5

Robot Sensors

> Robots are far from being at the end of their evolutionary tether. In the near future they will be made adaptable to a greater variety of manual tasks. The trick is to give robots at least one human sense, rudimentary eyesight. An otherwise puerile job is baffling to a robot if parts are not oriented at the pickup point.
>
> Joseph F. Engelberger (1980)

Almost all industrial robots use internal sensors, such as shaft encoders to measure rotary joint position and tachometers, which measure velocity to control their motions. Most controllers also provide interface capabilities so that signals from conveyors, machine tools, and the robot itself may be used to accomplish a task. However, external sensors, such as visual sensors, can provide a much greater degree of adaptability for higher level robot control as well as add automatic inspection capabilities to the industrial robot. Visual and other sensors are now used in such fundamental operations as material processing with immediate inspection, material-handling with adaptation, arc welding, and complex assembly tasks. A new industry of robot vision companies has emerged. This chapter provides an introduction to this new, important technology.

To understand what sensors might be desired or needed for robots to optimally perform tasks, it is useful to first review the human senses to see how they are used as we live and work. In so doing, we can appreciate the complexity of the human and perceive some of the problems encountered in equipping robots with simulated senses. We are all familiar with five of the seven senses of the human: sight, smell, taste, hearing, and touch. These give us an incredible amount of information about our environment that enables us to make decisions about how to adapt to new or unexpected situations.

Let's start with sight. One of the most important things sight allows us to do is to select proper, safe paths for motion. Binocular vision and other perceptual cues, such as object occlusion, permit us to judge the distance of objects. Color vision permits rapid discrimination of millions of different shades of light and color. The spatial acuity provided by high-density rod and cone receptors allows us to perceive a minute speck of dust in optimum lighting. Our brightness sensitivity is so acute that we can see the light of a single candle at a distance of 30 miles on a very dark night. Automatic brightness control permits rapid adjustment from very light to very dark environments. We depend on our eyes to give us most of the information we need. It is estimated that 70 percent of the information that reaches the brain comes through our visual sense. An interesting experiment (Buffington, 1983) illustrates the dominance of vision in our understanding of the world about us. It is also easy to perform. Close your eyes, and trace the number 2 with your finger on your forehead. Does it feel reversed? Now trace the same number on the back of your head. Does it feel normal? This experiment shows that, just because our eyes face forward and dominate our senses, other senses, such as tactility, are important for correctly performing tasks.

Hearing is, like vision, stereoscopic, and permits us to judge the direction and distance of a sound. This sense is well developed even before birth and works best when we are asleep. It is so acute that many mothers can hear the breathing of a newborn infant in another room. We also use hearing to select proper forms of motion, especially when visual cues are missing or obstructed, as when we hear a car coming before we see it. We also use our hearing in making decisions. For example, many very experienced automobile mechanics can simply listen to an engine running and correctly identify any problems. We are able to distinguish many different tones and wavelengths, which enables us to distinguish and identify millions of objects and phenomena in our world.

Smell is a chemical sense, olfaction, as is taste, gustation. Our olfactory senses enable us to distinguish many objects and phenomena without the use of any other senses. For example, we can usually tell what foods are edible or ripe by smelling them, even though our visual cues may indicate otherwise. Olfaction is particularly important in enabling us to identify invisible or hidden elements, such as gases. Taste is also important in determining the potability of food. The four taste qualities—bitter, sour, salty, and sweet—help provide us with the impetus to obtain essential nutrients. Many people would say their gustatory senses are too well developed. However, we use our taste sensors to again discriminate and distinguish many objects. For instance, we can detect the presence of minute amounts of metal in our foods (which is why you don't use metal utensils to cook foods with delicate flavors), as well as the presence of some gases and chemicals that are neither seen nor smelled.

Touch includes more sensitivity that we often think about. Sensors for pressure, temperature, and pain are embedded in our skin by the thousands. For example, there are about 3,000,000 pain sensors, 500,000 pressure sensors, and 200,000 temperature sensors distributed unevenly throughout the human body, mainly on the surface. For example, there are 232 pain sensors per square centimeter behind the knee, 60 per square centimeter on the thumb pad, and 44 per square centimeter on the tip of the nose. Again, we can use this sense to help us identify and distinguish objects and phenomena when our other senses fail or are obstructed. For instance, we can feel and identify a caterpillar crawling on our backs and take appropriate action without using our other senses.

These are the five commonly known senses, but we have two others that are also very important and of which we are not usually consciously aware. Some sense receptors located in the tendons, joints, and muscles inform the brain of the position and movements of the entire body. This kinesthetic sense permits us to walk without watching our legs, to tense a muscle without looking at it, or to touch our finger to our nose with our eyes shut. The ability to walk uses this kinesthetic sense, but it also requires a sense of balance. This balance is provided by the vestibular sense, located mainly in the inner ear, informing us whether we are upside down or right side up, speeding up or slowing down, or rising or falling. It is the manipulation of this sense that gives us such a thrill on a roller coaster ride.

There may be other senses humans possess, such as a magnetic sense, but these are still in some dispute. These seven certainly seem sufficient to allow us to adapt to our environment and perform our work. However, it is characteristic of humans to want to improve on even these extremely versatile senses with inventions that extend or amplify them. These extensions include microscopes, telescopes, sonar, radar, x-ray, and infrared devices or detectors, and Geiger counters. They give us far more information about our world than we could receive using our unaided senses. Equipping intelligent robots with sensors includes considering not only our own senses, but these extensions, as well, in enabling them to perform certain tasks. Sensing techniques, such as the ability to see in all directions at once, could be built into a machine. Parents often seem to have eyes in the back of their heads. The robot actually could. Similarly, sensors that measure force, torque, proximity, temperature, and magnetic or electric field properties could be implemented to measure these quantities at isolated points or at several neighboring points to permit recognition of an object. For the universal machine, all these senses may be needed.

5.1 Robot Sensor Classification

Sensors may first be classified as internal or external, in which the external sensors, such as vision or tactility, are not included as inherent components of the robot controller, but the internal sensors, such as shaft encoders, are built into the manipulator. This classification is indicative of the early stage of the evolution of robots. Eventually, it is likely that all useful sensors will be integrated into the robot design.

Another classification is based upon the function performed by the sensor. Although many types of sensors for robots are still being invented, there are some generic groupings of the different types, as shown in Figure 5–1. For the acquisition stage, in which information is being collected and the robot manipulator is not in contact with the part, a noncontact sensor must be used. This assumes that the only sensors are those on the robot. External sensors, such as a separate tactile table, could be used to sense shapes. Of the various types of noncontact sensors, there is a natural division into those that measure a point response and others that give a spatial array or measurements at neighboring points of information. An example of a point-measuring device may be found in the distance-measuring ultrasound devices developed by Polaroid. These devices measure the distance to the nearest object within a cone of information-

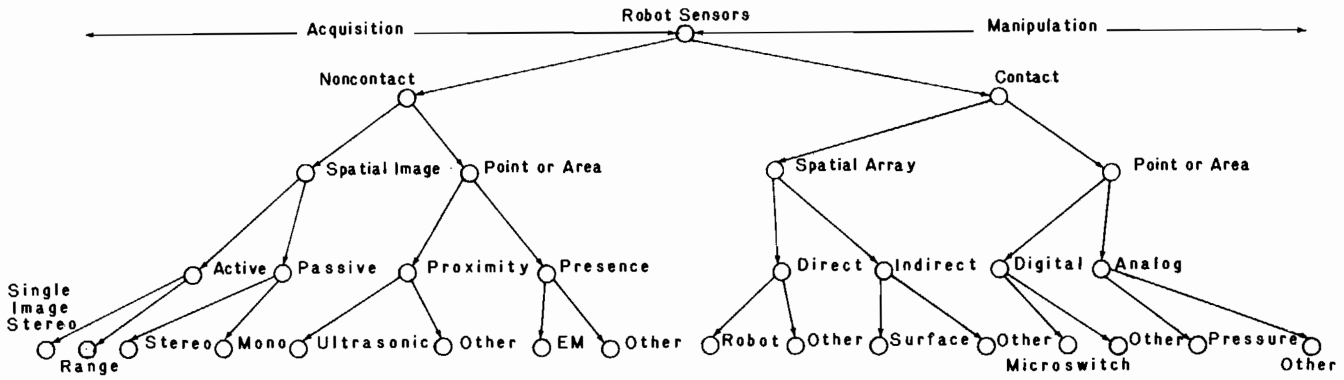


Figure 5–1. Classification of sensors available for use on an intelligent robot.

collection space. A camera is the most common example of a device that measures spatial information. Point or area sensors may be further divided into those that measure proximity and those that measure the presence of an object. Further subdivision of sensors into devices that measure spectral range, such as infrared, visible, and x-ray, may be developed.

A similar grouping may be developed for contact sensors, which are most useful in the manipulation stage of a process. Contact sensors may simply measure touch or may measure force or torque. The most common touch sensor is a simple switch that closes when it touches a part. Force or torque sensors may work according to Newton's law, that force is equal to the product of mass times acceleration, or torque is equal to the product of inertia times angular acceleration. A simple force sensor might measure acceleration with an accelerometer to sense force. These sensor categories may also be divided into whether a direct or indirect method of measurement is used. For example, the force could be measured directly at the robot hand or indirectly by its effect on the working surface. Force and touch sensors may be further subdivided into digital or analog or other categories. This listing of sensor types is not intended to be exhaustive, but does give some indication of the many types of sensors available for use on an intelligent robot. It is also interesting to note that most of the sensors are available in "hardened" form for use in environments of high temperature, pressure, and radiation, or corrosive, zero gravity, or other unusual environments.

5.2 Image Processing for Robot Vision

Let's now concentrate on one special sensor, the camera, and investigate what has been and may be possible through the use of computer vision. Computer vision research, development, and applications are aimed at understanding the complex visual processes so that simple, useful solutions to practical problems may be developed. Such areas as cognitive psychology, image processing, pattern recognition, machine intelligence, computer graphics, computer systems architectures and programming languages, engineering, science, mathematics, and even neurophysiology share common interests in the computer vision discipline.

Vision may be used in several ways in an intelligent robot. For accuracy, it may locate and track the robot hand to provide feedback control. The location, orientation, and recognition of parts to be picked up is another important application. Vision may be used to guide a seam-welding robot or to control the mating of two parts. Regardless of the application, the vision system must contain an illumination source, a camera system, and a computer interface. If ambient lighting is relied on for the illumination source, the imaging process may be called passive. This type of imaging is often used in military applications since the position of the viewer is not compromised. In industrial applications there is no such concern, so that controlled illumination or active imaging may be used. The camera system contains not only the camera detector but also, and very importantly, a lens system. The lens determines the field of view, the depth of focus, and other factors that directly affect the quality of the image detected by the camera. Novel techniques, such as using a fish-eye lens to obtain a 360-degree field of view with no focus adjustment, have recently been investigated and appear very useful in mobile robot applications.

