two groups: (1) low-temperature sensors with a range of –100 to +400°C, using specific sensing materials such as phosphors, semiconductors, and liquid crystals; and (2) high-temperature sensors with a range of 500 to 2000°C based on blackbody radiation.

Semiconductor Absorption Sensors

Many of these sensors can be located up to 1500 m away from the optoelectronic instruments. The operation of semiconductor temperature sensors is based on the temperature-dependent absorption of semiconductor materials. Because the energy and gap of most semiconductors decrease almost linearly with increasing temperature T, the band-edge wavelength $\lambda_g(T)$ corresponding to the fundamental optical absorption shifts toward longer wavelengths at a rate of about 3 Å/°C [for gallium arsenide (GaAs)] with T. As illustrated in Fig. 6.1, when a light-emitting diode with a radiation spectrum covering the wavelength $\lambda_g(T)$ is used as a light source, the light intensity transmitted through a semiconductor decreases with T.

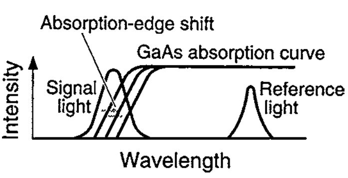


FIGURE 6.1 Operating principle of optical-fiber thermometer based on temperature-dependent GaAs light absorption.

Figure 6.2 shows the reflection-type sensing element. A polished thin GaAs chip is attached to the fiber end and mounted in a stainless-steel capillary tube of 2-mm diameter. The front face of the GaAs is antireflection-coated, while the back face is gold-coated to return the light into the fiber.

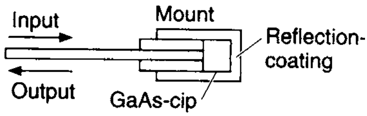


FIGURE 6.2 Sensing element of the optical-fiber thermometer with GaAs light absorber.

The system configuration of the thermometer is illustrated in Fig. 6.3. In order to reduce the measuring errors caused by variations in parasitic losses, such as optical fiber loss and connector loss, this thermosensor employs two LED sources [one aluminum gallium arsenide (AlGaAs), the other indium gallium arsenide (InGaAs)] with different wavelengths. A pair of optical pulses with different wavelengths $\lambda_s = 0.88$ μm and $\lambda_r = 1.3$ μm are guided from the AlGaAs LED and the InGaAs LED to the sensing element along the fiber. The light of λ_s is intensity-modulated by temperature. On the other hand, GaAs is transparent for the light of λ_r, which is then utilized as a reference light. After detection by a germanium avalanche photodiode (GeAPD), the temperature-dependent signal λ_s is normalized by the reference signal λ_r in a microprocessor.

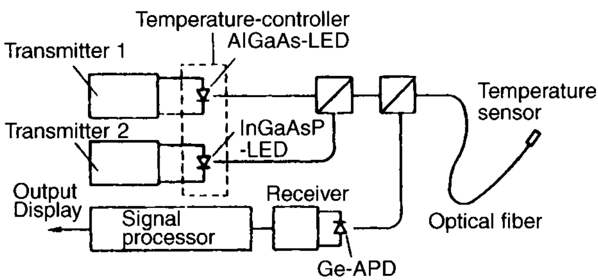


FIGURE 6.3 System configuration of the optical-fiber thermometer with GaAs light absorber.

The performance of the thermometer is summarized in Table 6.1. An accuracy of better than ±2°C is obtained within a range of −2° to +150°C. The principle of operation for this temperature sensor is based on the temperature-dependent direct fluorescent emission from phosphors.

TABLE 6.1 Characteristics of Semiconductor Sensors (Thermometer Performance)

Property	Semiconductor absorption sensor (Mitsubishi, Japan)	Semiconductor photoluminescence sensor (ASEA Innovation, Sweden)	Phosphor A sensor (Luxton, U.S.)	Phosphor B sensor (Luxton, U.S.)
Range, –C	−20 to +150	0 to 200	20 to 240	20 to 400
Accuracy, –C	±2.0	±1.0	±2.0	±2.0
Diameter, m	2	From 0.6	0.7	1.6
Time constant, s	0.5	From 0.3	From 0.25	
Fiber type	100-μm silica core	100-μm silica core	400-μm polymer clad	Fiber silica
Fiber length, m	300	500	100	300
Light source	AlGaAs LED	AlGaAs LED	Halogen lamp	Halogen lamp

Semiconductor Temperature Detector Using Photoluminescence

The sensing element of this semiconductor photoluminescence sensor is a double-heterostructure GaAs epitaxial layer surrounded by two $Al_xGa_{1-x}As$ layers. When the GaAs absorbs the incoming exciting light, the electron-hole pairs are generated in the GaAs layer. The electron-hole pairs combine and reemit the photons with a wavelength determined by temperature. As illustrated in Fig. 6.4, the luminescent wavelength shifts monotonically toward longer wavelengths as the temperature T increases. This is a result of the decrease in the energy gap E_g with T. Therefore, analysis of the luminescent spectrum yields the required temperature information. The double heterostructure of the sensing element provides excellent quantum efficiency for the luminescence because the generated electron-hole pairs are confined between the two potential barriers (Fig. 6.5).

The system is configured as shown in Fig. 6.6. The sensing element is attached to the end of the silica fiber (100-μm core diameter). The excitation light from an LED, with a peak wavelength of about 750 nm, is coupled into the fiber and guided to a special GRIN lens mounted to a block of glass. A first optical inference filter IF_1, located between the

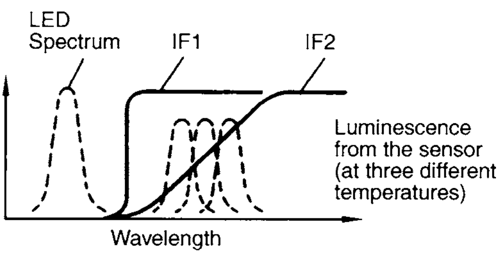


FIGURE 6.4 Operating principle of optical-fiber thermometer based on temperature-dependent photoluminescence from a GaAs epitaxial film.

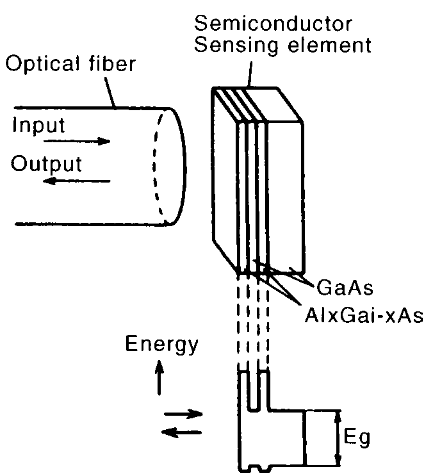


FIGURE 6.5 Sensing element of optical-fiber thermometer based on temperature-dependent photoluminescence.

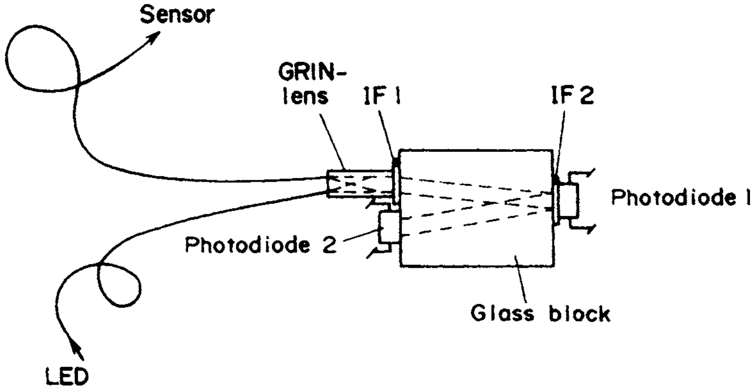


FIGURE 6.6 Optical system of optical-fiber thermometer based on temperature-dependent photoluminescence.

GRIN lens and the glass block, reflects the excitation light, which is guided to the sensing element along the fiber. However, this optical filter is transparent to the returned photoluminescent light. The reflectivity of the second interference filter IF_2 changes at about 900 nm. Because the peak wavelength of the luminescence shifts toward longer wavelength with temperature, the ratio between the transmitted and the reflected light intensifies if IF_2 changes. However, the ratio is independent of any variation in the excitation light intensity and parasitic losses. The two lights separated by IF_2 are detected by photodiodes 1 and 2. The detector module is kept at a constant temperature in order to eliminate any influence of the thermal drift of IF_2.

The measuring temperature range is 0 to 200°C, and the accuracy is ±1°C. According to the manufacturer's report, good long-term stability, with a temperature drift of less than 1°C over a period of nine months, has been obtained.

Temperature Detector Using Point-Contact Sensors in Process Manufacturing Plant

Electrical sensors are sensitive to microwave radiation and corrosion. The needs for contact-type temperature sensors have lead to the development of point-contact sensors that are immune to microwave radiation, for use in: (1) electric power plants using transformers, generators, surge arresters, cables, and bus bars; (2) industrial plants utilizing microwave processes; and (3) chemical plants utilizing electrolytic processes.

The uses of microwaves include drying powder and wood; curing glues, resins, and plastics; heating processes for food, rubber, and oil; device fabrication in semiconductor manufacturing; and joint welding of plastic packages, for example.

Semiconductor device fabrication is currently receiving strong attention. Most semiconductor device fabrication processes are now performed in vacuum chambers, and include plasma etching and stripping, ion implantation, plasma-assisted chemical vapor deposition, radio-frequency sputtering, and microwave-induced photoresist baking. These processes alter the temperature of the semiconductors being processed. However, the monitoring and controlling of temperature in such hostile environments is difficult with conventional electrical temperature sensors. These problems can be overcome by the contact-type optical-fiber thermometer.

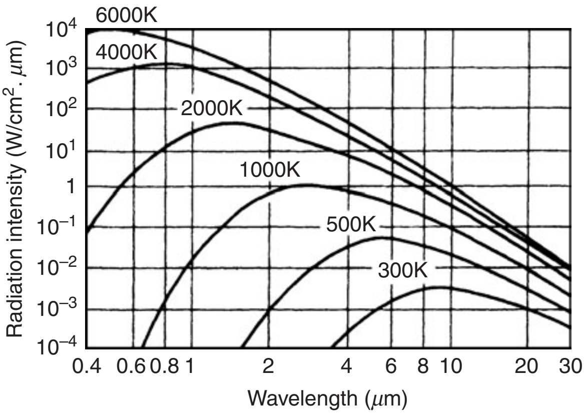


FIGURE 6.7 Spectral distribution of blackbody radiation.

Noncontact Sensors—Pyrometers

Because they are noncontact sensors, pyrometers do not affect the temperature of the object they are measuring. The operation of the pyrometer is based on the spectral distribution of blackbody radiation, which is illustrated in Fig. 6.7 for several different temperatures. According to the Stefan-Boltzmann law, the rate of the total radiated energy from a blackbody is proportional to the fourth power of absolute temperature and is expressed as:

$$W_t = \sigma T^4 \; W_t = \sigma T^4 \tag{6.2}$$

where σ is the Stefan-Boltzmann constant and has the value of $5.6697 \times 10^{-8}\,\mathrm{W/m^2 \cdot K^4}$

The wavelength at which the radiated energy has its highest value is given by Wien's displacement law,

$$\lambda_m T = 2.8978 \times 10^{-3}\,\mathrm{m \cdot K} \tag{6.3}$$

Thus, the absolute temperature can be measured by analyzing the intensity of the spectrum of the radiated energy from a blackbody. A source of measurement error is the emissivity of the object, which depends on the material and its surface condition. Other causes of error are deviation from the required measurement distance and the presence of any absorbing medium between the object and the detector.

Use of optical fibers as signal transmission lines in pyrometers allows remote sensing over long distances, easy installation, and accurate determination of the position to be measured by observation of a focused beam of visible light from the fiber end to the object. The sensing head consists of a flexible bundle with a large number of single fibers and lens optics to pick up the radiated energy (Fig. 6.8).

The use of a single silica fiber instead of a bundle is advantageous for measuring small objects and longer distance transmission of the picked-up radiated light. The lowest measurable temperature is 500°C, because of the unavoidable optical loss in silica fibers at wavelengths longer than 2 μm. Air cooling of the sensing head is usually necessary when the temperature exceeds 1000°C.

Optical-fiber pyrometers are one of the most successful optical-fiber sensors in the field of process control in manufacturing. Typical applications include:

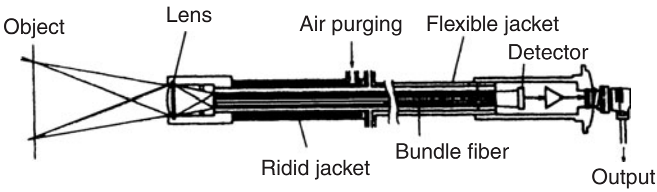


FIGURE 6.8 Schematic diagram of an optical-fiber pyrometer.

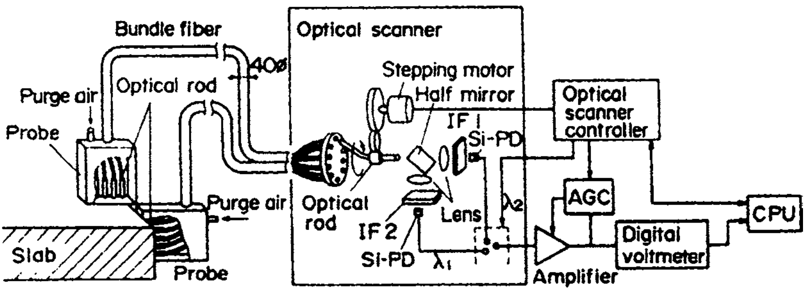


FIGURE 6.9 Temperature distribution measurement of steel slabs by an optical-fiber pyrometer using two-wavelength method.

- Casting and rolling lines in steel and other metal plants
- Electric welding and annealing
- Furnaces in chemical and metal plants
- Fusion, epitaxial growth, and sputtering processes in the semiconductor industry
- Food processing, paper manufacturing, and plastic processing

Figure 6.9 is a block diagram of the typical application of optical-fiber pyrometers for casting lines in a steel plant, where the temperature distribution of the steel slab is measured. The sensing element consists of a linear array of fused-silica optical rods, thermally protected by air-purge cooling. Light radiated from the heated slabs is collected by the optical rods and coupled into a 15-m-long bundle of fibers, which transmits light to the optical processing unit. In this system, each fiber in the bundle carries the signal from a separate lens, which provides the temperature information at the designated spot of the slabs. An optical scanner in the processing unit scans the bundle and the selected light signal is analyzed in two wavelength bands by using two optical interference filters.

PRESSURE SENSORS

If a pressure, P, acting on a diaphragm compresses a spring until an equilibrium is produced, the pressure can be represented as:

$$F(\text{kg}) = A(\text{m}^2) \times P(\text{kg/m}^2) \tag{6.4}$$

In this equation, F represents the force of the spring and A represents a surface area of the diaphragm. The movement of the spring is transferred via a system of levers to a pointer

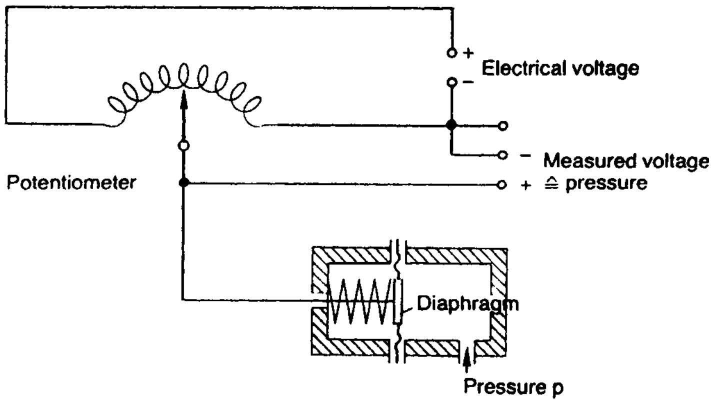


FIGURE 6.10 Deflection as a direct indication of pressure.

whose deflection is a direct indication of the pressure (Fig. 6.10). If the measured value of the pressure must be transmitted across a long distance, the mechanical movement of the pointer can be connected to a variable electrical resistance (potentiometer). A change in the resistance results in a change in the measured voltage, which can then easily be evaluated by an electronic circuit or further processed. This example illustrates the fact that a physical quantity is often subject to many transformations before it is finally evaluated.