The type of camera used is also important. A very popular type that has been used in several intelligent robots is a solid-state camera. This device offers greater sensitivity and ruggedness, is lightweight, and is easily interfaced to a computer. The camera may also contain important processing electronics. The camera/computer interface is also very important, since such factors as the number of resolution elements and distinct gray levels directly affect the quality of the image.

Several companies specialize in robot vision systems and are constantly improving their hardware and software. Two important hardware specifications of a computer imaging system are the number of points per square inch and the number of gray levels stored at each point.

If you consider the general scene, such as the one in front of you right now, you may see a set of objects sitting in some stable position on some supporting structure. Each object may be quite complex in both its three-dimensional surface shape and in its material properties. Just describing the scene so that it can be represented in computer data is a difficult problem in computer vision. The computer specialty that concentrates on displaying scenes is called computer graphics. It is now recognized that a complete representation of even a grain of salt is impossible even with our largest computer unless some form of model or pattern is used.

One way to model or pattern a scene is to break it into structures, then organize these structures into some sort of hierarchy. For instance, the structure of a scene can be thought of quite nicely in terms of a hierarchical tree, which is generally drawn as an inverted tree of several levels. In this case, the scene as a whole is the trunk, the objects that comprise the scene are the main branches, and the properties of those objects are smaller branches and twigs coming off the main branches.

To describe the surface of an object, let's imagine a "blocks world" in which the surface of each object is flat. We call these polygonal objects. Boxes, books, and houses fit quite well into this category of block objects. In the blocks world, each object is composed of a set of surfaces. For example, a cube has six surfaces. To represent this in our computer tree model, we let the three-dimensional cube be a main branch of the tree, the two-dimensional surfaces be small branches coming off the main branch, the one-dimensional properties of these surfaces be twigs coming off the small branches, and the zero-dimensional properties be the leaves.

Since each surface makes up one of the sides of the object, each surface must be represented by a separate, small branch connecting at the main branch. At this level of the tree, we are no longer dealing with three dimensions, but only two—length and width. To continue the decomposition, we note that each surface is bounded by straight lines or edges and is made of some material. These straight edges, being properties of the object's surface, are represented by twigs coming off the little branches representing each surface, and the material information is saved as a label on the trunk. This allows the computer to understand that each surface is bounded by straight lines or edges. At this level of the tree, we have only to deal with one dimension—length. As the final step in this decomposition, we may note that each edge is simply a line between its end points, which we will call vertices. Since the edges are represented by twigs, the vertices are now represented by leaves, and the computer understands that the vertex is an end point of the edge. This final dimensional reduction has taken us from one-dimensional to zero-dimensional entities, the vertex points.

In the real world, the situation is of course more complex because of the curved nature of most natural and many manufactured objects. These are much harder to break down into properties and represent. However, there is a mathematical theory of snowflake curves or "fractals," which introduces the complexities of even block-type objects. For example, a snowflake is a very complex block object. This theory also introduces the idea of a curve that can cover an entire surface and has a dimension greater than one but less than two. These fractional dimensional curves are only one interesting complexity. Another is related to the definition of a curved surface or even just a curve. Admittedly, we know about many types of curves, but how one represents the arbitrarily shaped curved object is still largely a puzzle. For now, we will assume that the surfaces encountered can be adequately represented by combinations of flat, quadric, and spline surfaces. Using all three types, scene description by hierarchical trees can still be made with only slightly more complexity. As we divide up a surface, we must segment it into simple surfaces, which are those that can be represented by planar, quartic, or spline surfaces. Also, for each surface the boundary now may be curved and perhaps described by data points or polynomial coefficients. This description is more realistic and more complex, as we might expect.

Given this method of describing a surface we may now consider two very important problems in the use of intelligent robots. The first is supplying the robot with a model of the scene with which it is working, that is, telling the robot what to expect. The second is enabling the robot to sense its environment and construct the scene so that it may compare sensed information with its modeled information. The first problem is in the province of computer graphics or computer-aided design. The second is the main emphasis of the robot vision field. Both may also be of interest in artificial intelligence, and both may need to be considered for some intelligent robot applications.

5.3 Robot Vision Specification

Let's now concentrate on the robot vision potential. Robot vision systems make the intelligent robot a useful machine in many applications. We are now in the third generation of robot vision systems. The first-generation systems work with silhouette images of objects and infer such parameters as the location, orientation, and size of an object from the shape of its silhouette. These systems are characterized by binary or two-level image processing, with the images produced by a back-lit scene. The second-generation robot vision systems use several gray levels to characterize objects. These systems can work with front-lit scenes and may differentiate texture patterns. The third-generation systems measure not only gray levels but also use stereo techniques to determine the three-dimensional coordinates of the visible objects in a scene. More advanced systems can even infer some information about surfaces that are not visible, such as the back of an object. We will now present some examples of each type of system.

There is an overall common structure between the human visual system and a machine vision system. One powerful computer peripheral that has only recently been made cost-effective and adaptable is the vision system. With computer vision systems, the capabilities of computer perception are greatly enhanced, and its potential for the performance of intelligent tasks is more apparent.

For robot applications, there are two main divisions of vision techniques. The first is for robot control, as illustrated in Figure 5–2. The special requirements of this application are high-speed computations required for robot control; therefore, the techniques used must be elegantly simple. Another class of robotic vision applications relates to inspection for quality control, as shown in Figure 5–3. Although there is still a desire for high-speed computations, the speed may be that of a conveyor rather than that of the robot. The situations are often stereotyped so that more sophisticated techniques may be used.

Illumination Systems

One of the first considerations in robot vision applications is the type of illumination to be used. In many other imaging applications, natural or "ambient" lighting is used. However, for industrial applications, ambient lighting is rarely sufficient. Therefore, an additional illumination system must be selected. Point, line, or area sources may be used. Also, spectral illumination may be selected to provide high contrast between the desired objects and background. Polarizing filters may be required to reduce undesirable specular "glare." Also, since a moving object may be involved, a "strobe" may be required to eliminate blurring resulting from motion. Various types of illuminators have been used with robot vision systems. Light tables that permit back-lighting are excellent for silhouette imaging. Line illumination produced with a cyclindrical lens is a key element in some robot vision systems. Laser point illumination is used in many three-dimensional measurement systems. The selection of an illumination system for robot vision is analogous to the selection of a staining method in microscopy. The proper selection often makes the problem much clearer.

Camera Positioning, Focus, Zoom, and Aperture Control

The selection of a method for camera positioning is also important. A fixed position is the simplest and is appropriate for some object recognition and inspection systems. In other applications, such as seam welding, a robot hand mounting is required. Often aperture, pan, tilt, zoom, and focus servo controls are required. The selection of these systems is application dependent and requires careful study.

Camera Selection

Camera selection is also an important consideration in robot vision systems. One cannot obtain reliable and accurate data from a camera that does not have reliable characteris-

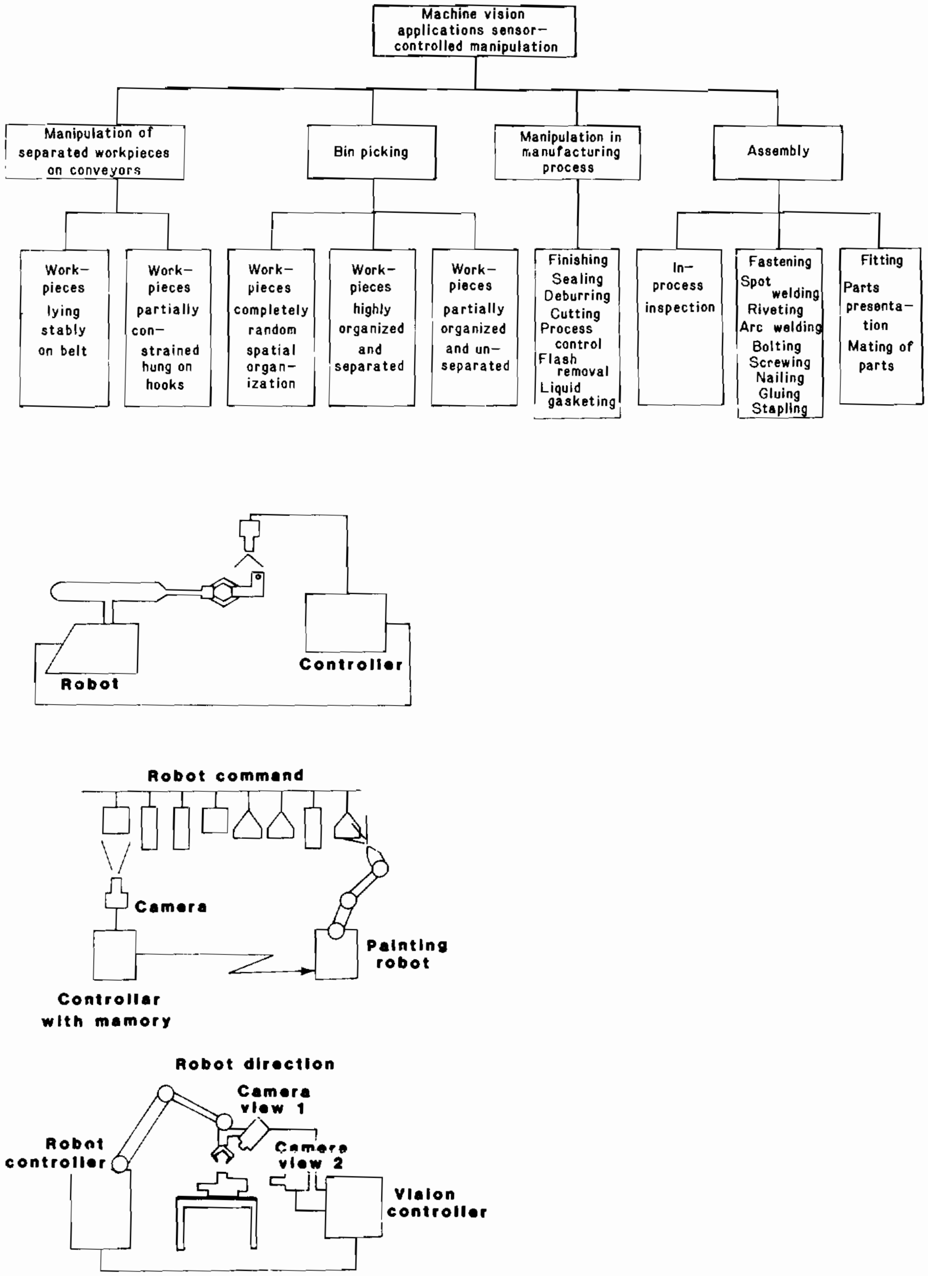


Figure 5–2. Machine vision applications for robot control. (Courtesy of Copperweld Robotics, Inc., Troy, Michigan.)

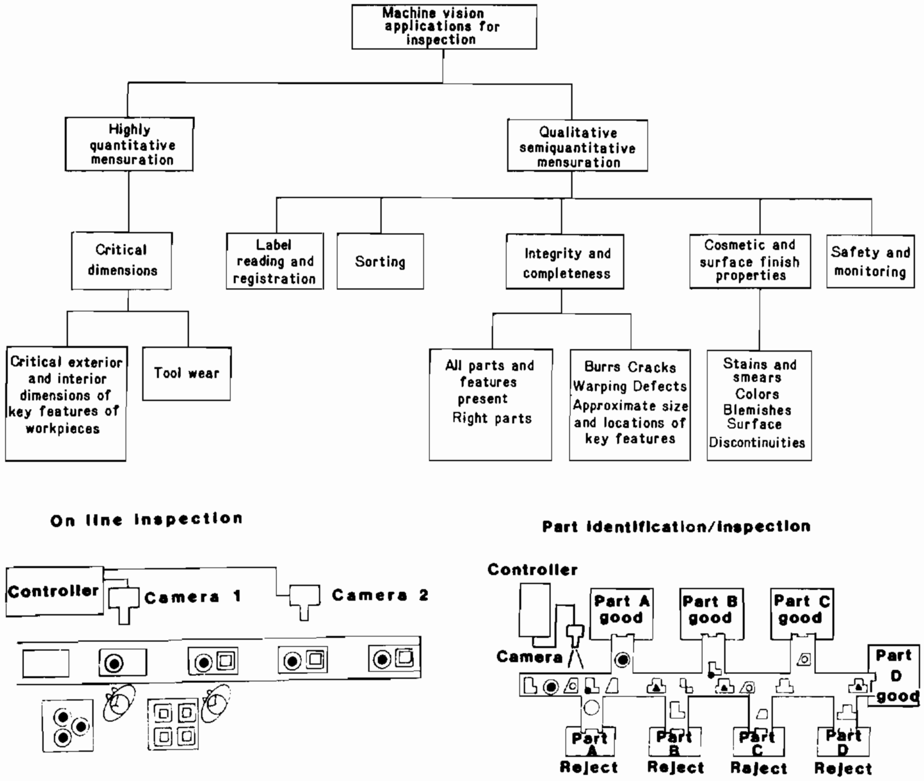


Figure 5–3. Machine vision applications for visual inspection. (Courtesy of Copperweld Robotics, Inc., Troy, Michigan.)

tics. Of the many different types of electronic camera sensors available, two are predominantly used in robotics applications—the vidicon, and the solid-state array. A vidicon is an electronic tube consisting of a glass faceplate with an embedded phosphor surface that is sensitive to light and has an electronic readout capability. The solid-state array camera is an integrated circuit array of discrete photodetectors with associated circuitry. The low cost of broadcast TV cameras often provides a false sense of cost for a measurement-quality camera. Only those cameras produced in large quantities are available at low cost. All these cameras are designed for image visualization, not image measurement. Therefore, a high-quality or special-purpose image-measurement camera can be quite costly. Camera specifications include scanning format, geometric precision and stability, band width, spectral response, signal-to-noise ratio, automatic gain control, and gain and offset stability.

Camera selection includes selection of the lens and camera, digitization and processing unit, and computer interface. Each of these components is important to the overall system. The lens determines the field of view and depth of focus. For a fixed camera location, these are not difficult to determine. However, for a robot hand-mounted system, some method of automatic focus may be required. A variety of cameras is available. A primary consideration is response time. With standard imaging cameras, a frame time of one-sixtieth of a second may be the smallest time increment involved. Tracking cameras are also available that provide continuous x and y locations of a spot. The digitization and control unit may have the capability of digitizing and storing an entire image frame in one-sixtieth or one-thirtieth of a second, or it may require up to several seconds to digitize a frame. The computer interface and processing software is the final element. Although a considerable amount of image processing software has been developed (Hall, 1979), only a small subset of this can be applied to real-time robot control with a general-purpose computer. New computer architectures are being developed for these new applications.

Digitizing an Image

The first step in the processing of an image for robot control is to convert the image into a computer-compatible form. This requires a two-step procedure called sampling and quantization. Sampling is the process of converting the continuous spatial information to discrete sample values at points on a usually equispaced grid. Sampling is usually done at resolutions that are powers of 2, such as 128 by 128, 256 by 256, or 512 by 512 picture elements, or pixels. Next, the brightness or intensity at each point is divided into discrete levels, such as 2 for binary images or perhaps 64, 128, or 256 for gray-scale images. These two steps provide an image that can be stored in computer memory for processing.

Note that the memory size required for storing an image grows quickly as the resolution is increased. For example, the memory required for a 64 by 64 binary image is 4096 bits, which is 4K of memory (1K is 1024 bits). This image could be stored in only 512 bytes since a byte is a group of 8 bits. The same size of spatial image stored at 8 bits/pixel for 256 shades of gray would require 4K bytes of memory. Increasing the spatial resolution to 256 by 256 pixels increases the memory requirements to 64K bytes for an 8-bit/pixel image. Note that this much storage would require an address length of 16 bits and would totally fill the memory on a microprocessor that has a 16-bit address length. Increasing the spatial resolution to 512 by 512 pixels again with 8 bits/pixel requires 256K bytes of memory. This resolution could not be stored in the direct memory of a microprocessor with a 16-bit address length but could be easily accommodated by one with a 32-bit address.

These large memory requirements are one indication of the complexity involved in image processing. A modest microprocessor could easily store several 64 by 64 binary images that require only 512 bytes of storage, but could only hold 16 images of this size at 8 bits/pixel. This same microprocessor could not accommodate a single 256 by 256 image at 8 bits/pixel without using the entire 64K memory.

Advanced architectures specifically designed for high-speed image processing are available; however, these have not been specifically designed for robot vision. The market is still young and growing, so we can expect to see many more robot vision systems develop over the next few years.

Another consideration for robot vision is processing speed. The U.S. standard TV image rate is 30 frames per second. A "frame grabber" that can digitize and store an image in one-thirtieth of a second is required for high-speed processing. Total processing times of the order of one-tenth of a second may be required to provide updates to a robot controller, which is at least 10 times higher than the robot's natural resonant frequency or that required to keep up with a moving conveyor line. More time may be available for inspection applications; however, if a conveyor is moving at 100 inches/second, 10 inches of material will pass by in one-tenth of a second. Processing speed is often crucial.

One image-processing operation that can easily be performed in hardware in one frame time is the computation of a histogram. A histogram is simply a tabulation of the number of image points at each particular gray level. For example, if an 8-bit/pixel, or 256 shades of gray, image is obtained from an analog-to-digital conversion of the camera's video signal, then only 256 storage cells are required to store the histogram. The histogram is computed by simply incrementing the storage cell for each pixel, using the gray level value to point to the correct memory location. A gray-level histogram is shown in Figure 5–4. If an image is to be converted into binary form, the valley in the histogram may be located and used as a threshold value. Any gray level below the threshold is set to 0, and any value above the threshold is set to 1 to produce a binary image.

Robot vision systems greatly enhance the abilities of both computers and robots. Many robotics systems are trained to perform specific tasks and require no human supervision once the training is complete. However, the training process alone can only be used if the task is well defined and does not change. This requires that the environment be well controlled and synchronized with the movements of the robot manipulator. If, for example, a robot is trained to pick up and place a part for assembly, each piece must be positioned in the proper place so the robot can locate it. Robot vision systems can reduce or eliminate this constraint by allowing the robot manipulator to locate each piece by itself. This enhancement to robotics can be used to provide a feedback path to adjust the trained procedure in accordance with the environment.

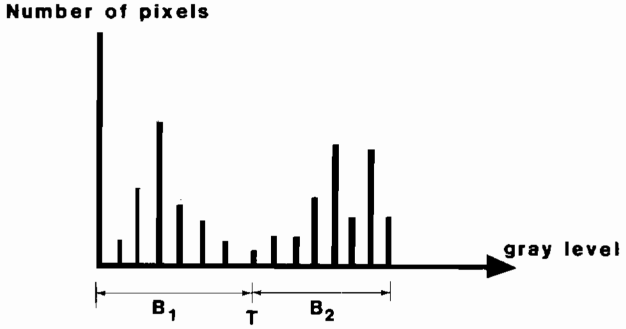


Figure 5–4. Gray-level histogram used to determine a threshold for producing a binary image.

Processing Examples

Some early vision systems located objects by interpreting polygonal surfaces in a block world like the one we described earlier. Objects of this type have distinguishable features that can be identified by segmentation methods. These features can be used as reference points of the surface for object location. However, a featureless curved object, like a ball, has no unique features on its surface, making it difficult to measure from a two-dimensional image. Unfortunately, this class of objects encompasses many machined workpieces.

The three-dimensional solution necessary for curved object location, recognition, and manipulation has many different approaches. The region method of representing polyhedra incorporates an approach called active imaging, which involves a projection system and a camera system. The projection system is used to impose features on the surface of the object, and the many resultant patterns can be used to obtain different types of images. The camera system acquires the object image with the imposed features, and necessary calculations are done to interpret the image for analysis. Active imaging techniques are ideally suited for industrial applications. One advantage of the active approach is that it provides a high degree of control over the environment by providing special illumination or markers for accurate location. The accuracy of the object location derived from a projected grid system can be controlled by the grid spacing. "Stereo" imaging with a single image may also be accomplished. This method uses the projected pattern as the first image and the received image as the second, and a standard stereo solution results. Other techniques involve the use of cooperative algorithms to determine the orientation from a single view using a single camera to recognize overlapping parts, or using geometric and relational reasoning to recognize a three-dimensional object from a single view.