Piezoelectric Crystals

Piezoelectric crystals may be utilized to measure pressure. Electrical charges are produced on the opposite surfaces of some crystals when they are mechanically loaded by deflection, pressure, or tension. The electrical charge produced in the process is proportional to the effective force. This change in the charge is very small. Therefore, electrical amplifiers are used to make it possible to process the signals (Fig. 6.11).

Pressure in this situation is measured by transforming it into a force. If the force produced by pressure on a diaphragm acts on a piezoelectric crystal, a signal that is proportional to the pressure measured can be produced by using suitable amplifiers.

Strain Gauges

Strain gauges can also measure pressure. The electrical resistance of a wire-type conductor is dependent, to a certain extent, on its cross-sectional area. The smaller the cross section

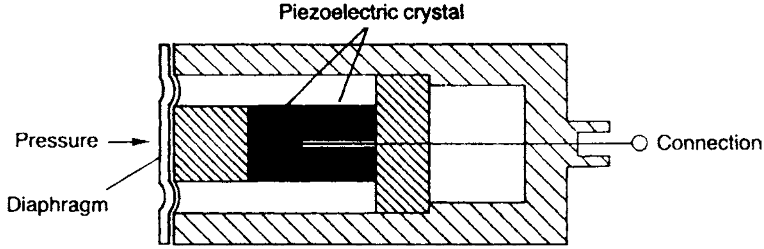


FIGURE 6.11 Electrical amplifiers are connected to a piezoelectric crystal.

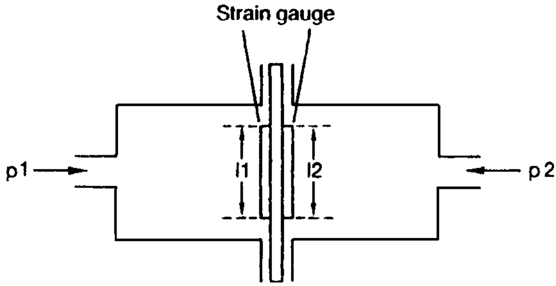


FIGURE 6.12 Strain gauge for measurement of pressure.

(i.e., the thinner the wire), the greater the resistance of the wire. A strain gauge is a wire that conducts electricity and stretches as a result of the mechanical influence (tension, pressure, or torsion) and thus changes its resistance in a manner that is detectable. The wire is attached to a carrier, which in turn is attached to the object to be measured. Conversely, for linear compression, which enlarges the cross-sectional area of a strain gauge, resistance is reduced. If a strain gauge is attached to a diaphragm (Fig. 6.12), it will follow the movement of the diaphragm. It is either pulled or compressed, depending on the flexure of the diaphragm.

FIBER-OPTIC PRESSURE SENSORS

A Y-guide probe can be used as a pressure sensor in process control if a reflective diaphragm, moving in response to pressure, is attached to the end of the fiber (Fig. 6.13). This type of pressure sensor has a significant advantage over piezoelectric transducers since it works as a noncontact sensor and has a high frequency response. The pressure signal is transferred from the sealed diaphragm to the sensing diaphragm, which is attached to the end of the fiber. With a stainless-steel diaphragm about 100 μm thick, hysteresis of less than 0.5 percent and linearity within ±0.5 percent are obtained up to the pressure level of 3×10^5 kg/m^2 (2.94 MPa) in the temperature range of −10 to +60°C.

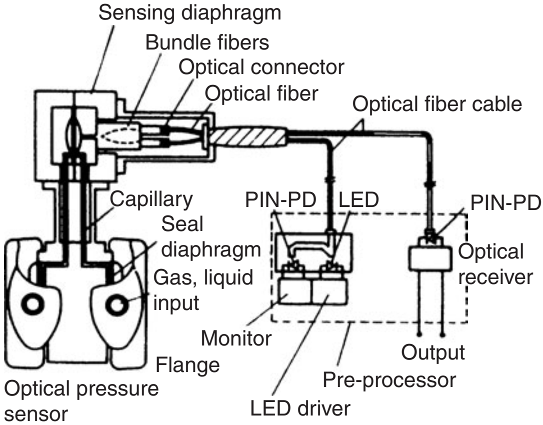


FIGURE 6.13 Schematic diagram of a fiber-optic pressure sensor using Y-guide probe with a diaphragm attached.

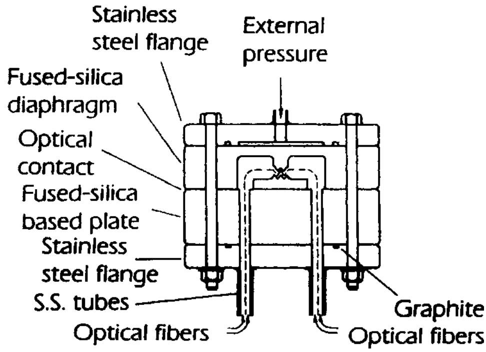


FIGURE 6.14 Fiber-optic microbend sensor.

The material selection and structural design of the diaphragm are important to minimize drift. Optical-fiber pressure sensors are expected to be used under severe environments in process control. For example, process slurries are frequently highly corrosive, and the temperature may be as high as 500°C in coal plants. The conventional metal diaphragm exhibits creep at these high temperatures. In order to eliminate such problems, an all-fused-silica pressure sensor based on the microbending effect in optical fiber has been developed (Fig. 6.14). This sensor converts the pressure applied to the fused silica diaphragm into an optical intensity modulation in the fiber.

A pressure sensor based on the wavelength filtering method has been developed. The sensor employs a zone plate consisting of a reflective surface, with a series of concentric grooves at a predetermined spacing. This zone plate works as a spherical concave mirror whose effective radius of curvature is inversely proportional to the wavelength. At the focal point of the concave mirror, a second fiber is placed that transmits the returned light to two photodiodes with different wavelength sensitivities. When broadband light is emitted from the first fiber to the zone plate, and the zone plate moves back and forth relative to the optical fibers in response to the applied pressure, the wavelength of the light received by the second fiber is varied, causing a change in the ratio of outputs from the two photodiodes. The ratio is then converted into an electrical signal that is relatively unaffected by any variations in parasitic losses.

DISPLACEMENT SENSORS FOR ROBOTIC APPLICATIONS

The operating principle of a displacement sensor using Y-guide probes is illustrated in Fig. 6.15. The most common Y-guide probe is a bifurcated fiber bundle. The light emitted from one bundle is back-reflected by the object to be measured and collected by another bundle (receiving fibers). As a result, the returned light at the detector is intensity-modulated to a degree dependent on the distance between the end of the fiber bundle and the object. The sensitivity and the dynamic range are determined by the geometrical arrangement of the array of fiber bundles and by both the number and type of the fibers. Figure 6.16 shows the relative intensity of the returned light as a function of distance for three typical arrangements: random, hemispherical, and concentric circle arrays. The intensities increase with distance and reach a peak at a certain discrete distance. After that, the intensities fall off very slowly. Most sensors use the high-sensitivity regions in these curves. Among the three arrangements, the random array has the highest sensitivity but the narrowest dynamic

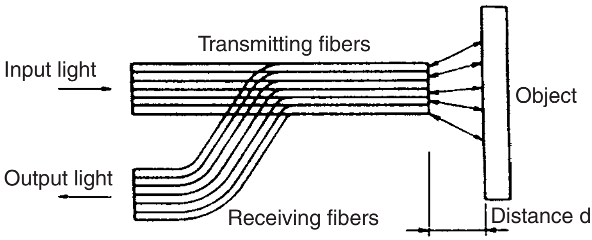


FIGURE 6.15 Principle of operation of fiber-optic mechanical sensor using a Y-guide probe.

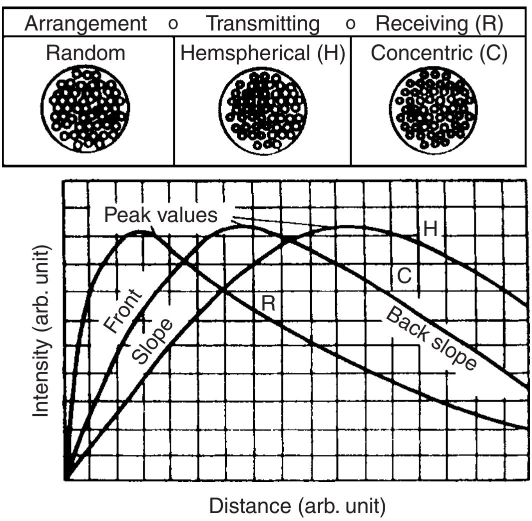


FIGURE 6.16 Relative intensity of returned light for three fiber-optic arrangements.

range. The displacement sensor using the Y-guide probe provides a resolution of 0.1 μm, linearity within 5 percent, and a dynamic range of 100 μm displacement. Y-guide probe displacement sensors are well-suited for robotics applications as position sensors and for gauging and surface assessment since they have high sensitivity to small distances.

One profound problem of this type of displacement sensor is the measuring error arising from the variation in parasitic losses along the optical transmission line. Recalibration is required if the optical path is interrupted, which limits the range of possible applications. In order to overcome this problem, a line-loss-independent displacement sensor with an electrical subcarrier phase encoder has been implemented. In this sensor, the light from an LED modulated at 160 MHz is coupled into the fiber bundle and divided into two optical paths. One of the paths is provided with a fixed retroreflector at its end. The light through the other is reflected by the object. The two beams are returned to the two photodiodes separately. Each signal, converted into an electric voltage, is electrically heterodyned into an intermediate frequency at 455 kHz. Then, the two signals are fed to a digital phase comparator, the output of which is proportional to the path distance. The resolution of the optical path difference is about 0.3 mm, but improvement of the receiver electronics will provide a higher resolution.

PROCESS CONTROL SENSORS MEASURING AND MONITORING LIQUID FLOW

According to the laws of fluid mechanics, an obstruction inserted in a flow stream creates a periodic turbulence behind it. The frequency of shedding the turbulent vortices is directly proportional to the flow velocity. The flow sensor in Fig. 6.17 has a sensing element consisting of a thin metallic obstruction and a downstream metallic bar attached to a multimode fiber-microbend sensor. As illustrated in Fig. 6.18, the vortex pressure produced at the metallic bar is transferred, through a diaphragm at the pipe wall that serves as both a seal and a pivot for the bar, to the microbend sensor located outside the process line pipe. The microbend sensor converts the time-varying mechanical force caused by the vortex shedding into a corresponding intensity modulation of the light. Therefore, the frequency of the signal converted into the electric voltage at the detector provides the flow-velocity information. This flow sensor has the advantage that the measuring accuracy is essentially independent of any changes in the fluid temperature, viscosity, or density, and in the light source intensity. According to the specifications for typical optical vortex-shedding flow sensors, flow rate can be measured over a Reynolds number range from 5×10^3 to 6000×10^3 at temperatures from −100 to +600°C. This range is high compared to that of conventional flow meters. In addition, an accuracy of ±0.4 and ±0.7 percent, respectively, is obtained for liquids and gases with Reynolds numbers above 10,000.

Flow Sensor Detecting Small Air Bubbles for Process Control in Manufacturing

Another optical-fiber flow sensor employed in manufacturing process control monitors a two-fluid mixture (Fig. 6.19). The sensor can distinguish between moving bubbles and liquid in the flow stream and display the void fraction, namely, the ratio of gas volume to the total volume.

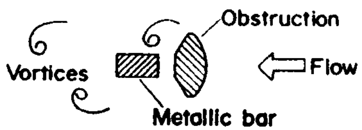


FIGURE 6.17 Principle of operation of a vortex-shedding flow sensor.

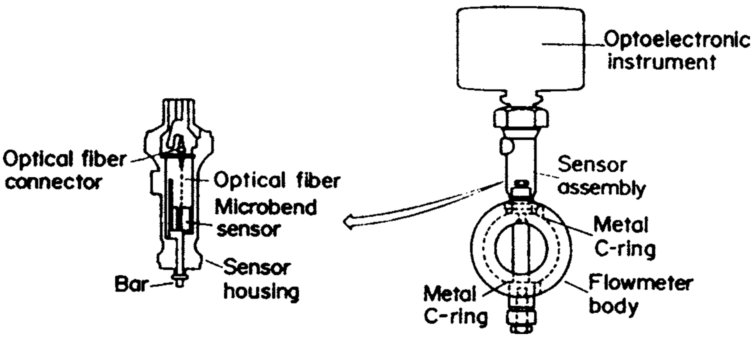


FIGURE 6.18 Schematic diagram of a vortex-shedding flow sensor.

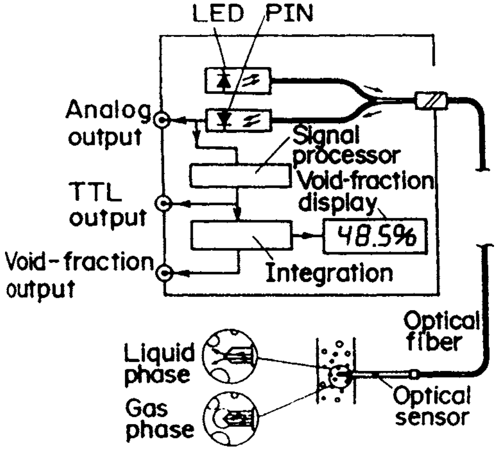


FIGURE 6.19 Flow sensor for two-phase mixtures.

The principle of operation is quite simple. The light from the LED is guided by the optical fiber to the sensing element, in which the end portion of the fiber is mounted in a stainless steel needle of 2.8-mm outer diameter. When liquid is in contact with the end of the fiber, light enters the fluid efficiently and very little light is returned. However, when a gas bubble is present, a significant fraction of light is reflected back. With this technique, bubbles as small as 50 μm may be detected with an accuracy of better than 5 percent and a response time of only 10 μs.

Potential applications of this flow sensor for the control of processes in manufacturing systems are widespread—for example, detection of gas plugs in production wells in the oil industry and detection of fermenters and distillers in the blood-processing and pharmaceutical industries.

An optical-fiber flow sensor for a two-phase mixture based on Y-guide probes is shown in Fig. 6.20. Two Y-guide probes are placed at different points along the flow stream to emit the input light and to pick up the retroreflected light from moving solid particles in the flow. The delay time between the signals of the two probes is determined by the average velocity of the moving particles. Therefore, measurement of the delay time by a conventional correlation technique provides the flow velocity. An accuracy of better than ±1 percent and a dynamic range of 20:1 are obtained for flow velocities up to 10 m/s. A potential problem of such flow sensors for two-phase mixtures is poor long-term stability, because the optical fibers are inserted into the process fluid pipes.

Liquid Level Sensors in Manufacturing Process Control for Petroleum and Chemical Plants

Several optical-fiber liquid level sensors developed in recent years have been based on direct interaction between the light and liquid. The most common method in commercial products employs a prism attached to the ends of two single optical fibers (Fig. 6.21). The input light from an LED is totally internally reflected and returns to the output fiber when the prism is in air. However, when the prism is immersed in liquid, the light refracts into the fluid with low reflection, resulting in negligible returned light. Thus, this device works as a liquid level switch. The sensitivity of the sensor is determined by the contrast ratio, which depends on the refractive index of the liquid. Typical examples of signal output change for liquids with different refractive indices are indicated in Table 6.2.