Successful developments in active imaging techniques include the CONSIGHT system developed by General Motors Technical Center (Holland et al., 1979). The CONSIGHT system uses a linear array camera and two projected light lines focused as one line on a belt, as shown in Figure 5–5. The system handles a wide class of manufactured parts in a nonideal industrial environment. The principal functions of the CONSIGHT system are object detection, position determination, part pickup, and part transfer to a stacking location. The overall system may be considered in two parts: the vision subsystem and the robot subsystem. The vision system uses a linear array camera with 256 elements focused on the conveyor belt. It also uses two line beams formed by cylindrical lenses to project a line on the conveyor belt directly under the camera. When an object comes under the camera, a portion of the line is shifted horizontally in proportion to the height of the object. The camera control senses the edges of the object by the absence of the line. Features can then be computed from the edge image, including the position of the object, the area, a bounding rectangle, the area of the largest hole, and the centroid and orientation of the largest object. Recognition is accomplished by comparing the computed features with those of prestored prototypes. Given this information, the robot subsystem is then able to track and pick up the object and place it in an appropriate location. The position on the part at which the robot can pick it up is also determined by the vision system.

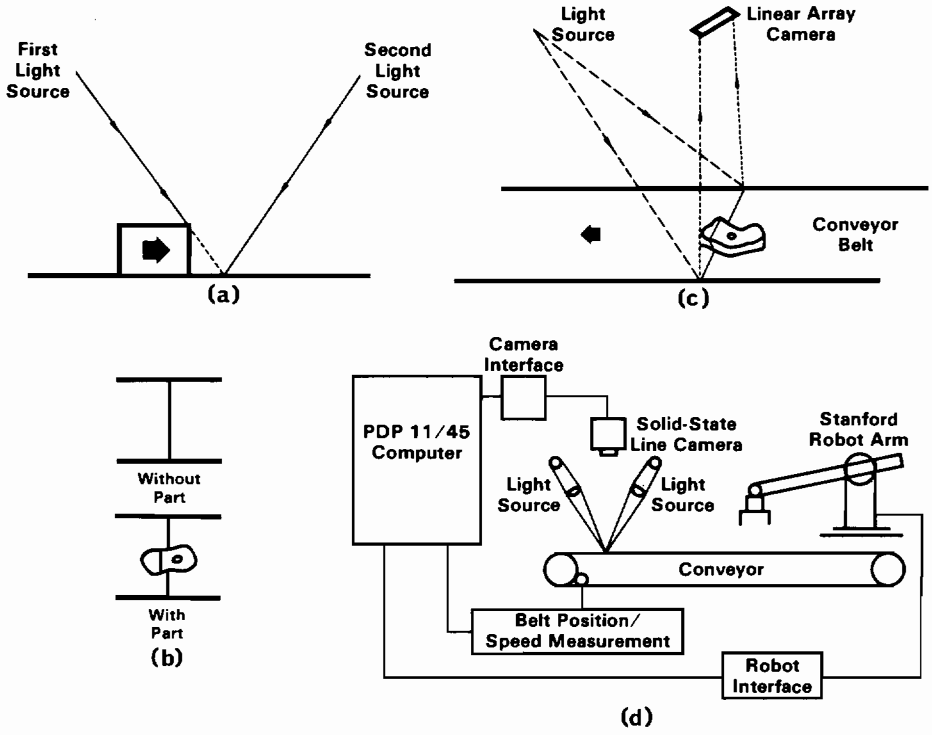
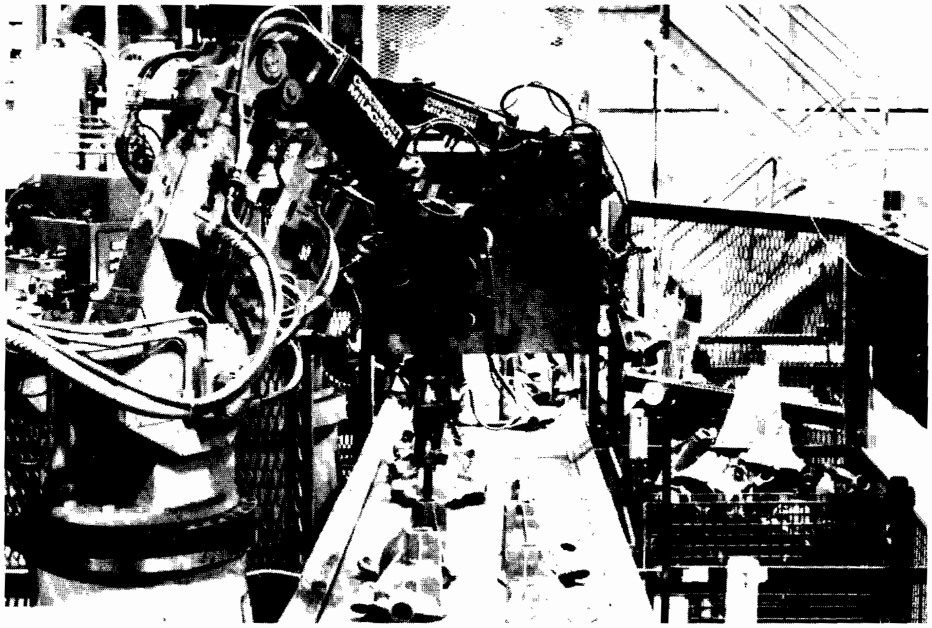


Figure 5–5. The CONSIGHT system developed at General Motors Technical Center. (a) Basic lighting principle. The principle of the lighting apparatus is illustrated here. A narrow and intense line of light is projected across the belt surface. The line camera is positioned to image the target line across the belt. When an object passes into the beam, it intercepts the light before it reaches the belt surface. When viewed from above, the line appears deflected from its target wherever a part is passing on the belt. Therefore, wherever the camera sees brightness, it is viewing the unobstructed belt surface; wherever the camera sees darkness, it is viewing the passing part. (b) Computer's view of parts. (c) Improved lighting arrangement. Unfortunately, a shadowing effect causes the object to block the light before it actually reaches the imaged line. The solution is to use two (or more) light sources, all directed at the same strip across the belt. When the first light source is prematurely interrupted, the second will normally not be. By using multiple light sources and by adjusting the angle of incidence appropriately, the problem is essentially eliminated. (d) CONSIGHT hardware schematic. Since its speed is neither constant nor predictable, the belt is provided with a position and speed detection device. Position and speed information are necessary because the camera scans the belt at a constant rate, independent of belt speed. For each equal increment of belt travel, the vision subsystem records one of these scans. Belt travel increments must therefore be measured precisely. (e) A realization of the CONSIGHT system on a robot assembly line. (Courtesy of General Motors Research Laboratories, Warren, Michigan.)



(e)

Figure 5–5 (continued)

Another system was developed by Albus at the National Bureau of Standards (Albus, 1981). Albus's system also uses a plane of light to determine the position and orientation of parts on a table. However, the camera and a strobed line light source are mounted on the robot arm. This gives the system the flexibility of approaching a set of objects from many different directions. Also, the camera used is a two-dimensional array so that the amount of displacement in the line may be used to determine the distance to the object. The line projection technique works quite well on locating the edges of an object but requires scanning to locate all the object surface points.

The use of grid coding for image segmentation based on the spatial frequencies of the projected images has also been proposed. Other computation techniques have been used recently on the laser/shutter/space encoding systems. These systems use the time and space coding of dots to measure surface coordinates.

The human visual system features two perspective views of objects from which distance or depth information is derived. A computer vision system can do the same thing by using two cameras to view an object from two different perspectives. The human visual system also has the ability to detect depth changes from shading information. Depth information can also be derived from images obtained by a computer vision system through an analysis of the shading of the surface along with a knowledge of lighting conditions. A model can be obtained, given a format for the model, using numerical or analytic techniques of least-squares curve fitting. Once a model is obtained, the mathematical model may be decomposed into a qualitative description of the surface. The model could then be used to describe the power of the return signal as a function of the incidence angle between the surface normal and the transmitted signal vector. There is a concern for characterizing the reflection of light on surfaces with different orientations. Numerous computer-aided design systems, such as those marketed by General Electric, Computer Vision Corporation, and Applicon Systems, can simulate the appearance of a mechanical part or architectural structure through graphic reconstruction. For these reconstructions to appear natural, an understanding of the reflection characteristics of surfaces is required.

Polygonal surfaces are among the simplest to describe. Simple techniques are used in their reconstruction on graphic displays. One popular program, called MOVIE.BYU, developed at Brigham Young University (Christenson, 1978), uses a vertex list and a vertex connection table to describe a surface. This representation provides sufficient information to reconstruct the projection of the surface onto a plane positioned anywhere in space and oriented in any way relative to a global coordinate system. Other data structures for describing polygonal surfaces use node or vertex locations with a variety of schemes for describing the connection of points in graphic reproduction. However, similar techniques cannot be applied to curved surface representation since there exist no vertices, edges, or sides on a curved surface, such as a sphere.

Polygonal surfaces are constructed of multiple intersecting planes, each of which may be described by a linear equation. The line defined by the intersection of two or more planes is considered an edge, and a vertex is the point of intersection of two or more edges. The edges of a polygonal object are discontinuities in the surface description. Since an object of this type is constructed of bounded planes that form the polygons, the edges are line segments at which one element of the surface ends and another begins. Curved surfaces may also have edges formed by discontinuities if the surface is formed by multiple bounded curved surfaces where each of the elemental curved surfaces describes only a bounded region of the entire surface. With no restriction on the order of the equations required to describe a general surface, the complexity of the mathematical model could approach that of the molecular-level description. Again, the search for an accurate description of a general surface becomes futile. Therefore, the methods employed in the representation of arbitrary curved surfaces must rely on approximation techniques. Many manufactured objects and structures can often be described by low-order equations for which a mathematical model can be derived. In such cases, a mathematical model would be the preferred method of representing the surface, since a single equation may be used to obtain any point on the surface. Numerous methods of describing curved surfaces, both by mathematical models and by approximations, are known.

A surface recognition technique requires a sufficient sample of surface points, as required by the sampling theorem, for adequate surface description. Stereo vision has long been accepted as a valid method of surface measurement; however, since the stereo vision technique requires the identification of corresponding points in the two images, this method is difficult to apply to curved surfaces. The vertices of polygonal solids are the features most commonly used to match two points on a pair of images. Curved surfaces may not contain vertices, but still be closed surfaces. If a featureless curved surface, such as a sphere, is viewed from two different positions, the two images may appear identical under uniform lighting. One solution to curved surface measurement utilizes the characteristics of the change in surface reflectivity as the illumination source is moved from one position to another.