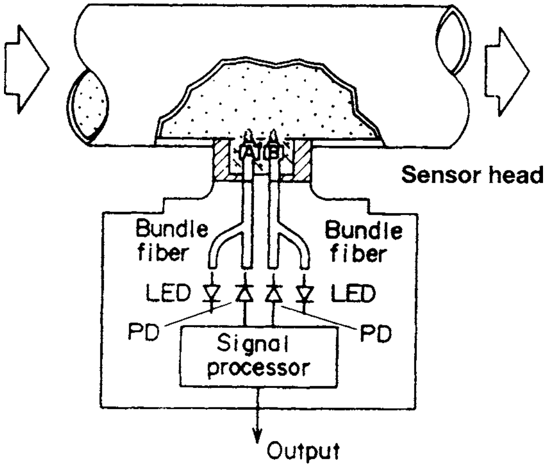


FIGURE 6.20 Flow sensor using two Y-guided probes based on a correlation technique.

The output loss stays at a constant value of 33 dB for refractive indices higher than 1.40. The signal output of a well-designed sensor can be switched for a change in liquid level of only 0.1 mm.

Problems to be solved for this sensor are dirt contamination on the prism surface and bubbles in the liquid. Enclosing the sensing element with a fine filter helps keep it clean and simultaneously reduces level fluctuations caused by bubbles. Since optical-fiber liquid level sensors have the advantages of low cost and electrical isolation, their use is widespread in petroleum and chemical plants, where the hazardous environment causes difficulties with conventional sensors. They are used, for example, to monitor storage tanks in a petroleum plant.

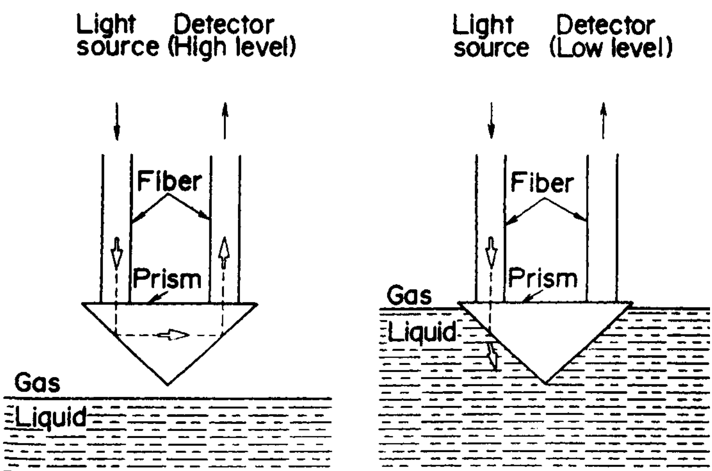


FIGURE 6.21 Principle of operation of a liquid level sensor with a prism attached to two optical fibers.

TABLE 6.2 Refractive Index Versus Output

Refractive index, n	Loss change, dB
1.333	2.1
1.366	4.6
1.380	6.0
1.395	31.0

Another optical-fiber liquid level sensor, developed for the measurement of boiler-drum water level, employs a triangularly shaped gauge through which red and green light beams pass. The beams are deflected as it fills with water, so that the green light passes through an aperture. In the absence of water, only red light passes through. Optical fibers transmit red or green light from individual gauges to a plant control room located up to 150 m from the boiler drum (Fig. 6.22). The water level in the drum is displayed digitally.

This liquid level sensor operates at temperatures up to 170°C and pressures up to 3200 lb/in^2 gauge. Many sensor units are installed in the boiler drum, and most have been operating for seven years. This sensor is maintenance-free, fail-safe, and highly reliable.

On-line Measuring and Monitoring of Gas by Spectroscopy

An optical spectrometer or optical filtering unit is often required for chemical sensors because the spectral characteristics of absorbed, fluorescent, or reflected light indicate the presence, absence, or precise concentration of a particular chemical species (Fig. 6.23).

Sensing of chemical parameters via fibers is usually done by monitoring changes in a suitably selected optical property—absorbance, reflectance, scattering (turbidity), or luminescence (fluorescence or phosphorescence), depending on the particular device. Changes in parameters such as the refractive index may also be employed for sensing purposes. The change in light intensity due to absorption is determined by the number of absorbing species in the optical path, and is related to the concentration C of the absorbing species by the Beer-Lambert relationship. This law describes an exponential reduction of light intensity with distance (and also concentration) along the optical path. Expressed logarithmically,

$$A = \log I_0/I = \eta l C \tag{6.5}$$

where A is the optical absorbance, l is the path length of the light, η is the molar absorptivity, and I_0 and I are the incident and transmitted light, respectively. For absorption measurements via optical fibers, the medium normally must be optically transparent.

An accurate method for the detection of leakage of flammable gases such as methane (CH_4), propane (C_3H_8), and ethylene (C_2H_4) is vital in gas and petrochemical plants in order to avoid serious accidents. The recent introduction of low-loss fiber into spectroscopic measurements of these gases offers many advantages for process control in manufacturing:

- Long-distance remote sensing
- On-line measurement and monitoring
- Low cost
- High reliability

The most commonly used method at present is to carry the sample back to the measuring laboratory for analysis. Alternatively, numerous spectrometers may be used at various

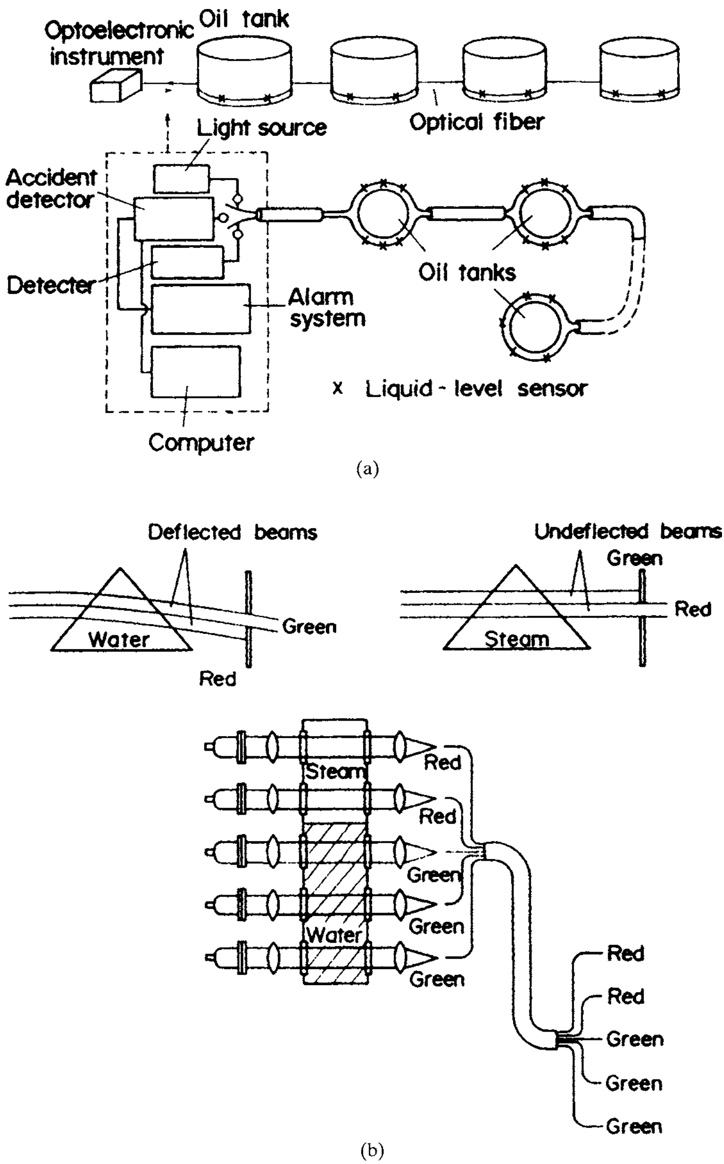


FIGURE 6.22 (*a*) Principle of operation of liquid level sensor for the measurement of boiler-drum water. (*b*) Liquid level sensor for the measurement of boiler-drum water with five-port sensors.

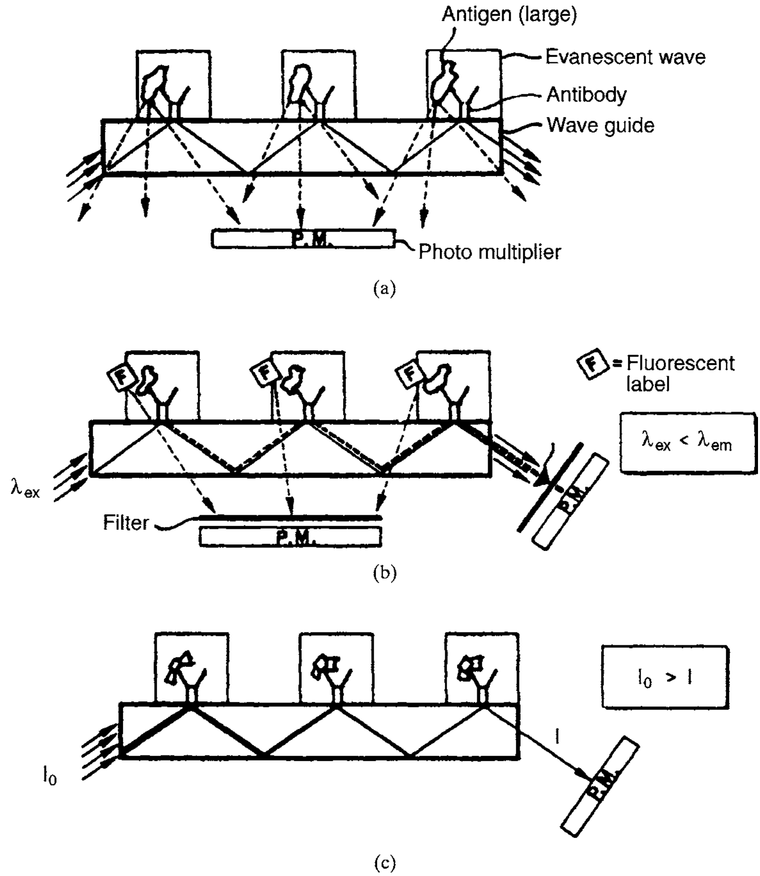


FIGURE 6.23 (*a*) Light scattering from the waveguide surface increases as the antigen-antibody complexes are formed, and is detected by a side-mounted photodetector. (*b*) Fluorescence is excited in a fluorescent marker, attached to an antigen molecule, by light passing through the waveguide. The fluorescent light can be collected either sideways from the guide, or as light that is retrapped by the guide and directed to a photodetector. (*c*) Absorption of light by antibody-antigen complexes on the surface attenuates light traveling down the waveguide.

points around the factory. The new advances in spectroscopic measurements allow even CH_4 to be observed at a distance of 10 km with a detection sensitivity as low as 5 percent of the lower explosion limit (LEL) concentration. The optical-fiber gas measuring system employs an absorption spectroscopy technique, with the light passing through a gas-detection cell for analysis. The overtone absorption bands of a number of flammable gases are located in the near-infrared range (Fig. 6.24).

The optical gas sensing system can deal with a maximum of 30 detection cells (Fig. 6.25). The species to be measured are CH_4, C_3H_8, and C_2H_4 molecules. Light from a halogen lamp (an infrared light source) is distributed into a bundle of 30 single optical fibers. Each of the distributed beams is transmitted through a 1-km length of fiber to a corresponding gas detection cell. The receiving unit is constructed of three optical switches, a rotating sector with four optical interference filters, and three Ge photodiodes. Each optical switch can select any ten returned beams by specifying the number of the cell. The peak transmission

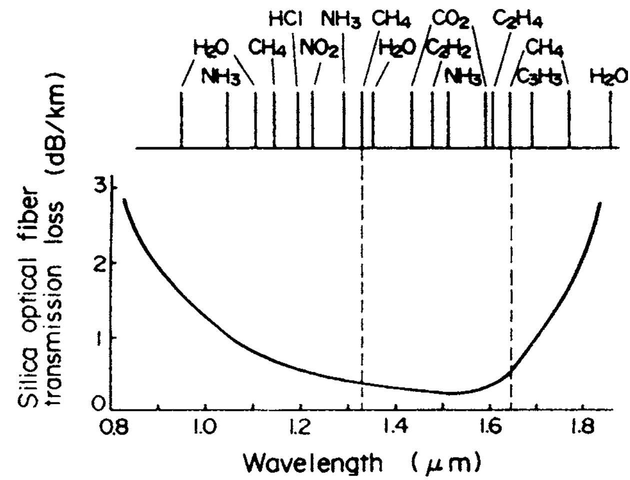


FIGURE 6.24 Absorption lines of typical flammable gases in the near-infrared and the transmission loss of silica fiber.

wavelength of the optical filter incorporated in the sensor is 1.666 μm for CH_4, 1.690 μm for C_3H_8, 1.625 μm for C_2H_2, and 1.600 μm for a reference beam. After conversion to electrical signals, the signal amplitudes for the three gases are normalized by the reference amplitude. Then the concentration of each gas is obtained from a known absorption-concentration calibration curve stored in a computer.

An intrinsic distributed optical-fiber gas sensor for detecting the leakage of cryogenically stored gases such as CH_4, C_2H_4, and N_2 has also been developed. The sensor's operation is based on the temperature-dependent transmission loss of optical fiber—that is, the optical fiber is specially designed so the transmission loss increases with decreasing temperature by choosing the appropriate core and cladding materials. Below the critical temperature, in the region of –55°C, most of the light has transferred to the cladding layer,

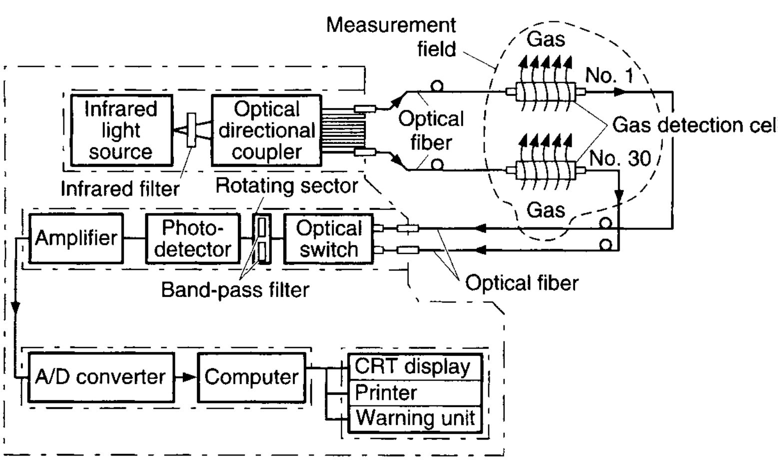


FIGURE 6.25 Gas detection system with 30 detection cells.

and the light in the core is cut off. By connecting this temperature-sensitive fiber between a light source and a detector and monitoring the output light level, the loss of light resulting from a cryogenic liquid in contact with the fiber can be detected directly.