The models that are applied to image synthesis suggest the ability of the vision system to recognize surface shapes from shading information. The process of surface shading involves the assignment of an intensity or color to every picture element in the image that accurately simulates the viewing situation. The shading of a surface point depends on the surface reflection characteristics, the surface geometry, and the lighting conditions. Each of these properties must be considered in the development of a surface shading model. Models used in image synthesis for the shading of the surface of an object can be applied to image analysis to obtain the surface shape from the shading information. However, the ability to perform this task with sufficient accuracy depends upon the ability to select an appropriate shading model that closely approximates the reflection characteristics of the surface material. The problem of selecting an appropriate shading model is a topic that requires continued research. Present techniques of calibration are difficult to apply to automated systems, since a new set of calibration parameters must be obtained for every surface to be examined.

The technique of stereo vision for surface measurement is well known in photogrammetry and has long been applied to three-dimensional data acquisition and object description. The stereo vision process requires that the camera models be known for each view of the object. Through a knowledge of the camera model parameters of location, orientation, and focal length, the perspective projection transformation matrix for each camera can be determined. Since it is often difficult to obtain the model parameters directly, the camera calibration procedure may be applied using six points for which the global coordinates and image coordinates are known. Once the transformation matrices are known, the global coordinates of a surface point may be computed.

Curved surface representation techniques have been described whereby a model of the surface may be obtained that is applicable to surface recognition. These models may consist of mathematical models or numerical approximations. To obtain the necessary sample surface points to describe the surface geometry adequately, several methods of curved surface measurements have been developed. One method is based upon the relations between surface shading and surface geometry. A second method utilizes the stereo vision approach to surface measurement but requires only a single image of the surface to be analyzed.

Curved surface representation is a necessary step in curved surface recognition. Because objects comprised of curved surfaces may have no edges or vertices, surface models must be used to describe their geometry in three dimensions. These models may consist of mathematical models or numerical approximations, such as splines. Mathematical modeling is often the best approach to surface description, since the information contained in the model may be easily decomposed into qualitative descriptors, including shape, size, orientation, and location. A description of this type provides an object recognition technique independent of the location or orientation of the object.

To describe an object, some level of knowledge must exist for which a qualitative description can be derived. The level of knowledge obtained by the two techniques to be considered is a set of sample surface point coordinates. One method of obtaining sample surface points is by examining the shading of the surface under controlled lighting conditions. It has been demonstrated that the models used in image synthesis for surface shading can be applied in image analysis to obtain the surface shape. If an appropriate shading model for the surface material and texture is obtained, a measurement of the surface normal vector for the surface's points viewed by each pixel in the image can be computed. This method does not require a high-resolution image to accurately obtain a sufficient number of sample surface points to describe the object. Obtaining an adequate shading model for a particular surface material and texture is a difficult problem that will require further research.

A second method is described that utilizes the stereo vision principle. Vertices and edges are the features most commonly matched in a pair of images, and by triangulation, the three-dimensional location of the feature is obtained. Featureless curved surfaces may have no edges or vertices and still be closed surfaces. The solution to this problem is to impose features onto the surface through "active imaging." The system described uses both a camera and a projection system. Because the projection process is identical to the imaging process with respect to the process models, this system performs stereo vision with a single image (Hall et al., 1982).

5.4 Commercial Robot Vision Systems

Three different types of robot vision systems for robot control and inspection are commercially available. The first type deals with two-dimensional binary images, or silhouette imagery. Many of these systems are based upon the pioneering research done at SRI International in Menlo Park, California. The second type uses gray-scale imagery from which more discrimination is possible but for which longer computation times may be required. The third type uses some type of structured light and stereo triangulation to determine the three-dimensional surface of objects. Each approach requires a lighting system for illuminating the objects, a camera for gathering the image, and a storage device for recording the image during processing.

Silhouette Robot Vision

As an example of a silhouette robot vision system, let us consider the Machine Intelligence Corporation's VS-100 system. The system may be divided into two major parts: a camera processor and a feature processor, as shown in Figure 5–6. The camera processor can control a strobe illuminator and accept input from one or more cameras. The LSI-11 microcomputer performs the supervisory tasks, computes features, and performs the object recognition. The system recognizes objects by comparing features on the test object with those of stored prototypes. The prototype features are determined during a training session in which exemplars are shown to the system and the recognition features determined. Several features, including geometric properties, such as the area

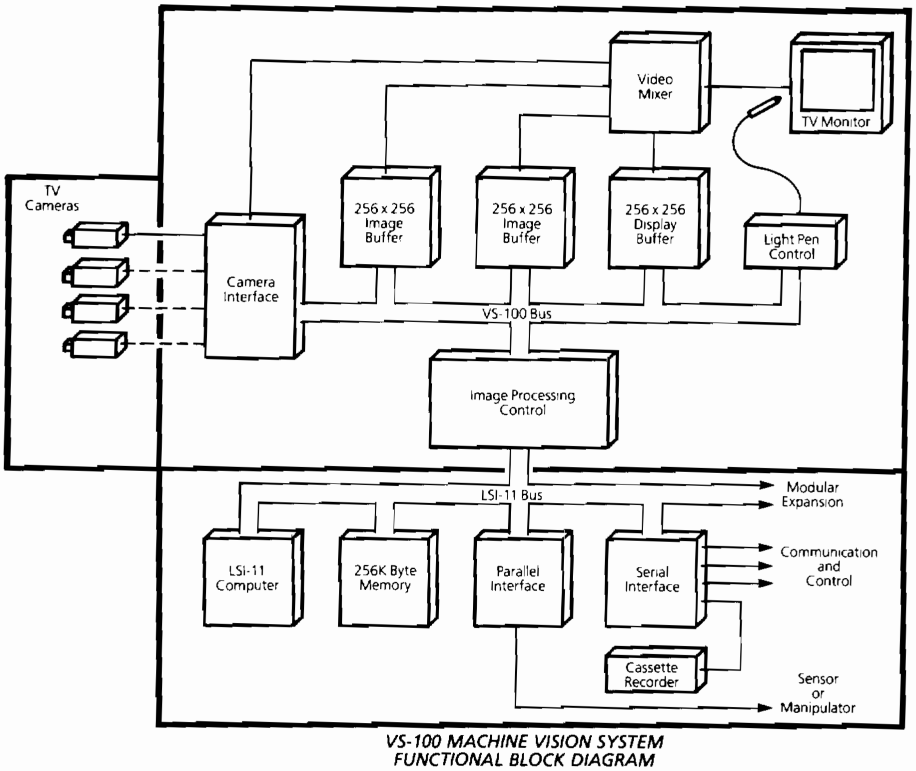


Figure 5–6. Machine Intelligence Corporation's VS-100 system functional block diagram showing divisions into camera and feature processor. (Courtesy of Machine Intelligence Corporation, Sunnyvale, California.)

and perimeter, as well as orientation properties, such as the centroid and principal orientation, may be selected for computation. Note that the geometric features are invariant to location and orientation and permit recognition invariant to the factors.

During training an image is scanned, a histogram of the gray levels is computed, and a threshold value is selected. The threshold is usually at the valley between the light and dark peaks in the histogram. If a clear valley does not occur, this indicates that the lighting should be adjusted to produce a sharper contrast between the object and the background. One illumination method that produces a very sharp contrast is backlighting the object. This is easily accomplished using a light table. High contrast can also be achieved in most cases using front lighting with an intense source.

Recognition of objects involves several steps. First, the image is scanned and, using the predetermined threshold, is stored in binary form. Next, the outlines of the objects are determined by finding edge points. Starting at the upper left-hand corner and moving line by line from top to bottom, the edge transition points are located. A compressed image consisting of the length of pixels of the same brightness or run length code is also determined. Whenever an edge point is determined, the computer determines which object it belongs to by examining neighboring points. It then enters this edge point into the list of edge points for that object. The edge lists are considered tentative until the scan is completed so that objects with holes or protrusions may be accommodated. The edge lists are processed at the end of the scan and merged if necessary.

After the edge lists have been processed, the features, such as the perimeter length and area, are computed, as well as the location and orientation. Finally, the system compares the computed features with the prestored prototype features to recognize the object. An illustration of the processing and the VS-100 system is shown in Figure 5–7.

As another example, let's consider the first vision system offered by a robot manufacturer. The Copperweld Robotics, Opto-Sense vision system could be used alone or interfaced to a robot. Typical applications have included measuring automobile frames on a moving assembly line, looking for missing parts in metal subassemblies, detecting missing objects in packages prior to sealing, and inspecting complex stamp-

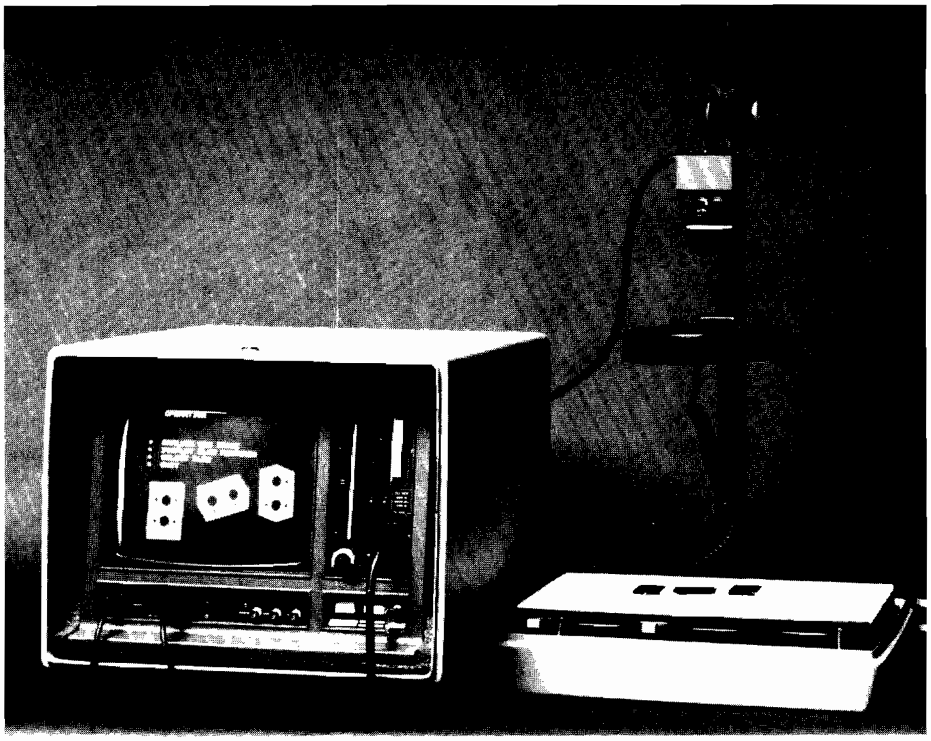


Figure 5–7. The Machine Intelligence Corporation system showing operation and images. (Courtesy of Machine Intelligence Corporation, Sunnyvale, California.)

ings for the presence and location of holes and other details. When used with a robot, the system provided the robot with visual feedback. The Opto-Sense system consisted of a computer controller that controls one or more solid-state cameras, such as the General Electric TN 2500. The image was divided into an array of 244 by 244 pixels. Software was usually tailored to particular applications using a subroutine package for various tasks.