CRACK DETECTION SENSORS FOR COMMERCIAL, MILITARY, AND SPACE INDUSTRY USE

Accurate and precise detection of crack propagation in aircraft components is of vital interest for commercial and military aviation and the space industry. A system has been recently developed to detect cracks and crack propagation in aircraft components. This system uses optical fibers of small diameter (20 to 100 μm), which can be etched to increase their sensitivity. The fibers are placed on perforated adhesive foil to facilitate attachment to the desired component for testing. The fiber is in direct contact with the component (Fig. 6.26). The foil is removed after curing of the adhesive. Alternatively, in glass-fiber-reinforced plastic (GFRP) or carbon-fiber-reinforced plastic (CFRP), materials used more and more in aircraft design, the fiber can be easily inserted in the laminate without disturbing the normal fabrication process. For these applications, bare single fiber or prefabricated tape with integrated bundles of fibers is used. The system initially was developed for fatigue testing of aircraft components such as frames, stringers, and rivets. In monitoring mode, the system is configured to automatically interrupt the fatigue test. The system has also been applied to the inspection of the steel rotor blades of a 2-MW wind turbine. A surveillance system

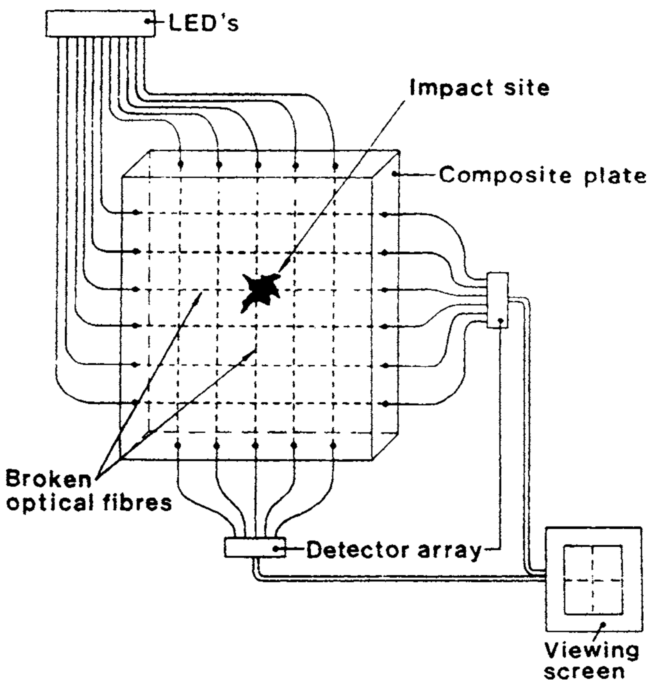


FIGURE 6.26 Schematic diagram of a fiber-optic system showing the location of impact damage in a composite structure.

has been developed for the centralized inspection of all critical components of the Airbus commercial jetliner during its lifetime. This fiber nervous system is designed for in-flight monitoring and currently is accessible to flight and maintenance personnel.

An optical-fiber mesh has been tested for a damage assessment system for a GFRP submarine sonar dome. Two sets of orthogonally oriented fibers are nested in the laminate during the fabrication process. When the fibers of the mesh are properly connected to LEDs and the detectors, the system can be configured to visualize the location of a damaged area.

As an alternative, a video camera and image processing are applied to determine the position of the damaged area. The fiber end faces at the detection side of the mesh are bundled and imaged into the camera tube. Two images are subtracted: the initial image before the occurrence of damage and the subsequent image. If fibers are broken, their location is highlighted as a result of this image subtraction.

CONTROL OF INPUT/OUTPUT SPEED OF CONTINUOUS WEB FABRICATION USING LASER DOPPLER VELOCITY SENSOR

A laser Doppler velocimeter (LDV) can be configured to measure any desired component velocity, perpendicular or parallel to the direction of the optical axis. An LDV system has been constructed with a semiconductor laser and optical fibers and couplers to conduct the optical power. Frequency modulation of the semiconductor laser (or, alternatively, an external fiber-optic frequency modulator) is used to introduce an offset frequency. Some commercial laser Doppler velocimeters are available with optical-fiber leads and small sensing heads. However, these commercial systems still use bulk optical components such as acoustooptic modulators or rotating gratings to introduce the offset frequency.

With an LDV system, the velocity can be measured with high precision in a short period of time. This means that the method can be applied for real-time measurements to monitor and control the velocity of objects as well as measure their vibration. Because the laser light can be focused to a very small spot, the velocity of very small objects can be measured, or if scanning techniques are applied, high spatial resolution can be achieved. This method is used for various applications in manufacturing, medicine, and research. The demands on system performance with respect to sensitivity, measuring range, and temporal resolution are different for each of these applications.

In manufacturing processes, for example, LDV systems are used to control continuous roll milling of metal (Fig. 6.27), to control the rolling speed of paper and films, and to monitor fluid velocity and turbulence in mixing processes. Another industrial application is vibration

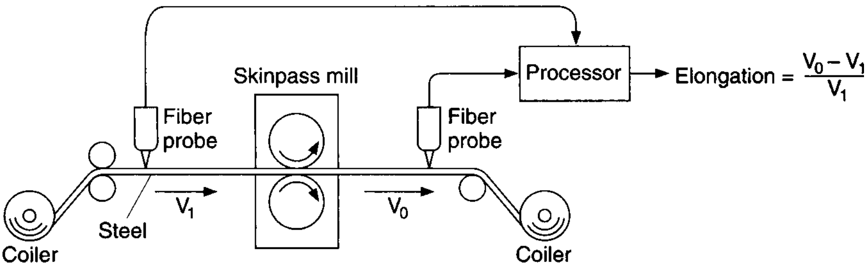


FIGURE 6.27 Fiber-optic laser Doppler velocimeter at a rolling mill controls pressure by measuring input speeds.

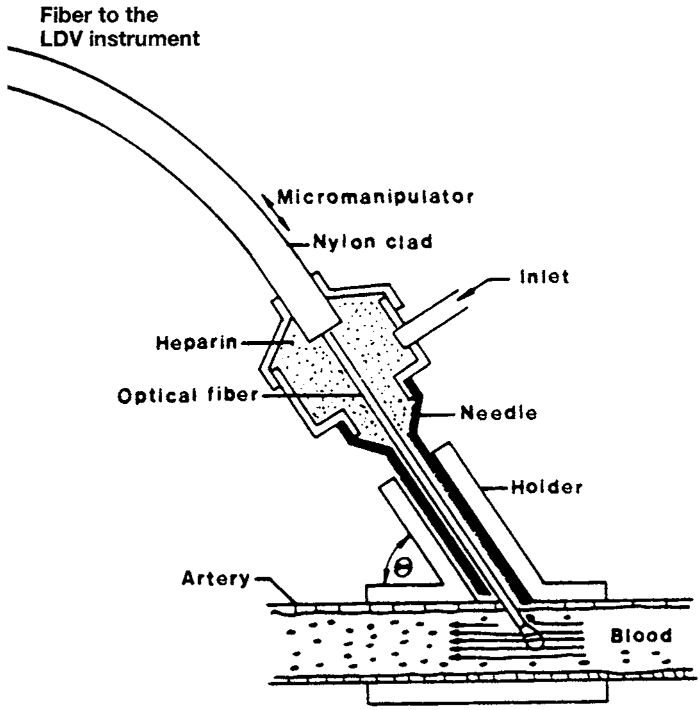


FIGURE 6.28 Special probe for measurement of blood velocity.

analysis. With a noncontact vibrometer, vibration of machines, machine tools, and other structures can be analyzed without disturbing the vibrational behavior of the structure.

Interestingly, the LDV system proved useful in the measurement of arterial blood velocity (Fig. 6.28), thereby providing valuable medical information. Another application in medical research is the study of motion of the tympanic membrane in the ear.

ULTRASONIC/LASER NONDESTRUCTIVE EVALUATION SENSOR

Ultrasonic/laser optical inspection is a relatively new noncontact technique. A laser system for generating ultrasound pulses without distortion of the object surface is shown in Fig. 6.29. A laser pulse incident on a surface will be partly absorbed by the material and will thus generate a sudden rise in temperature in the surface layer of the material. This thermal shock causes expansion of a small volume at the surface, which generates thermoelastic strains. Bulk optical systems have been used previously to generate the laser pulse energy. However, the omnidirectionality of bulk sources is completely different from other well-known sources, and is regarded as a serious handicap to laser generation.

To control the beamwidth and beam direction of the optically generated ultrasonic waves, a fiber phased array has been developed. In this way, the generated ultrasonic beam

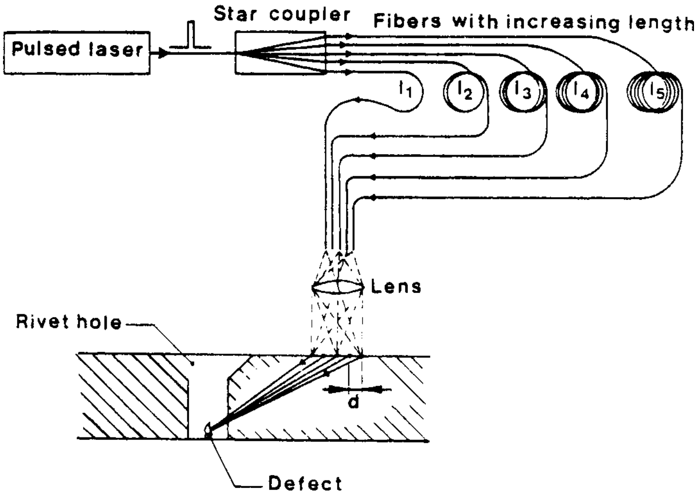


FIGURE 6.29 Setup for beam steering of laser-generated ultrasound by fiber-optic phased array.

can be focused and directed to a particular inspection point below the surface of an object (Fig. 6.29). This system has been optimized for the detection of fatigue cracks at rivet holes in aircraft structures.

The combination of laser-generated ultrasound and an optical-fiber interferometer for the detection of the resultant surface displacement has led to a technique that is useful for a wide variety of inspection tasks in manufacturing, including areas difficult to access and objects at high temperature, as well as more routine inspection and quality control in various industrial environments. Such a system can be applied to the measurement of thickness, velocity, flaws, defects, and grain size in a production process.

PROCESS CONTROL SENSOR FOR ACCELERATION

The principle of operation of the process control acceleration sensor is illustrated in Fig. 6.30. The sensor element, consisting of a small cantilever and a photoluminescent material, is attached to the end of a single multimode fiber. The input light of wavelength λ_s is transmitted along the fiber from a near-infrared LED source to the sensor element. The sensor element returns light at two different wavelengths—one of which serves as a signal light and the other as a reference light—into the same fiber. The signal light at wavelength λ_s is generated by reflection from a small cantilever. Since the relative angle of the reflected light is changed by the acceleration, the returned light is intensity-modulated. The reference light of wavelength λ_r is generated by photoluminescence of a neodymium-doped glass element placed close to the sensor end of the fiber.

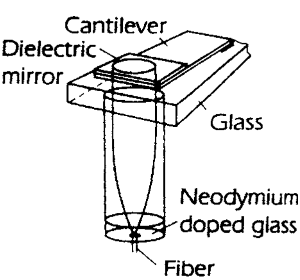


FIGURE 6.30 Cantilever-type acceleration sensor.

The optoelectronic detector module has two optical filters to separate the signals λ_s and λ_r, and also two photodiodes to convert the signal and the reference light into separate analog voltages. The signal processing for compensation is then merely a matter of electrical division. A measuring range of 0.1 to 700 m/s^2 and a resolution of 0.1 m/s^2 is obtained over the frequency range of 5 to 800 Hz.

AN ENDOSCOPE AS IMAGE TRANSMISSION SENSOR

An imaging cable consists of numerous optical fibers, typically 3000 to 100,000, each of which has a diameter of 10 μm and constitutes a picture element (pixel). The principle of image transmission through the fibers is shown in Fig. 6.31. The optical fibers are aligned regularly and identically at both ends of the fibers. When an image is projected on one end of the image fiber, it is split into multiple picture elements. The image is then transmitted as a group of light dots with different intensities and colors, and the original picture is reduced at the far end. The image fibers developed for industrial use are made of silica glass with low transmission loss over a wide wavelength band from visible to near infrared, and can therefore transmit images over distances in excess of 100 m without significant color changes. The basic structure of the practical optical-fiber image sensing system (endoscope) is illustrated in Fig. 6.32. It consists of the image fiber, an objective lens to project the image on one end, an eyepiece to magnify the received image on the other end, a fiber protection tube, and additional fibers for illumination of the object.

Many examples have been reported of the application of image fibers in process control. Image fibers are widely employed to observe the interior of blast furnaces and the burner flames of boilers, thereby facilitating supervisory control. Image fibers can operate at temperatures up to 1000°C, when provided with a cooling attachment for the objective lens and

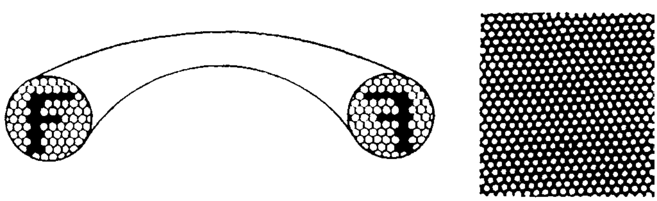


FIGURE 6.31 Image transmission through an image fiber.

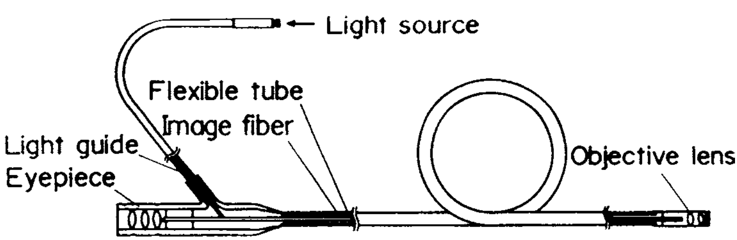


FIGURE 6.32 Basic structure of fiber scope.

its associated equipment. Another important application of the image fiber bundles is observation, control, and inspection of nuclear power plants and their facilities. Conventional image fibers cannot be used within an ionizing radiation environment because ordinary glass becomes colored when exposed to radiation, causing increasing light transmission loss. A high-purity silica core fiber is well-known as a radiation-resistant fiber for nuclear applications.

The endoscope has demonstrated its vital importance in medical and biochemical fields such as:

- Angioplasty
- Laser surgery
- Gastroscopy
- Cystoscopy
- Bronchoscopy
- Cardioscopy

SENSOR NETWORK ARCHITECTURE IN MANUFACTURING

In fiber-optic sensor networks, the common technological base with communication is exploited by combining the signal generating ability of sensors and the signal transmitting capability of fiber optics. This combination needs to be realized by a suitable network topology in various manufacturing implementations. The basic topologies for sensor networking are illustrated in Fig. 6.33. These topologies are classified into six categories:

- Linear array network with access-coupled reflective sensors (Fig. 6.33*a*).
- Ring network with in-line transmissive sensors (Fig. 6.33*b*).
- Star network with reflective sensors (Fig. 6.33*c*).
- Star network with reflective sensors; one or more sensors can be replaced by a separate star network, in order to obtain a tree network (Fig. 6.33*d*).

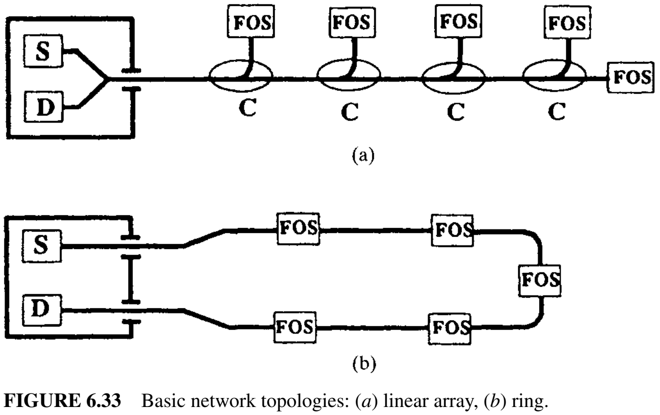


FIGURE 6.33 Basic network topologies: (*a*) linear array, (*b*) ring.

- Star network that can also be operated with transmissive sensors (Fig. 6.33*e*).
- Ladder network with two star couplers. A star coupler is replaced by several access couplers, the number required being equal to the number of sensors (Fig. 6.33*f*).

Topological modifications, especially of sensor arrays and ladder networks, may be desirable in order to incorporate reference paths of transmissive (dummy sensors) or reflective sensors (splices, open fiber end). The transmit and return fibers, or fiber highway, generally share a common single-fiber path in networks using reflective sensors.

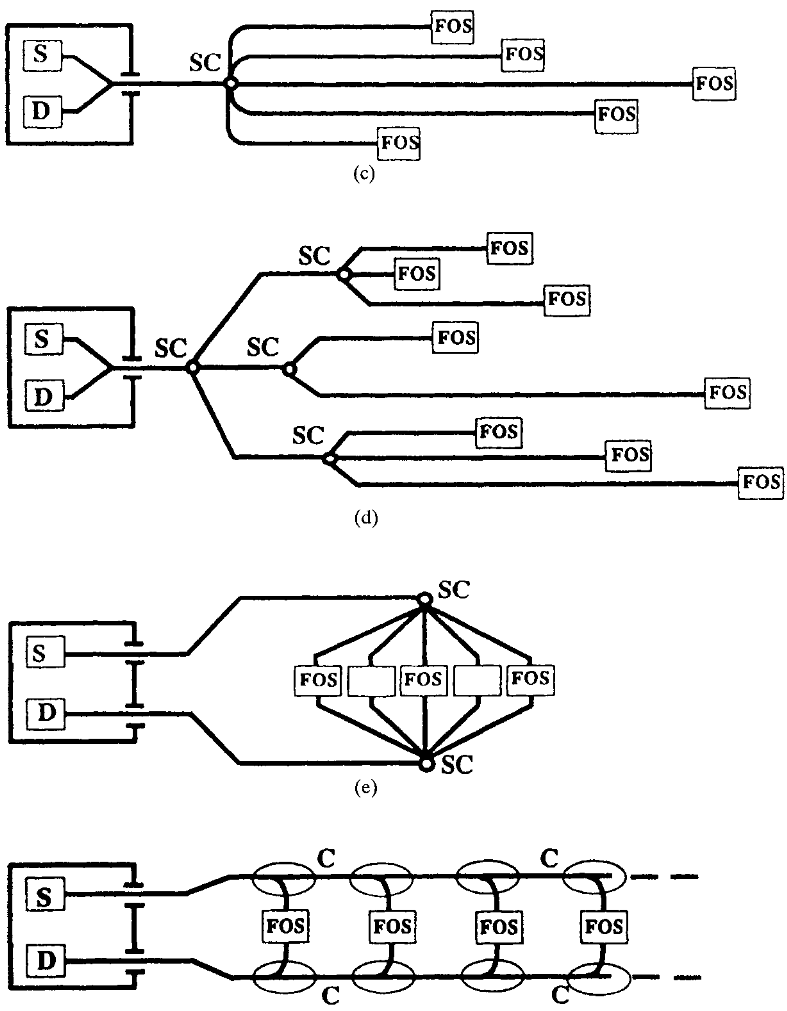


FIGURE 6.33 *(Continued)* Basic network topologies: (*c*) reflective star, (*d*) reflective tree, (*e*) transmissive star, and (*f*), ladder network.

When a suitable fiber-optic network topology is required, various criteria must be considered:

- The sensor type, encoding principle, and topology to be used
- The proposed multiplexing scheme, required number of sensors, and power budget
- The allowable cross-communication level
- The system cost and complexity constraints
- The reliability (i.e., the effect of component failure on system performance)

POWER LINE FAULT-DETECTION SYSTEM FOR POWER GENERATION AND DISTRIBUTION INDUSTRY

In power distribution lines, faults such as short circuits, ground faults, and lightning strikes on the conductors must be detected in a very short time to prevent damage to equipment and power failure, and to enable quick repair. If the transmission line is divided in sections and a current or magnetic-field sensor is mounted in each section, a faulty section can be determined by detection of a change of the level and phase of the current on the power line. A system was developed as a hybrid optical approach to a fault-locating system that detects the phase and current difference between two current transformers on a composite fiber-optic ground wire (OPGW) wherein, due to induction, current is constantly passing (Fig. 6.34). The signal from a local electrical sensor, powered by solar cells and batteries, is transmitted over a conventional optical-fiber communication link. By three-wavelength multiplexing, three sensor signals can be transmitted over a single fiber. Seven sensors, three at each side of a substation and one at the substation itself, can monitor one substation on the power line, using one fiber in the OPGW.

Another system uses current transformers to pick up lightning current and thus detect lightning strikes. The signal is transmitted to a central detection point using the OPGW. Every sensor has its own OPGW fiber. This system is on a 273-kV power line in Japan.

The OPGW opens the possibility for using all kinds of sensors along the transmission line. These sensors may not only be for locating faults, but also for monitoring structural

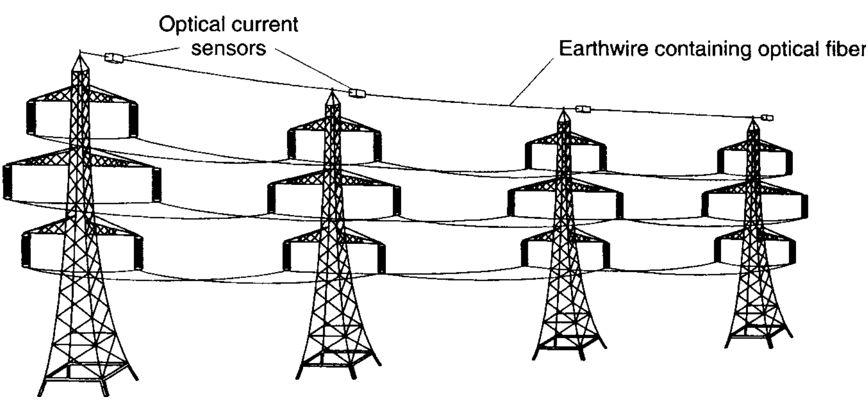


FIGURE 6.34 Fault locating system based on current and phase in ground wire.

integrity. The use of optical time-domain reflectometry combined with passive intrinsic (distributed) sensors along the OPGW has future potential for providing a convenient and powerful monitoring method for power lines.

REFERENCES

1. Bailey Control Systems, Wickliffe, Ohio.
2. Bartman, R. K., B. R. Youmans, and N. M. Nerheim. "Integrated Optics Implementation of a Fiber Optic Rotation Sensor: Analysis and Development," *Proc. SPIE* 719, 122–134.
3. Berthold, J. W., "Industrial Applications of Optical Fiber Sensors," *Fiber Optic and Laser Sensors III*, *Proc. SPIE* 566, 37–44.
4. Carrol, R., C. D. Coccoli, D. Cardelli, and G. T. Coate, "The Passive Resonator Fiber Optic Gyro and Comparison to the Interferometer Fiber Gyro," *Proc. SPIE* 719, 169–177 (1986).
5. Chappel, A. (ed.), *Optoelectronics—Theory and Practice*, McGraw-Hill, New York, 1978.
6. Crane, R. M., A. B. Macander, D. W. Taylor, and J. Gagorik, "Fiber Optics for a Damage Assessment System for Fiber Reinforced Plastic Composite Structures," *Rev. Progress in Quantitative NDE*, 2B, Plenum Press, New York, 1419–1430 (1982).
7. Doeblin, E. O., *Measurement Systems—Application and Design*, 4th ed., McGraw-Hill, New York, 1990.
8. Fields, J. N., C. K. Asawa, O. G. Ramer, and M. K. Barnoski, "Fiber Optic Pressure Sensor," *J. Acoust. Soc. Am.* 67, 816 (1980).
9. Finkelstein, L., and R. D. Watts, "Fundamental of Transducers—Description by Mathematical Models," *Handbook of Measurement Science*, vol. 2, P. H. Sydenham (ed.), Wiley, New York, 1983.
10. Friebele, E. L. and M. E. Gingerich, "Radiation-Induced Optical Absorption Bands in Low Loss Optical Fiber Waveguides," *J. Non-Crust. Solids* 38(39), 245–250 (1980).
11. Henze, M., "Fiber Optics Temperature and Vibration Measurements in Hostile Environments," Technical Material, *ASEA Research and Innovation*, CF23-1071E (1987).
12. Hofer, B., "Fiber Optic Damage Detection in Composite Structure," *Proc. 15th Congress. Int. Council Aeronautical Science*, ICAS-86-4.1.2, 135–143 (1986).
13. Kapany, N. S., *Fiber Optics, Principles and Applications*, Academic Press, London, 1976.
14. Lagakos, N., et al., "Multimode Optical Fiber Displacement Sensor," *Appl. Opt.* 20, 167 (1981).
15. Liu, K., "Optical Fiber Displacement Sensor Using a Diode Transceiver," Fiber Optic Sensors II, A. M. Sheggi (ed.), *Proc. SPIE* 798, 337–341 (1987).
16. Mizuno, Y., and T. Nagai, "Lighting Observation System on Aerial Power Transmission Lines by Long Wavelength Optical Transmission," *Applications of Fiber Optics in Electrical Power Systems in Japan*, C.E.R.L. Letterhead, paper 5.
17. Mori, S., et al., "Development of a Fault-Locating System Using OPGW," *Simitomo Electric Tech. Rev.* 25, 35–47.
18. Norton, H. N., *Sensors and Analyzer Handbook*, Prentice-Hall, Englewood Cliffs, N.J., 1982.
19. Neubert, H.K.P., *Instrument Transducers*, 2d ed., Clarendon Press, Oxford, 1975.
20. Ogeta, K., *Modern Control Engineering*, 2d ed. Prentice-Hall, Englewood Cliffs, N.J., 1990.
21. Petrie, G. R., K. W. Jones, and R. Jones, "Optical Fiber Sensors in Process Control," *4th Int. Conf. Optical Fiber Sensors, OFS'86*, Informal Workshop at Tsukuba Science City, VIII I–VIII 19, (1986).
22. Place, J. D., "A Fiber Optic Pressure Transducer Using a Wavelength Modulation Sensor," *Proc. Conf. Fiber Optics '85* (*Sira*), London, (1985).
23. Ramakrishnan, S., L. Unger, and R. Kist, "Line Loss Independent Fiberoptic Displacement Sensor with Electrical Subcarrier Phase Encoding," *5th Int. Conf. Optical Fiber Sensors, OFS '88*, New Orleans, 133–136 (1988).

24. Sandborn, V. A., *Resistance Temperature Transducers*, Metrology Press, Fort Collins, Colo., 1972.
25. Scruby, C. B., R. J. Dewhurst, D. A. Hutchins, and S. B. Palmer, "Laser Generation of Ultrasound in Metals," *Research Techniques in Nondestructive Testing*, 15, R. S. Sharpe (ed.), Academic Press, London, 1982.
26. Tsumanuma, T., et al., "Picture Image Transmission-System by Fiberscope," *Fujikura Technical Review* 15, 1–10 (1986).
27. Vogel, J. A., and A.J.A. Bruinsma, "Contactless Ultrasonic Inspection with Fiber Optics," *Conf. Proc. 4th European Conf. Non-Destructive Testing*, Pergamon Press, London, 1987.
28. Yasahura, T., and W. J. Duncan, "An Intelligent Field Instrumentation System Employing Fiber Optic Transmission," *Advances in Instrumentation*, ISA, Wiley, London, 1985.

CHAPTER 7
SENSORS IN FLEXIBLE MANUFACTURING SYSTEMS

INTRODUCTION

Flexibility has become a key goal in manufacturing, hence the trend toward flexible manufacturing systems. These are designed to produce a variety of products from standard machinery with a minimum of workers. In the ultimate system, raw material in the form of bars, plates, and powder would be used to produce any assembly required without manual intervention in manufacture. Clearly, this is a good breeding ground for robots.

But it should be emphasized that the early FMSs are, in fact, direct numerical control (DNC) systems for machining. And it must be acknowledged that an NC machine tool is really a special-purpose robot. It manipulates a tool in much the same way as a robot handles a tool or welding gun. Then, with no more than a change in programming, it can produce a wide range of products. Moreover, the controllers for robots and NC machines are almost the same. But for an NC machine to be converted into a self-supporting flexible system, it needs some extra equipment, including a handling device. It then forms an important element in an FMS.

The principle of flexible manufacturing for machining operations is that the NC machining cells are equipped with sensors to monitor tool wear and tool breakage. Such cells are able to operate unmanned so long as they can be loaded by a robot or similar device, since the sensors will detect any fault and shut the operation down if necessary. The requirements can be summarized as:

- CNC system with sufficient memory to store many different machining programs.
- Automatic handling at the machining tool either by robot or other material handling system.
- Workpieces stored near the machine to allow unmanned operation for several hours. A guided vehicle system may be employed if workpieces are placed at a designated storage and retrieval system away from the machine.
- Various sensors to monitor, locate, and/or diagnose any malfunction.

THE ROLE OF SENSORS IN FMS

The monitoring sensor devices are generally situated at the location of the machining process, measuring workpiece surface textures, cutting-tool vibrations, contact temperature between cutting tool and workpiece, flow rate of cooling fluid, electrical current fluctuations, and so on. Data in the normal operating parameters are stored in memory with data on acceptable manufacturing limits. As the tool wears, the tool changer is actuated. If the current rises significantly, along with other critical signals from sensors, a tool failure is indicated. Hence, the machine is stopped. Thus, with the combination of an NC machine, parts storage and retrieval, handling devices, and sensors, the unmanned cell becomes a reality. Since the control system of the NC machine, robot, and unmanned guided vehicles are similar, central computer control can be used effectively.

Systems based on these principles have been developed. In Japan, Fanauc makes numerical controllers, small NC and EDM machines, and robots; Murata makes robot trailers as well as a variety of machinery including automated textile equipment; and Yamazaki makes NC machines. In France, Renault and Citroen use FMS to machine gear boxes for commercial vehicles and prototype engine components, respectively, while smaller systems have been set up in many other countries.

The central element in establishing an error-free production environment is the availability of suitable sensors in manufacturing. The following represents a summary of sensing requirements in manufacturing applications:

- Part identification
- Part presence or absence
- Range of object for handling
- Single-axis displacement of measurement
- Two-dimensional location measurement
- Three-dimensional location measurement

Current Available Sensor Technology for FMS

The currently available sensors for manufacturing applications can be classified into four categories:

- Vision sensors
  - Photodetector
  - Linear array
  - TV camera
  - Laser triangulation
  - Laser optical time-domain reflectometry
  - Optical fiber
- Tactile sensors
  - Probe
  - Strain gauges
  - Piezoelectric
  - Carbon material
  - Discrete arrays
  - Integrated arrays
- Acoustic sensors
  - Ultrasonic detectors and emitters

- Ultrasonic arrays
- Microphones (voice control)
- Passive sensors
  - Infrared
  - Magnetic proximity
  - Ionizing radiation
  - Microwave radar

Integrating vision sensors and robotics manipulators in flexible manufacturing systems presents a serious challenge in production. Locating sensors on the manipulator itself, or near the end effector, provides a satisfactory solution to the position of sensors within the FMS. Locating an image sensor above the work area of a robot may cause the manipulator to obscure its own work area. Measurement of the displacement of the end effector also may suffer distortion, since the destination will be measured in relative terms, not absolute. Placing the sensor on the end effector allows absolute measurement to be taken, reducing considerably the need for calibration of mechanical position and for imaging linearity. Image sensory feedback in this situation can be reduced to the simplicity of range finding in some applications.

Extensive research and development activities were conducted recently to find ways to integrate various sensors close to the gripper jaws of robots. The promise of solid-state arrays for this particular application has not entirely materialized, primarily because of diversion of effort resulting from the commercial incentives associated with the television industry. It might be accurate to predict that, over the next decade, imaging devices manufactured primarily for the television market will be both small and affordable enough to be useful for robotics applications. However, at present, array cameras are expensive and, while smaller than most thermionic tube cameras, are far too large to be installed in the region of a gripper. Most of the early prototype arrays of modest resolution (developed during the mid-1970s) have been abandoned.