In one application of robotic inspection at the Chevrolet Motor Division of General Motors in Flint, Michigan, the Opto-Sense was used with four cameras to determine if valve covers for engines were properly assembled. The cameras were mounted overhead in protective enclosures with transparent windows. Since the cameras were focused on the assembly line, dust on the lenses had very little effect. Reflected light and shadows permitted the imaging system to discern the presence or absence of detail. The first task performed by the vision system was to establish part identity, that is, whether a left- or right-side cover was being inspected. The second task was to determine the presence or absence of all necessary characteristics, such as clinch nuts, metal brackets, baffles, and holes. This step also determined whether any extra parts were added. This information was analyzed, and an accept or reject decision was made. This decision information was then given to the robot controller, which also had information available from other tests. The robot controller then determined whether the part should be put into an accept chute, a visual reject chute, or a leak test reject chute. The robot then put the part in the appropriate location. The system increased production from 300 to 1200 parts per hour. This was a 400 percent increase in productivity with 100 percent inspection. The system diagram is shown in Figure 5–8.

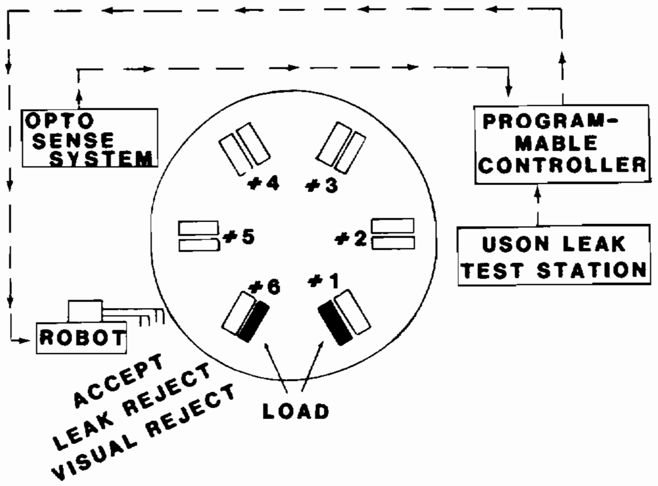
Gray-Scale Robot Vision Systems

As an example of a robot vision system capable of gray-scale processing, let's consider the Control Automation, Inc., InterVision 1000 system, shown in Figure 5–9. This system processes 6 bits/pixel, or 64 gray-level images, at rates up to 10 per second on the entire image or 80 window inspections per second.

Another robot vision system capable of processing gray-scale images is the Automatix, Inc., Cybervision III system, shown in Figure 5–10. The Automatix system is based upon a special hardware architecture using 68000 microprocessors and is programmable in a high-level language called RAIL.

Three-Dimensional Robot Vision

A striking example of three-dimensional robot vision is the system made by Robotic Vision Systems, Inc., shown in Figure 5–11. This system was designed for guiding a seam-welding robot and consists of a light projector and camera mounted in protective covers on the robot wrist. The system measures the orientation, position, width, and depth of a seam to be welded. This information is then transmitted to the robot controller to guide the robot path and weld parameters. This eliminates the need for expensive special-purpose fixtures to precisely locate the seam with respect to the robot. Other



INSPECTION SEQUENCE:

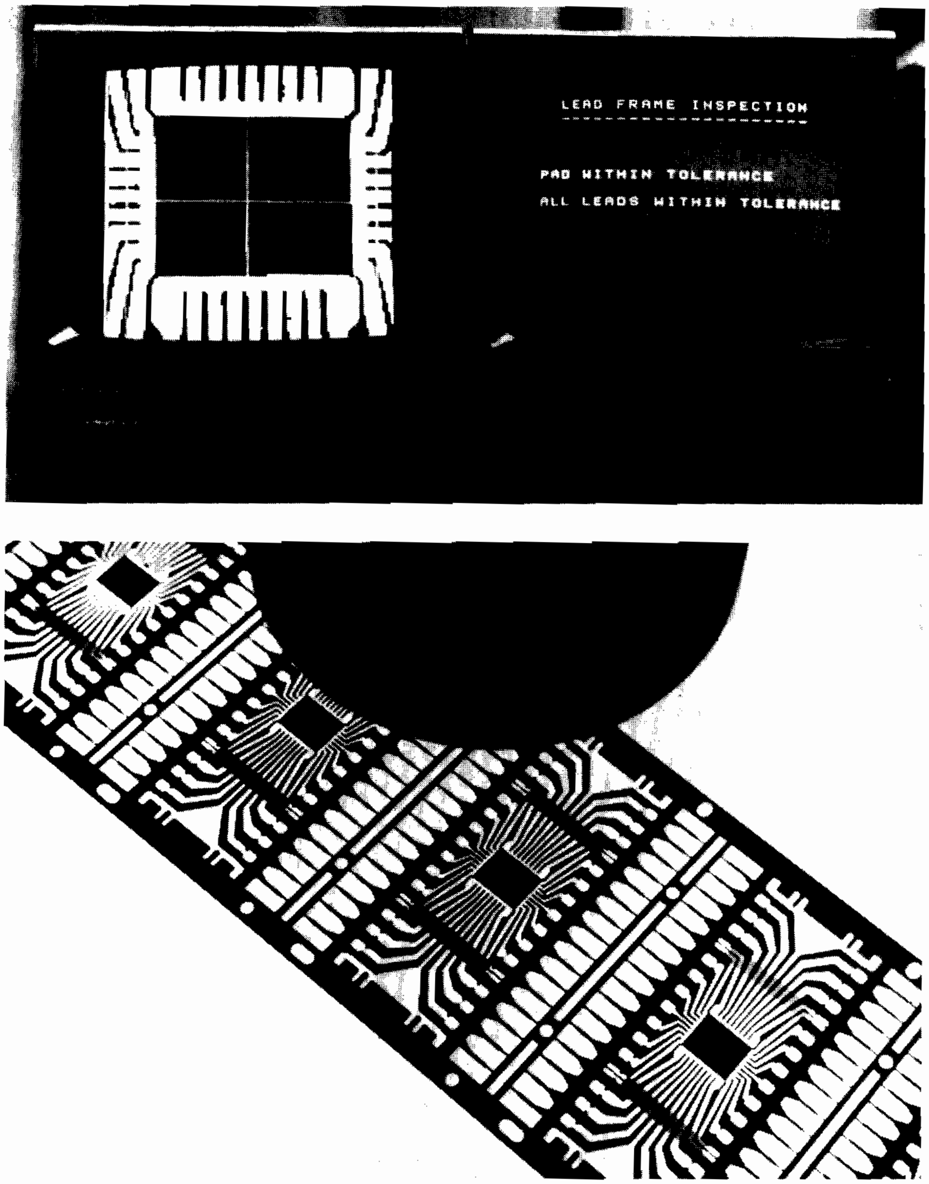
St. No. 1: (left-hand fixture) load
St. No. 2: Uson Leak Tester
St. No. 3: Idle position
St. No. 4: Opto Sense System
St. No. 5: Auto-Place Robot sorts and unloads valve covers
St. No. 6: (right-hand fixture) load

Figure 5–8. Copperweld's Opto-Sense inspection sequence diagram. (Courtesy of Copperweld, Inc., Troy, Michigan.)

applications of the system include automatic control of robot grinding, assembly, and materials handling and routing; automatic inspection of cast, forged, and machined parts; inspecting for missing parts; and three-dimensional copying and scaling of parts.

5.5 Range and Proximity Sensors

A range sensor measures the distance from the sensor to an object. Both point or proximity sensors and range image sensors have been developed for robotics applications. One of the most popular ranging devices is the Polaroid ultrasonic device originally developed for focus control of cameras. This device is now used on many robots, especially mobile robots for obstacle avoidance. A control computer generates an initiate signal to the Polaroid ultrasonic electronics control board. The ultrasonic board then generates a series of 56 pulses of four different ultrasonic frequencies for a total duration of 1 millisecond. This signal is applied to the transducer and can be heard



LEAD FRAME INSPECTION

PAD WITHIN TOLERANCE

ALL LEADS WITHIN TOLERANCE

Figure 5–9. Control Automation's InterVision 1000 vision system is a computer-based inspection system with solid-state camera that combines binary and 64-level gray-scale processing to inspect products at 4800 parts per minute. InterVision can be used for noncontact measurement, profile recognition, and presence or absence inspections of randomly oriented, moving, or fixtured products, such as lead frames, leads of components, and populated printed circuit boards. (Courtesy of Control Automation, Inc., Princeton, New Jersey.)

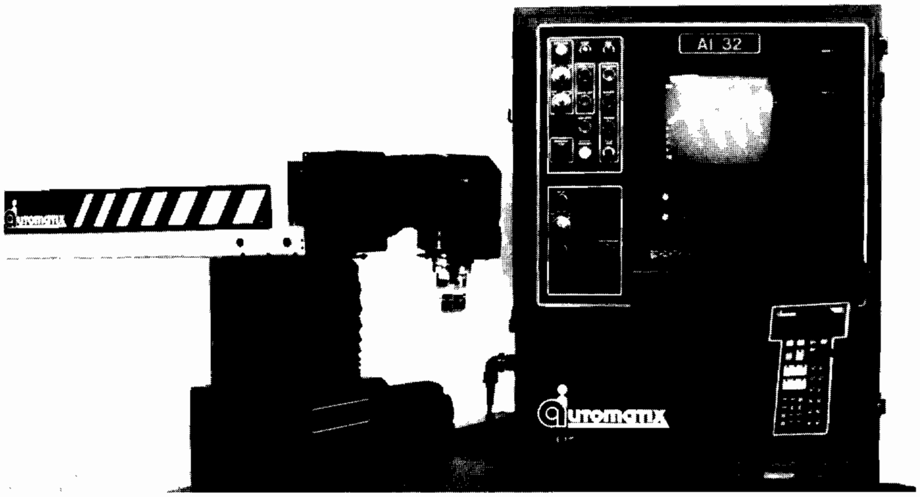


Figure 5–10. Automatix, Inc., vision system, the Cybervision III system. The system is programmed in a structured language that incorporates special vision features and is easily interfaced with some robot controllers. (Courtesy Automatix, Inc., Billerica, Massachusetts.)

as an audible "chirp." A pulse-sent signal is then sent to the computer and control circuitry at the same time that the pulse is transmitted. This signal causes the range count to increment at a rate of 3.2 kilohertz. Upon detection of the reflected echo, the Polaroid board sends an echo-received signal to terminate the count. A maximum range count of 127 corresponds to a distance of 22.5 feet. A single count represents a distance increment of about 2 inches. The range represents the minimum range encountered in a cone of sensitivity corresponding to the radiation pattern of the ultrasonic beam.