Some researchers have attacked the problem of size reduction by using coherent fiber optics to retrieve the image from the gripper array, which imposes a cost penalty on the total system. This approach can, however, exploit a fundamental property of optical fiber in that a bundle of coherent fibers can be subdivided to allow a single high-resolution imaging device to be used to retrieve and combine a number of lower-resolution images from various paths of the work area, including the gripper, with each subdivided bundle associated with its own optical arrangement.

Linear arrays have been used for parts moving on a conveyer in such a way that mechanical motion is used to generate one axis of a two-dimensional image. The same technique can be applied to a robot manipulator by using the motion of the end effector to generate a two-dimensional image.

Tactile sensing is required in situations involving placement. Both active and passive compliant sensors have been successfully applied in the field. This is not the case for tactile array sensors because they are essentially discrete in design, are inevitably cumbersome, and have very low resolution.

Acoustic sensors, optical sensors, and laser sensors are well developed for effective use in manufacturing applications. Although laser range-finding sensors are well developed, they are significantly underused in FMS, especially in robotic applications. Laser probes placed at the end effector of an industrial robot will form a natural automated inspection system in manufacturing.

Sensing for robot applications does not depend on a relentless pursuit for devices with higher resolution; rather, the fundamental consideration is selecting the optimum resolution for the task to be executed. There is a tendency to assume that the higher the resolution, greater the application range for the system. However, considerable success has been achieved with a resolution as low as 50 × 50 picture elements. With serial processing architectures, this resolution will generate sufficient gray-scale data to test the ingenuity of image processing algorithms. Should its processing time fall below 0.5 s, an algorithm can be used for robots associated with handling. However, in welding applications, the image processing time must be faster.

ROBOT CONTROL THROUGH VISION SENSORS

An increasing number of manufacturing processes rely on machine-vision sensors for automation. The tasks for which vision sensors are used vary widely in scope and difficulty. Robotic applications in which vision sensing has been used successfully include inspection, alignment, object identification, and character recognition.

Human vision involves transformation, analysis, and interpretation of images. Machine-vision sensing can be explained in terms of the same functions: image transformation, image analysis, and image interpretation.

Image Transformation

Image transformation involves acquiring camera images and converting them to electrical signals that can be used by a vision computer (Fig. 7.1). After a camera image is transformed into an electronic (digitized) image, it can be analyzed to extract useful information in the image such as object edges, alignment, regions, boundaries, colors, and absence or presence of vital components.

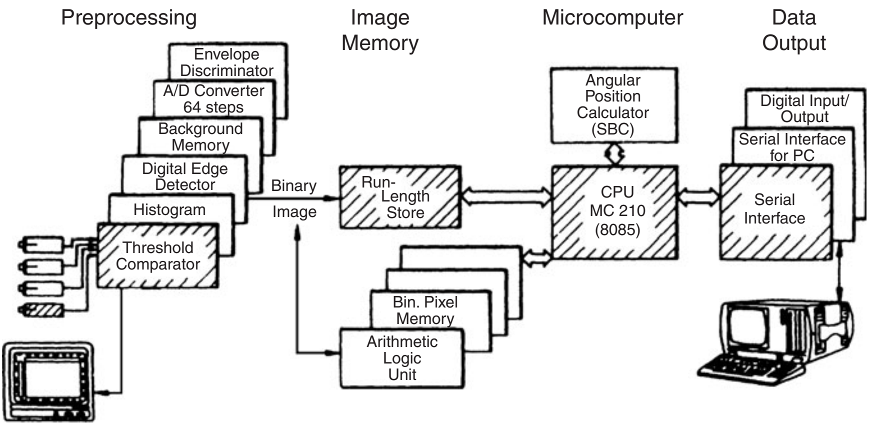


FIGURE 7.1 Image transformation involves the acquisition and conversion of camera images to electrical signals that can be used by a vision computer.

Once the image is analyzed, the vision sensing system can interpret what the image represents so the robot can continue its task. In robot vision execution, design considerations entail cost, speed, accuracy, and reliability.

Robot Vision and Human Vision

Given that robot vision systems typically execute only part of what is normally called *seeing*, some similarities with human vision nevertheless arise. Other similarities arise in the "hardware" of human and robot visual systems.

The similarities in hardware could include an analogy between eyes and video cameras—both have lenses to focus an image on a sensitive "retina" that produces a visual signal interpreted elsewhere. In both human and robot vision, this signal is passed to a device that can remember important aspects of the image for a time, perform specialized image processing functions to extract important information from the raw visual signal, and analyze the image in a more general way.

Some similarities in performance follow. Human and robot vision work well only where lighting is good. Both can be confused by shadows, glare, and cryptic color patterns. Both combine size and distance judgments, tending to underestimate size when they underestimate distance, for instance.

However, humans and machines have far more differences than similarities. A human retina contains several million receptors, constantly sending visual signals to the brain. Even the more limited video camera gathers over 7 Mbytes of visual information per second. Many of the surprising aspects of machine vision arise from the need to reduce this massive flow of data so it can be analyzed by a computer system.

Machine-vision systems normally only detect, identify, and locate objects, ignoring many of the other visual functions. However, they perform this restricted set of functions very well, locating and even measuring objects in a field of view more accurately than any human can.

Robot Vision and Visual Tasks

Several standard visual tasks are performed by robot-vision systems. These tasks include recognizing when certain objects are in the field of view, determining the location of visible objects, assisting a robot hand with pickup and placement, and inspecting known objects for the presence of certain characteristics (usually specific flaws in manufacture) (Fig. 7.2).

A robot-vision system must exercise some judgment in performing *visual tasks*—those for which the input is a visual image (normally obtained from an ordinary video camera). Which visual tasks are relatively easy for machine vision, and which are hard? The distinction is not so much that some tasks are hard and others easy; rather, it is the detail within a task that distinguishes easy problems from hard ones.

Contributing factors affecting problem complexity:

- Objects that vary widely in detail. (Examining stamped or milled product may be easy, while molded or sculpted items may be more difficult. Natural objects are by far the hardest with which to deal.)
- Lighting variations, including reflections and shadows, as well as fluctuations in brightness (as found in natural sunlight). These variations may be unnoticed by human inspectors, however, they can make otherwise easy problems difficult or impossible for robot vision.
- In general, ignoring "unimportant" variations in an image while responding to "significant" ones is very hard.

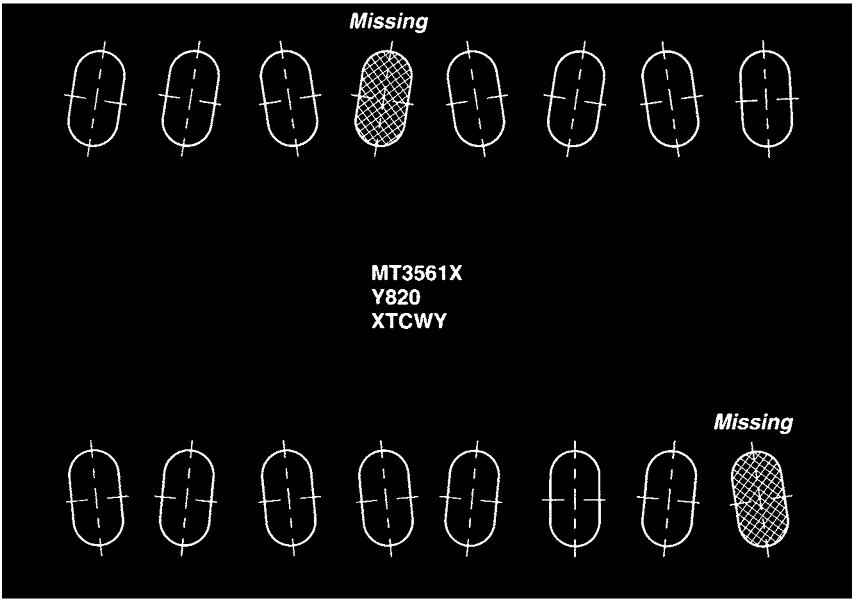


FIGURE 7.2 Image of a flaw.

Robot Visual Sensing Tasks

Robots will work in unpleasant locations. Health hazards are of no concern to them. Special safety equipment is not required for a robot spraying paint, welding, or handling radioactive materials or chemicals. All this adds up to reduced production costs. As the day progresses, the tired worker has a tendency to pay less attention to details and the quality of the finished product may suffer. This is especially noticeable in automobiles where spray paint can run or sag and weld joints may not be made perfectly. The panels of the car may not be aligned, and the finished product may not operate properly, with predictable customer dissatisfaction. In pharmaceutical production, too, an operator inspecting and verifying lot number and expiration date may become fatigued and fail to inspect for sensitive information.

Robots, on the other hand, do not tire or change their work habits unless programmed do so. They maintain the same level of operation throughout the day. With vision-sensing systems and robots, it is possible for American manufacturers to compete against lower labor costs in foreign countries. The initial investment is the only problem. After the initial investment, the overall operation costs of the production line are reduced or held constant. Small educational robots can be used to retrain humans to operate and maintain robots.

The roles of robots with machine vision can be summarized as follows:

- *Handling and assembly.* Recognizing position/orientation of objects to be handled or assembled, determining presence or absence of parts, and detecting parts not meeting required specifications
- *Part classification.* Identifying objects and recognizing characters

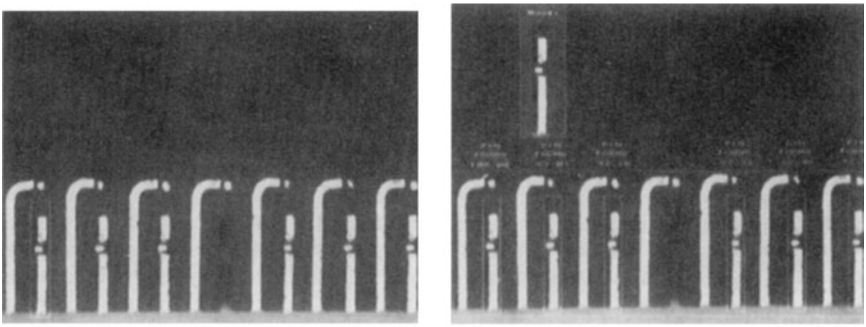
- *Inspection.* Checking for assembly and processing, surface defects, and dimensions
- *Fabrication.* Making investment castings, grinding, deburring, water-jet cutting, assembling wire harnesses, gluing, sealing, puttying, drilling, fitting, and routing
- *Welding.* Automobiles, furniture, and steel structures
- *Spray painting.* Automobiles, furniture, and other objects

Robots Utilizing Vision Systems to Recognize Objects

A basic use of vision is recognizing familiar objects. It is easy to see that this ability should be an important one for robot vision. It can be a task in itself, as in counting the number of each kind of bolt in a mixed lot on a conveyer belt. It can also be an adjunct to other tasks—for example, recognizing a particular object before trying to locate it precisely, or before inspecting it for defects.

It is important to note that this task actually has two distinct parts: first, object familiarization—that is, learning what an object looks like; and then object recognition (Fig. 7.3).

Learning to recognize objects can be done in many ways. Humans can learn from verbal descriptions of the objects, or they can be shown one or more typical items. A brief description of a *pencil* is enough to help someone identify many unfamiliar items as *pencil*. Shown a few *pencils*, humans can recognize different types of *pencils*, whether they look exactly like the samples or not.



(a) (b)

FIGURE 7.3 Recognition of a learned object.

Robot-vision systems are not so powerful, but both these approaches to training still apply to them. Robots can be given descriptions of what they are to recognize, perhaps derived from CAD data to guide a machine tool (Fig. 7.4) or they can be shown samples, then be expected to recognize objects more or less like the samples.

Recognizing objects once they have been learned is the second, and more difficult, part of the task. Several basic questions arise in virtually every recognition task. Among them are "What are the choices?" and "What can change?"

Specifying the actual task completely is normally the hardest part of the application. When this has been accomplished, the basic question is, for what particular vision features should the search begin to recognize an object, and, when the features are found, how should they be analyzed to decide which object (if any) has been found?

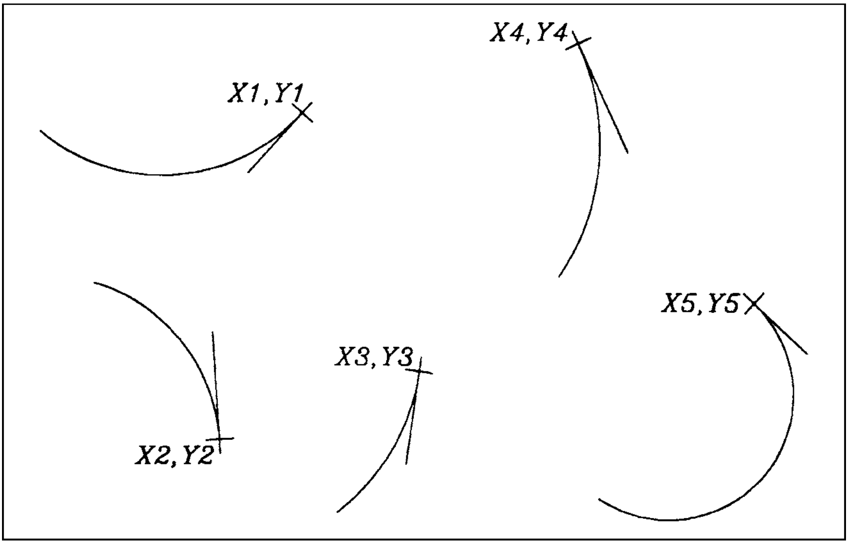


FIGURE 7.4 Machine guidance.

ROBOT VISION LOCATING POSITION

Humans use several techniques for gauging distances, especially triangulation on the left- and right-eye views, feedback from the eye's focusing mechanism, and the apparent motion produced by small head movements. This kind of object location may make use of knowledge about the object being located. By knowing how large the object really is, one can judge how far away it is from the size of its retinal image.

Few robot-vision systems use binocular vision, autofocus feedback, or moving cameras to estimate distances. However, with a rigidly mounted camera, it is possible to interpret each visible point as lying on a particular line of sight from the camera. Accurate information about the true distances between visible points on an object allows the robot-vision system to accurately calculate its distance. Similarly, if an object is resting on a platform at an accurately known distance, the robot vision system can interpret distances in the camera image as accurate distances on the object.

However locations are determined, the visual location task for a robot-vision system normally includes calibration as well as object location. Calibration normally is carried out by providing the system with a view that has easily identifiable points at known spatial locations. Once calibrated, the system can then locate objects in its own coordinate system (pixels) and translate the position into work cell coordinates (inches, millimeters, and so on).

ROBOT GUIDANCE WITH VISION SYSTEM

Another use of machine vision is in robot guidance—helping a robot handle and place parts and providing it with the visual configuration of an assembly after successive tasks. This can involve a series of identification and location tasks. The camera can be attached to a mobile arm, making the location task seem somewhat more like normal vision. However, the camera typically is mounted on a fixed location to reduce system complexity.

While each image can give the location of certain features with respect to the camera, this information must be combined with information about the current location and orientation of the camera to give an absolute location of the object. However, the ability to move the camera for a second look at an object allows the unambiguous location of visible features by triangulation. Recognition is a useful tool for flexible manufacturing systems within a CIM environment. Any of several parts may be presented to a station where a vision system determines the type of part and its exact location. While it may be economical and simpler to send the robot a signal giving the part type when it arrives, the ability to detect what is actually present at the assembly point, not just what is supposed to be there, is of real value for totally flexible manufacturing.

Robot Vision Performing Inspection Tasks

Visual inspection can mean any of a wide variety of tasks, many of which can be successfully automated. Successful inspection tasks are those in which a small number of reliable visual cues (features) are to be checked and a relatively simple procedure is used to make the required evaluation from those cues.

FIGURE 7.5 Presence or absence of items in an assembly.

The differences between human and robot capabilities are most evident in this kind of task, where the requirement is not simply to distinguish between good parts and anything else—a hard enough task—but usually to distinguish between good parts or parts with harmless blemishes, and bad parts. Nevertheless, when inspection can be done by robot vision, it can be done very predictably.

Many inspection tasks are well suited to robot vision. A robot-vision system can dependably determine the presence or absence of particular items in an assembly (Fig. 7.5), providing accurate information on each of them. It can quickly gauge the approximate area of each item passing before it, as long as the item appears somewhere in its field of view (Fig. 7.6).

Components of Robot Vision

Figure 7.7 is a schematic diagram of the main components of a robot in a typical vision process for manufacturing. A fixed camera surveys a small, carefully lighted area where the objects to be located or inspected are placed. When visual information is needed (as signaled by some external switch or sensors), a digitizer in a robot vision system converts the camera image into a "snapshot": a significant array of integer brightness values (called gray levels). This array is sorted in a large random-access memory (RAM) array in the robot-vision system called an *image buffer* or a *frame buffer*. Once sorted, the image can be displayed on a monitor at any time. More importantly, the image can be analyzed or manipulated by a vision computer, which can be programmed to solve robot vision problems. The vision computer is often connected to a separate general-purpose (host) computer, which can be used to load programs or perform tasks not directly related to vision.

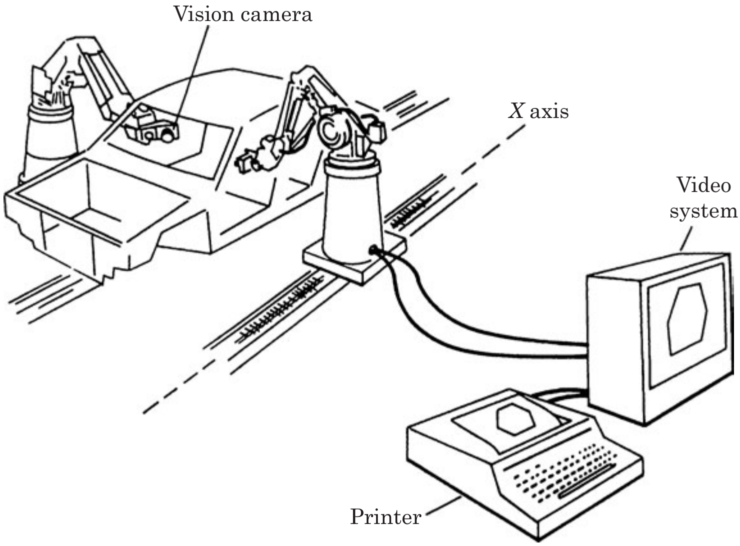


FIGURE 7.6 Object in field of view.

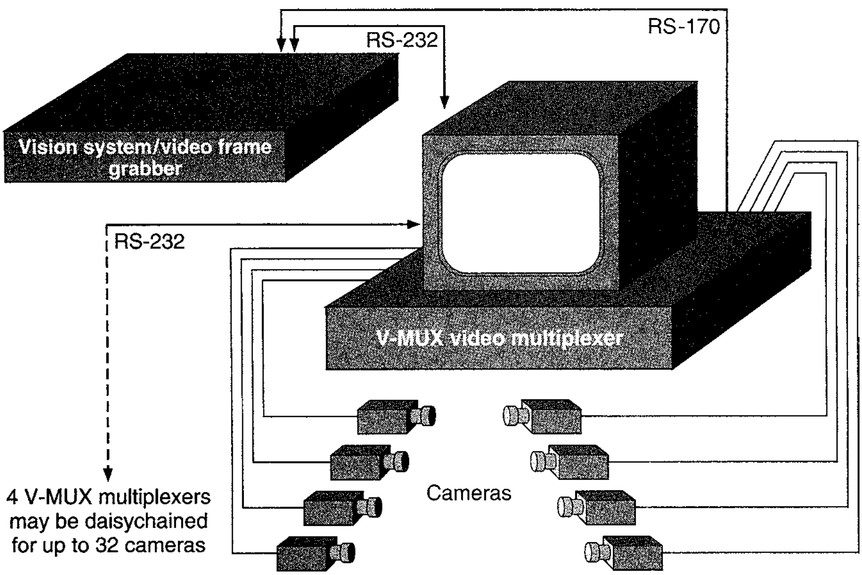


FIGURE 7.7 Schematic diagram of typical vision process.

Once an image is acquired, vision processing operations follow a systematic path. Portions of the image buffer may first be manipulated to suppress information that will not be valuable to the task at hand and to enhance the information that will be useful. Next, the vision program extracts a small number of cues from the image—perhaps allowing the region of interest to be reduced to exclude even more extraneous data.

At this stage, the vision program calculates, from the selected image region, the cues (features) of direct importance to the task at hand and makes a decision about the presence of a known part or its location in the field of view, or perhaps about the presence of specific defects in the object being inspected. Finally, the robot-vision system activates control lines based on the decisions, and (perhaps) transfers a summary of the conclusion to a data storage device or another computer.

END EFFECTOR CAMERA SENSOR FOR EDGE DETECTION AND EXTRACTION

A considerable amount of development of synchronized dual camera sensors at a strategic location on a robot end effector has been conducted for processing two-dimensional images stored as binary matrices. A large part of this work has been directed toward solving problems of character recognition. While many of these techniques are potentially useful in the present context, it is valuable to note some important differences between the requirements of character recognition and those associated with visual feedback for mechanical assembly.

Shape and Size

All objects presented to the assembly machine are assumed to match an exact template of the reference object. The object may have an arbitrary geometric shape, and the number of possible different objects is essentially unlimited. Any deviation in shape or size, allowing for errors introduced by the visual input system, is grounds for rejection of the object (though this does not imply the intention to perform 100 percent inspection of components). The derived description must therefore contain all the shape and size information originally presented as a stored image. A character recognition system must, in general, tolerate considerable distortion, or style, in the characters to be recognized, the most extreme example being handwritten characters. The basic set of characters, however, is limited. The closest approach to a template-matching situation is achieved with the use of a type font specially designed for machine reading, such as optical character recognition (Fig. 7.8).

Position and Orientation

A component may be presented to the assembly machine in any orientation and any position in the field of view. Though a position- and orientation-invariant description is required in order to recognize the component, the measurement of these parameters is also an important function of the visual system to enable subsequent manipulation. While a line character may sometimes be skewed or bowed, individual characters are normally presented to the recognition system in a relatively constrained orientation, a measurement of which not required.

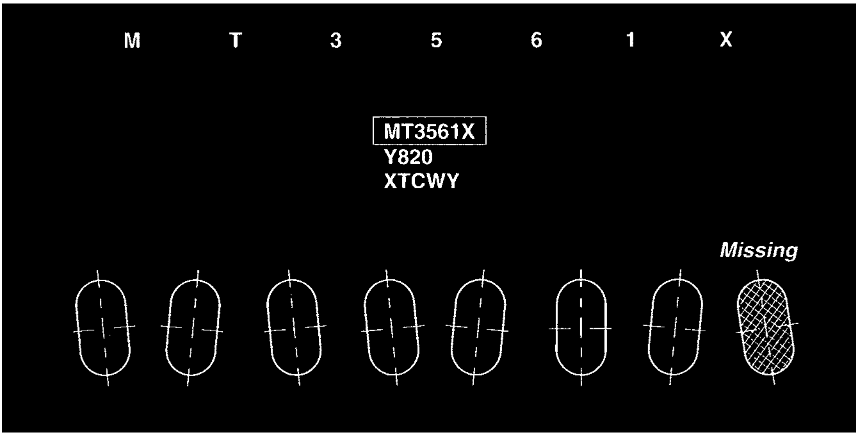


FIGURE 7.8 Optical character recognition.

Multiple Objects

It is a natural requirement that the visual system for an assembly machine should be able to accommodate a number of components randomly positioned in the field of view. The corresponding problem of segmentation in character recognition is eased (for printed characters) by *a priori* knowledge of character size and pitch. Such information has fostered techniques for the segmentation of touching characters. No attempt is made to distinguish between touching objects. Their combined image will be treated by the identification procedures as that of a single, supposedly unknown object.

The essentially unlimited sizes of the set of objects that must be accommodated by the recognition system demands that a detailed description of shapes be extracted for each image. However, a number of basic parameters may be derived from an arbitrary shape to provide valuable classification and position information. These include:

- Area
- Perimeter
- Minimum enclosing rectangle
- Center of the area
- Minimum radius vector (length and direction)
- Maximum radius vector (length and direction)
- Holes (number, size, position)

Measurements of area and perimeter provide simple classification criteria that are both position- and orientation-invariant. The dimensionless shape factor area/perimeter has been used as a parameter in object recognition. The coordinates of the minimum enclosing rectangle provide some information about the size and shape of the object, but this information is orientation-dependent. The center of the area is a point that may be readily determined for any object, independent of orientation, and is thus of considerable importance for recognition and location purposes. It provides the origin for the radius vector, defined as a line in the center of the area to a point on the edge of an object. The radius vectors of maximum and minimum length are potentially useful parameters for determining both identification and orientation. Holes are common features of engineering components, and the number present in a part is a further suitable parameter. The holes themselves may also be treated as objects, having shape, size, and position relative to the object in which they are found.

The requirements for the establishment of connectivity in the image and the derivation of detailed descriptions of arbitrary geometric shapes are most appropriately met by an edge-following technique. The technique starts with the location of an arbitrary point on the black/white edge of an object in the image (usually by a raster scan). An algorithm is then applied that locates successive connected points on the edge until the complete circumference has been traced and the starting point is reached. If the direction of each edge point relative to the previous point is recorded, a one-dimensional description of the object is built up that contains all the information present in the original shape. Such chains of directions have been extensively studied by Freeman. Measurements of area, perimeter, center of area, and enclosing rectangle may be produced while the edge is being traced, and the resulting edge description is in a form convenient for the calculation of radius vectors.

Edge following establishes connectivity for the object being traced. Continuing the raster scan in search of further objects in the stored image then presents the problem of the rediscovery of an already traced edge.

A computer plot of the contents of the frame with the camera viewing a square and a disk is illustrated in Fig. 7.9, and the result of applying the edge-extracting operation is illustrated in Fig. 7.10. The edge-following procedure may now be applied to the image in the same way it was to the solid object. The procedure is arranged, however, to reset each edge point as it is traced. The tracing of a complete object thus removes it from the frame and ensures it will not be subsequently retraced.

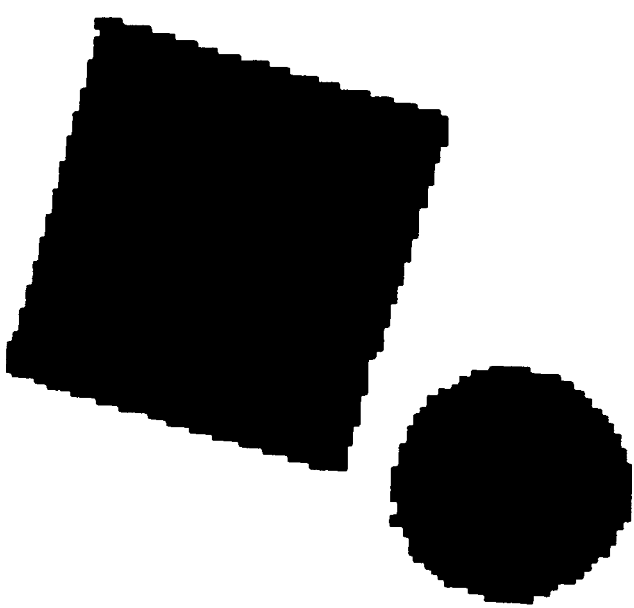


FIGURE 7.9 Computer plot of the contents of a frame.

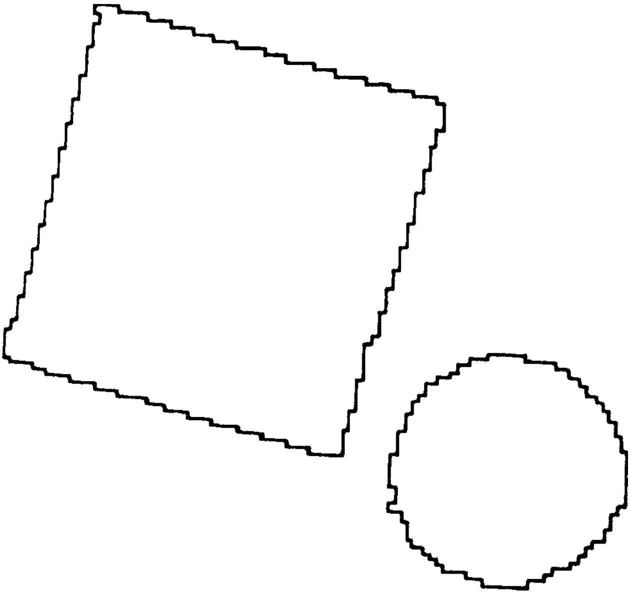


FIGURE 7.10 Result of applying an edge extraction operation.

END EFFECTOR CAMERA SENSOR DETECTING PARTIALLY VISIBLE OBJECTS

A new method of locating partially visible two-dimensional objects has been developed. The method is applicable to complex industrial parts that may contain several occurrences of local features, such as holes and corners. The matching process utilizes clusters of mutually consistent features to hypothesize objects and uses templates of the objects to verify these hypotheses. The technique is fast because it concentrates on key features that are automatically selected on the basis of the detailed analysis of CAD-type models of the objects. The automatic analysis applies general-purpose routines for building and analyzing representations of clusters of local features that could be used in procedures to select features for other locational strategies. These routines include algorithms to compute the rotational and mirror symmetries of objects in terms of their local features. The class of tasks that involve the location of the partially visible object ranges from relatively easy tasks, such as locating a single two-dimensional object, to the extremely difficult task of locating three-dimensional objects jumbled together in a pallet. In two-dimensional tasks, the uncertainty is in the location of an object in a plane parallel to the image plane of the camera sensor. This restriction implies, on one hand, a simple one-to-one correspondence between sizes and orientations in the image, and on the other hand, the sizes and orientations in the plane of the object.

This class of two-dimensional tasks can be partitioned into four subclasses that are defined in terms of the complexity of the scene:

1. A portion of one of the objects
2. Two or more objects that may touch one another
3. Two or more objects that may overlap one another
4. One or more objects that may be defective

This list is ordered roughly by the increasing amount of effort required to recognize and locate the object.

Figure 7.11 illustrates a portion of an aircraft frame member. A typical task might be to locate the pattern of holes for mounting purposes. Since only one frame member is visible at a time, each feature appears at most once, which simplifies feature identification. If several objects can be in view simultaneously and can touch one another, as in Fig. 7.12, the

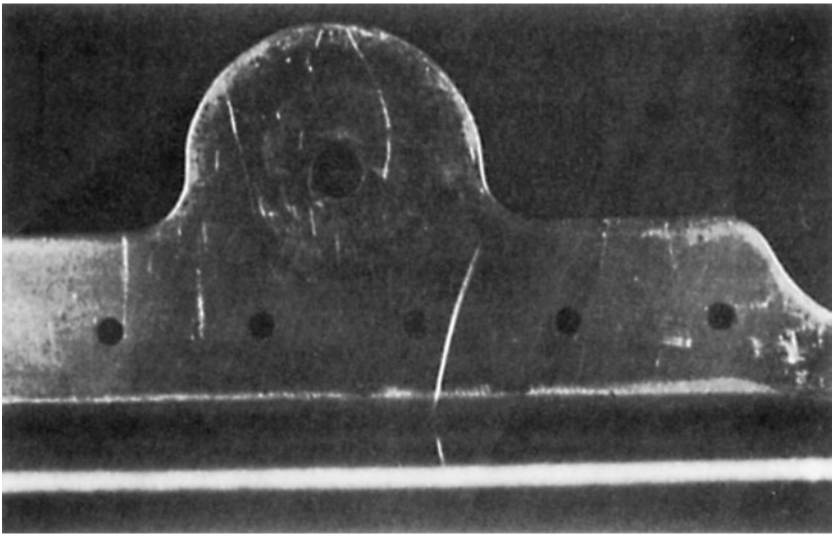


FIGURE 7.11 Portion of an aircraft frame member.