Laser range finders and imaging devices have also been developed for robotics applications. For example, the time of flight can be measured directly for large distances. However, since light travels at a speed of 1 foot per nanosecond, very high frequency counters are required. An alternative approach, which is more appropriate for the distances encountered in industrial applications, is to measure the phase shift between the transmitted and received beams. Another approach is to use the laser as a light source in a stereo triangulation system to measure the distance to an object. One possible problem with laser devices is the eye hazard to humans. They must be used with care.

5.6 Tactile Sensors

When a robot hand is close enough to touch an object, contact sensors may be used. Just as in the human skin, which contains pressure, pain, and temperature sensors, a variety

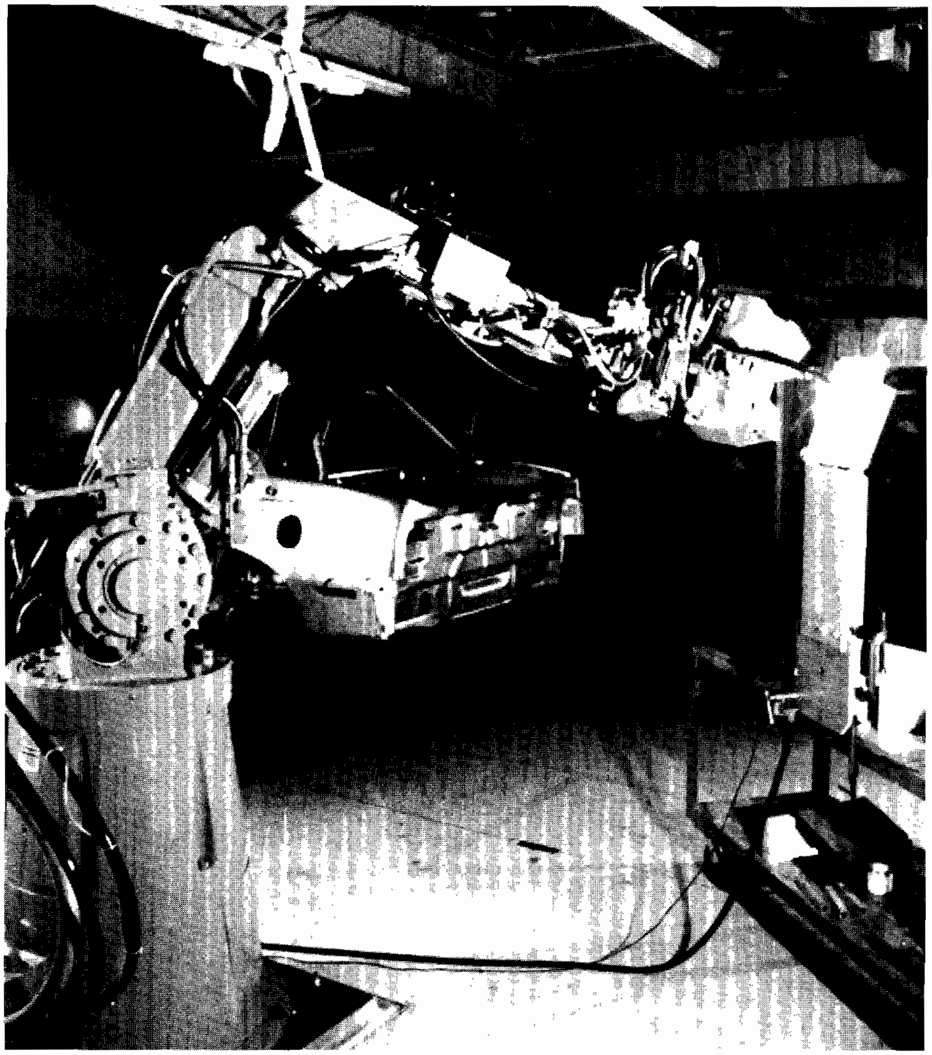


Figure 5–11. Robotic Vision, Inc., three-dimensional vision system mounted on a robot and used for guiding a seam-welding gun. (Courtesy of Robotic Vision, Inc., Melville, New York.)

of contact sensors for robots have been developed. Force or pressure, which is simply force per unit area, can be measured at a single point to help adjust the position of a tool, or at several points to provide enough information for the robot to recognize an object.

Perhaps the simplest sensor in this category is a touch sensor implemented with a microswitch. Such sensors are often used to stop the motion of the robot in a particular axis or direction. For example, the feedback on a nonservo robot may be supplied by such a stop or switch in a "move until touch" mode. Also, the closing of a gripper may be halted when a switch is closed. Other applications include sensing that a target has been reached, such as in spot welding, preventing collision, centering the robot gripper on a object without moving it, measuring object dimensions using the switch in conjunction with the high-precision joint encoders, and determining object presence. Point touch sensors are excellent for such tasks in which a single point of contact is sufficient. They are often inadequate for such tasks as part placement or recognition in which several contact measurements are needed.

Force sensors employ a transducer, such as a piezoelectric element, which provides a signal proportional to the deflection and therefore force applied to the point of contact. Such a measurement may be used to provide force feedback for collision recovery or to permit a robot gripper to grasp a delicate object, such as an egg, without crushing it.

In many assembly applications, a sensor that can measure the three-dimensional forces and torques may be very useful. For example, in inserting a bearing into its holder, a very high tolerance may be required. A single force measurement would only indicate that the bearing is in contact with the receptacle. However, three force measurements in the x, y, and z directions could indicate the displacement of the part from the hole. Additional measurements of the three torques (force times distance) may be sufficient to also yield rotational offsets that may be used with the displacement information to permit insertion of the part. Astez Corporation makes such a force-torque sensor that may be attached directly to a robot wrist.

Arrays of touch or force sensors may provide sufficient information to recognize an object, determine how the object is resting, determine its center of mass or pickup point, or its orientation. Lord Corporation of Erie, Pennsylvania, has used such an array sensor for object recognition and location for assembly. The Lord LTS-100 tactile sensor is 3.18 inches square and 1.12 inches thick. It contains 64 sensing sites mounted on an 8 by 8 array with 0.3-inch spacing. The device employs two stages of transduction to convert the impressions on its touch surface into easily processed electrical signals. One stage is a mechanical deflectometer that has an elastic compliant touch surface in which an 8 by 8 array of pins constructed of the same elastic material are embedded. A deflection of any pin is measured by a light emitter and detector pair. As the pin is deflected, it shadows a portion of the light. The maximum amount of light is received by the detector when no deflection is present, and the minimum amount when full deflection of the pin shadows most of the light. The deflections range from 0 to 0.8 inch at 1 pound force. The sensor light signals are converted into 64 levels of displacement. Since the device measures displacement, force measurements are developed by comparing the amount of deflection to the force deflection characteristics of the elastic material. The sensor has been used in conjunction with a Unimate 560 robot for assembly of a flexible diaphragm and plastic cap. In this application, the tactile sensor was located on a bench. The robot using a vacuum gripper picks up the part and places it on the tactile array to determine which side is facing the array. If the wrong side is presented, the robot picks up and turns the part. The part is again placed on the tactile array to determine its center location. With this position known, the robot can then place the diaphragm correctly into the plastic holder. Lord's newer sensor, the LTS 200, incorporates several major changes. Its overall size is reduced to 1.75 inches long by 1.125 inches wide and 0.65 inch thick. The sensing array is 12 by 8 on 0.1-inch centers. Each site is sensitive to loads as low as 0.05 pound and up to 25 pounds perpendicular to the touch surface. It also includes a new module that measures gross normal and sheer forces. The Lord system is shown in Figure 5–12.

Another example of a commercial touch-sensing device is the tactile sensor developed by Barry Wright Corporation of Watertown, Massachusetts. The device is called the TS402 and consists of a touch sense pad on an electronic interface that connects to the robot controller. The active area is 1.56 by 1.56 inch square and contains a 16 by 16 array of sensors located on 0.1-inch centers. The overall size is 2.50 inches wide by 2.60 inches long and 0.35 inch thick. The device uses elastomers that provide data that may be used to determine force, position, and part orientation.

A survey conducted by Case Western Reserve University (Harmon, 1982) of tactile sensing tasks needed in industry indicated that many of the desired operations could be satisfied with rather gross sensing using fewer than 10 sensor elements per square inch and a 4 by 4 array of elements. However, at least 7 percent of the applications did require fine tactile sensing, with more than 10 elements per square inch and larger than 16 by 16 arrays. Slip sensing was also determined to be required in about 53 percent of the tasks studied.

One major result of Harmon's study was that the required tactile spatial resolution required for a wide range of assembly tasks was surprisingly small. Arrays of no more than 8 by 8 pressure-sensitive points with interpoint spacing of 0.1 inch and capable of slip sensing appear to promise considerable capacity. Such arrays, mounted on two-fingered grippers and three-fingered anthropomorphic hands, were concluded to be able to replace humans in 80 percent of the assembly tasks studied.

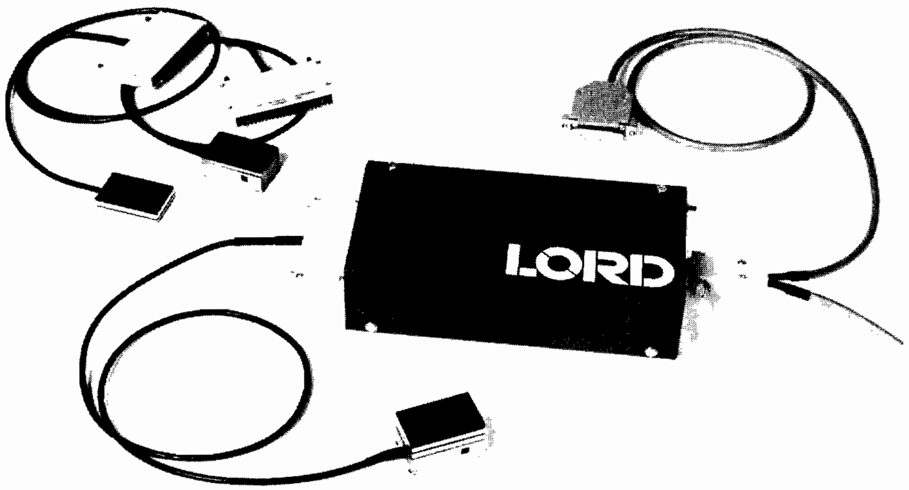


Figure 5–12. Lord Corporation's tactile sensors and interface unit. (Courtesy of Lord Corporation.)