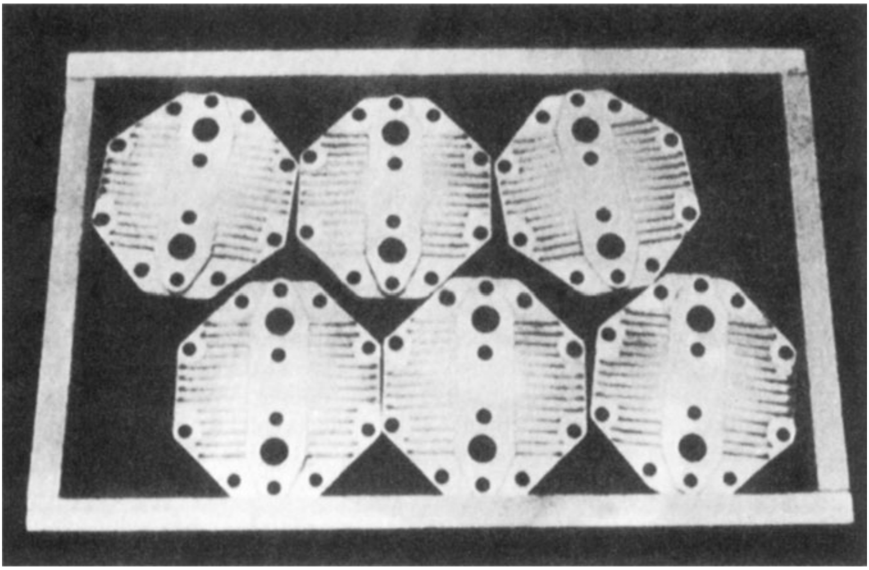


FIGURE 7.12 Objects touching each other.

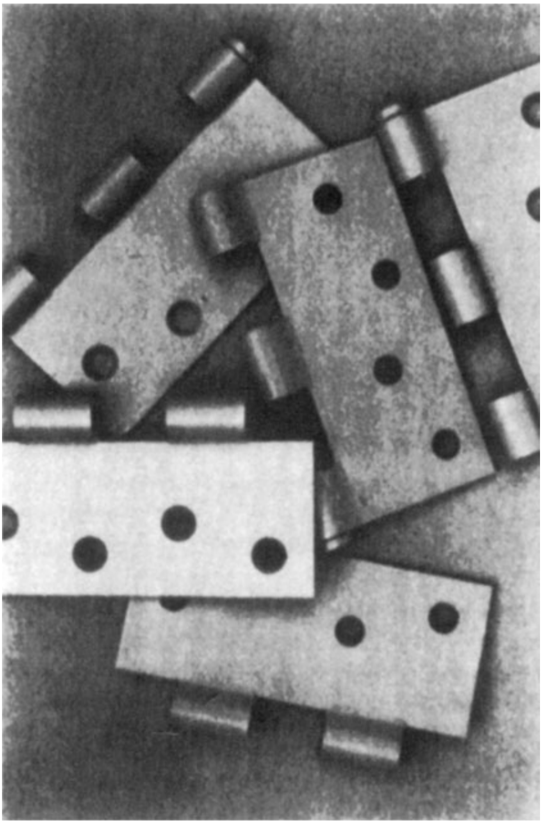


FIGURE 7.13 Objects lying on top of each other.

features may appear several times. Boundary features such as corners may not be recognizable, even though they are in the picture, because the objects are in mutual contact. If the objects can lie on one another (Fig. 7.13), even some of the internal holes may be unrecognizable because they are partially or completely occluded. And, finally, if the objects are defective (Fig. 7.14), the features are even less predictable and hence harder to find.

Since global features are not computable from a partial view of an object, recognition systems for these more complex tasks are forced to work with either local features, such as small holes and corners, or extended features, such as a large segment of an object's boundary. Both types of feature, when found, provide constraints on the position and the orientations of their objects. Extended features are in general computationally more expensive to find, but they provide more information because they tend to be less ambiguous and more precisely located.

Given a description of an object in terms of its features, the time required to match this description with a set of observed features appears to increase exponentially with the number of features. The multiplicity of features precludes the straightforward application of any simple matching technique. Large numbers of features have been identified by locating a few extended features instead of many local ones. Even though it costs more to locate extended features, the reduction in the combinatorial explosion is often worth it. The other

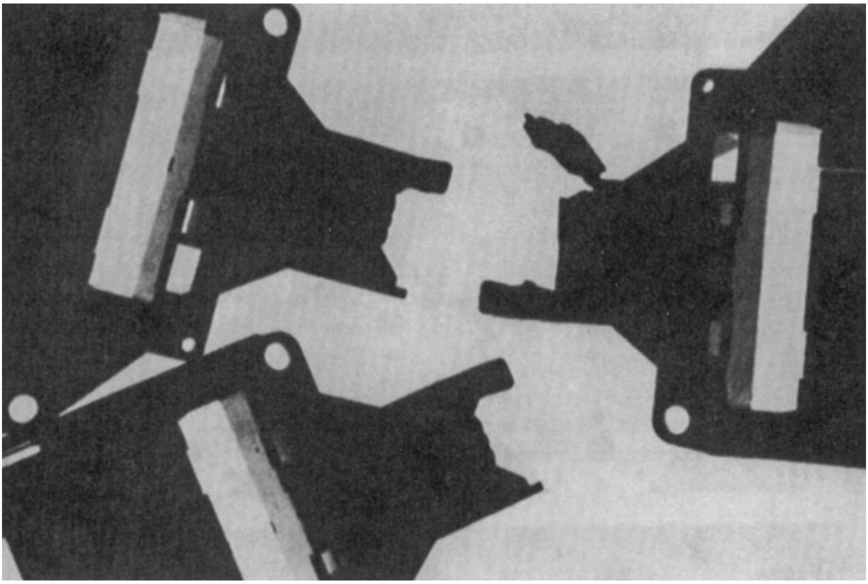


FIGURE 7.14 A trained image.

approach is to start by locating just one feature and using it to restrict the search area for nearby features. Concentrating on one feature may be risky, but the reduction in the total number of features to be considered is often worth it. Another approach is to sidestep the problem by hypothesizing massively parallel computers that can perform matching in linear time. Examples of these approaches include graph matching, relaxation, and histogram analysis. The advantage of these applications is that the decision is based on all the available information at hand.

The basic principle of the local-feature-focus (LFF) method is to find one feature of an image, referred to as the *focus feature*, and use it to predict a few nearby features to look for. After finding some nearby features, the program uses a graph-matching technique to identify the largest cluster of image features matching a cluster of object features. Since the list of possible object features has been reduced to those near the focus feature, the graph is relatively small and can be analyzed efficiently.

The key to the LFF method is an automatic feature-selection procedure that chooses the best focus features and the most useful sets of nearby features. This automatic-programming capability makes possible quick and inexpensive application of the LFF method to new objects. As illustrated in Fig. 7.15, the training process, which includes the selection of features, is performed once and the results are used repeatedly.

Run-Time Phase

The run-time phase of the LFF acquires images of partially visible objects and determines their identities, positions, and orientations. This processing occurs in four steps:

1. Reading task information
2. Locating local features

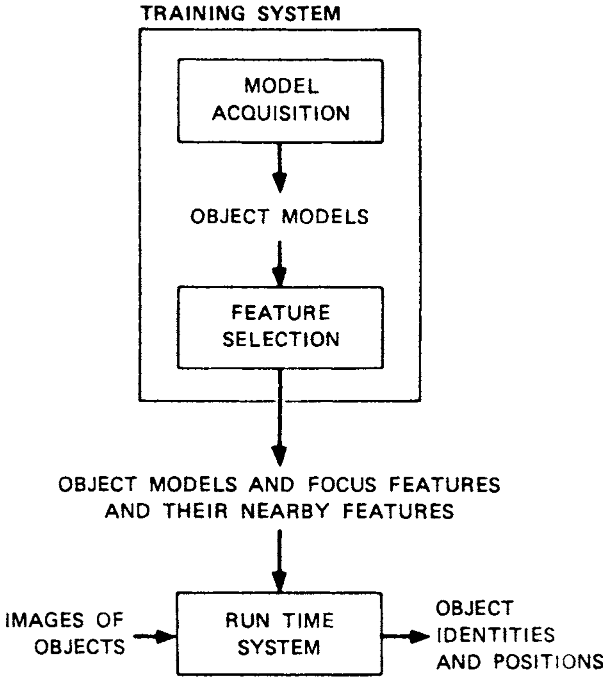


FIGURE 7.15 A run-time phase procedure.

3. Hypothesizing objects
4. Verifying hypotheses

The procedure (Fig. 7.15) is to input the object model together with the list of focus features and their nearby cofeatures. Then, for each image, the system locates all potentially useful local features, forms clusters of them to hypothesize object occurrences, and finally performs template matching to verify these hypotheses.

ULTRASONIC END EFFECTORS

An end effector on a welding robot (Fig. 7.16) contains an ultrasonic sensor for inspection of the weld. An ultrasonic sensor detects such flaws as tungsten inclusions and lack of penetration of a weld. The end effector determines the quality of a weld immediately after the weld contact has been made, while the workpiece is still mounted on the weld apparatus; a weld can be reworked in place, if necessary. The delay caused by the paperwork and setup involved in returning the workpiece for rework is thereby avoided.

The ultrasonic end effector can be mounted on any standard gas tungsten arc welding torch. It may also be equipped with a through-the-torch vision system. The size of the ultrasonic end effector is the same as that of a gas cup with cathode.

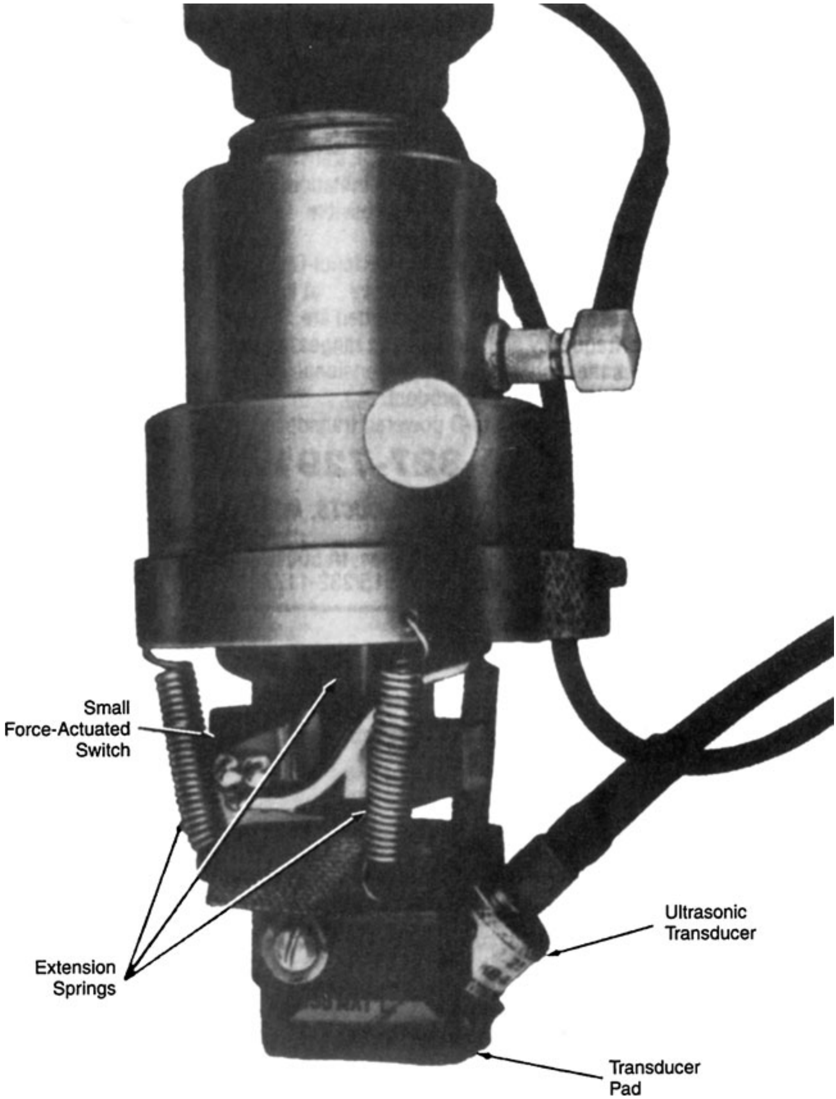


FIGURE 7.16 An end effector on a welding robot.

A set of extension springs stabilizes the sensor and ensures its elastomeric dry-couplant pad fits squarely in the weldment surface. The sensor can be rotated 360° and locked into alignment with the weld lead. A small force-actuated switch halts downward travel of the robot arm toward the workpiece and sets the force of contact between the sensor and the workpiece.

END EFFECTOR SOUND-VISION RECOGNITION SENSORS

The sound recognition sensor consists of a source that emits sound waves to an object and a sound receiver that receives the reflected sound waves from the same object (Fig. 7.17). The sound recognition sensor array consists of one sound source and from 1 to 16 receivers fitted intricately on an end effector of a robot.

The sound-vision recognition sensor array measures reflections from some surface of interest on the object, called the *measured surface*, which is perpendicular to the sound waves emitted from the sound source (Fig. 7.18). Four conditions govern the performance of sound-vision sensors:

- Standoff
- Large surfaces
- Small surfaces
- Positioning

Standoff

Standoff is how far the array must be located from the measured surface. The standoff, like other measurements, is based on the wavelength of the sound used. Three different

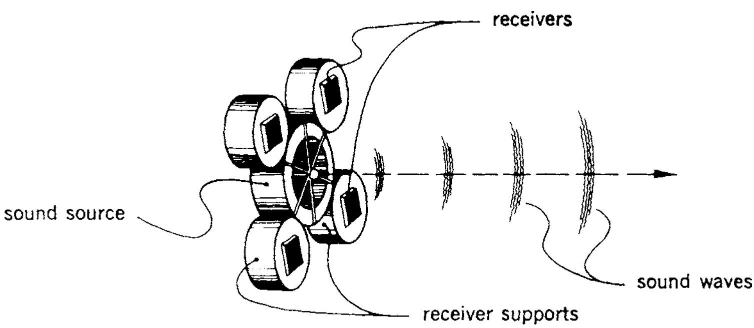


FIGURE 7.17 An end effector sound-vision recognition system.

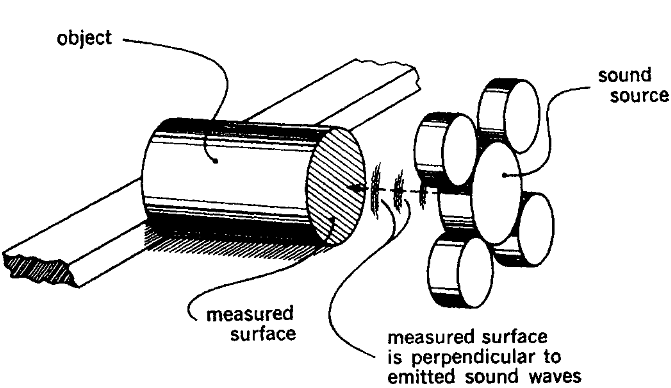


FIGURE 7.18 The measured surface should be perpendicular to the sound waves emitted from the sound source.