5.7 Sensors for Mobile Robots

Autonomous, self-guiding, mobile robots require special sensors that are not necessary for stationary robots. The purpose of this section is to review the sensor requirements for mobile robots and some solutions for the various problems encountered in using sensor devices. Safety considerations make it most desirable for mobile robots to be equipped with sensory devices that can enable the robot to, for instance, avoid collisions or use sensory feedback information for guidance and position determination and often for target location. These requirements include contact tactile sensors, proximity sensors, local and global position sensors, and level sensors. Many applications for intelligent mobile robots exist. Some of these include industrial carts for material transport, military sentry duties, medical patient care, and domestic duties, as well as lawn mowing and vacuum cleaning.

The "Shakey" robot, developed by SRI International about 10 years ago, is an excellent model for the type of robot we wish to consider (Raphael, 1976). It consisted of a mobile wheeled base, an on-board computer, and several types of sensors, including simple contact switches, a ranging sensor, and a TV camera. Shakey could maneuver through a room, avoiding obstacles and performing simple tasks.

Hans Moravec (1982) of Carnegie-Mellon University (CMU) has developed a camera-equipped mobile robot to support research in control, perception, planning, and related issues. The CMU rover has a cylindrical shape and is about a meter tall and 55 centimeters in diameter. It has a steerable wheel assembly consisting of three independently steerable wheel assemblies. It also carries a TV camera on a pan-tilt-slide mount, as well as several short-range infrared and long-range proximity detectors.

Bart Everett (1982) developed a mobile sentry robot, which he called ROBART, at the Naval Postgraduate School. This robot was designed as a development robot for autonomous sentry applications with emphasis on testing appropriate sensors and their associated interface hardware. ROBART is designed to function autonomously and is equipped with a complex scheme for collision avoidance. Also, the battery condition is constantly monitored by its on-board computer, which can activate a radio-controlled homing beacon on a nearby recharging station when needed. When a low battery condition is detected, the beacon activates and the robot homes in on this station, connects to the charger, and replenishes its energy supply.

Dr. Hall and his students at the University of Tennessee, built two mobile robots. The first, called Micromutt, contained two independent drive wheels, which permitted a zero turning radius, an on-board microprocessor, audio sensors to permit homing, and sonar sensors for collision avoidance. The second robot, called MERV, contained many of the same capabilities but with additional sonar sensors to not only detect an obstacle but also to determine the direction of motion. These robots were built as prototype educational robots and proved an excellent experience for the student designers.

Another class of mobile robots included the popular show and home entertainment types, such as Brains on Board (BOB), which is a 3-foot-tall mobile robot that uses, along with other sensors, an infrared sensor attuned to the wavelength of the human body that permits it to locate people in a room. It is mobile and is equipped with two microprocessors and a voice synthesizer. It is built by Androbot, Inc., of Sunnyvale, California. Other robots that are currently being offered include HERO I by Heath Company, ComRo by Comro, Inc., and RB5X by RB Robot Company. Common features include on-board microcomputers, ranging sensors, light sensors, sound sensors, and voice synthesizers. A commercially available mobile robot vacuum cleaner is also available as an attachment to the RB5X robot and is offered by The Sharper Image.

Several special sensors are required for mobile robots. Furthermore, since the sensors may be used for controlling a robot that may be moving at velocities in the range of 1 to 20 feet/second, high-speed algorithms may be required. Contact switches, proximity detectors, homing signals, light and sound detectors, level indicators, and local and global positioning devices may also be required.

Contact switches placed around the periphery of the robot provide a last line of protection for collision avoidance. The switches could be used to simply stop the robot; however, if they are arranged properly, a new direction of motion can also be inferred that allows the robot to continue on its way. Contact switches that simply stop the robot when it touches an obstacle are used on the popular mail cart robots. Several switches around the periphery are needed to provide the motion control.

Proximity sensors, such as the Polaroid sonar ranging device, may be used to change the robot's path before it encounters an obstacle. A single sensor is of limited use; however, one proximity sensor may be used to slow down the robot before it encounters the obstacle. Three proximity sensors may be used to provide steering information. For example, if an object is detected in front of the robot, side-mounted proximity sensors may be used to guide the robot in the direction in which it has the maximum unobstructed traveling distance. Peripherally located proximity switches may also be used to provide a safety or warning system for the robot, which tells the robot if something or someone is approaching it. Infrared proximity sensors for intrusion detection have been used by Everett for use on security applications for mobile robots.

One of the most important and difficult sensor systems required for a mobile robot is a position location device. Both local and global position information may be required. The accuracy of this information is also very important in determining the control strategy of the robot, since the success and accuracy of the manipulator's task is directly related to the success and accuracy of positioning the robot. Local positioning information may be achieved by implementing shaft encoders on the wheels. These encoders can provide accurate information for short distances; however, wheel slippage and other factors can cause large errors to accumulate over large distances. Therefore, some global method for determining updates to the position may be required.

Several methods can be used for providing position updates. The LORAN system, used in navigation, uses radiofrequency beacons and receivers that determine the time delay from the known location beacons to permit triangularization of position. Unfortunately, the accuracy of these devices is of the order of tens of feet. A local LORAN-type system could also be built; however, it would require active transmitters as well as receivers.

A global positioning system using modulated mercury arc lamps modulated at different frequencies and a scanning optical detector to measure the angles to the known positions was built into a mobile robot cart (Anbe et al., 1972). Other methods for obtaining global positioning may use passive reflectors or targets. The advantage of a passive approach is that the positioning location device can be located on the mobile robot and powered from the same source.

The use of a global positioning device may also require that a map be programmed into the robot's memory, so that a strategy based upon its current position and desired position may be developed. This realization has led some researchers to develop methods for mapping the robot's environment. For example, a ranging device on the mobile robot would permit the collection of range data to objects in the surrounding area and, with further processing, be used to produce a map.

Example: The PEGASUS System. As an example of the type of sensors and controls required for a mobile robot, the PEGASUS system design developed by Dr. Hall and his students will now be described. The goal of this project was to develop a system for the autonomous operation of a commercial Hustler lawn mower. The system was to be interfaced to the lawn mower in such a manner as to provide safe, autonomous control requiring minimum human supervision. PEGASUS contained both training and automatic modes. A remote control override capability would also be used to ensure safe operation by keeping a human to supervise the operation.

During the training mode, the operator was to drive the robot around the perimeter of the field and indicate any critical obstacles. During the automatic mode, the robot would select and implement a mowing strategy. The sensors would permit responses to changes in the environment, such as avoiding obstacles, updating its position, or responding to variations in terrain and avoiding any unforeseen obstacles. Both local and global positioning systems were used. For local control, shaft encoders provided the basic position data. A local imaging device was also used to permit the mower to follow the grass cut line and adapt to small directional changes. A global positioning system would provide periodic updates to correct the positional data.

Several mowing strategies could be followed. The strip strategy consists of translating the mower back and forth across the field oriented along the greatest length of the field. At the end of the field, a 180-degree rotation is executed to permit the next strip to be mowed. The perimeter strategy consists of outlining the perimeter and mowing in ever-decreasing areas until the field is mowed. A sector strategy consists of dividing the field into sectors and using either the strip or perimeter method for each sector. It is also possible to combine the strategies. For example, the operator could outline the perimeter of the field and major obstacles in the training mode, then let the computer sectionize the field into sectors, prioritize these, then mow each using a strip pattern. With any of the strategies, obstacles can be avoided in two ways. In the training mode, the operator could indicate the location of obstacles by moving the mower along the outer boundaries of the obstacles. This would permit mowing around lakes or large flower beds. During the automatic mode, small obstacles, such as trees, could be automatically avoided using the sonar obstacle avoidance. This would free the operator from locating a large number of small obstacles during training. In the automatic mode, the contact switches would also prevent the mower from getting stuck between obstacles, such as between two large rocks or trees.

An integral part of the fail-safe nature of the design was the perimeter contact switches. They provided a last line of defense in obstacle avoidance. The contact switches completely encircled the mower to permit not only stopping the mower but also determining the new direction of motion when an obstacle was encountered. The switches were arranged at a height of 6 inches above the ground and 6 inches from the mower. The switch information provided interrupts to the drive system and caused the mower to stop, back away from, and then proceed around the obstacle.

Shaft encoders fixed to the hydrostatic drive motors monitored the mower's position, translation, and speed. The x and y location of the mower relative to a starting or home position was determined from the angular information provided by the shaft encoders. Some compensation for wheel slippage, such as that detected by an increase in instantaneous velocity increases, could also have been implemented. The hydrostatic drive motors were controlled by linear actuators attached to a hydraulic pump. The two drive wheels were mounted on an axis of symmetry and encoded, providing phase pulses and directional information. Wheel slippage was detected by comparing the instantaneous angular velocity to a running average velocity. Any increase in the instantaneous velocity over the average velocity greater than a threshold was considered the start of wheel slippage. When the velocity difference became less than the threshold, the end of the slippage was assumed. The time interval of slippage and the average velocity permitted an estimate of the true distance traveled.

The major system for obstacle avoidance was a perimeter array of proximity detectors. The sonar system was an assembly of sonar detectors, such as those used on the MERV project. However, this system was designed to detect the location of objects near the mower in all directions. This permitted the implementation of a 20-foot safety zone around the mower, monitored by the sensors that automatically stopped the mower, and disengaged the blade if someone approached it from the rear or sides, as well as directional control for obstacles approached from the front. The design consisted of 12 sonar transducers and four stepper motors. The stepper motors rotated their respective transducers through 270 degrees on the sides of the mower, and slightly less than 180 degrees on the front and rear. Four of the stepper motors were located at the four corners of the mower. Each of these had two sonar transducers, one to detect low obstacles and the other to detect high obstacles. The other four motors were located on the front, sides, and back of the mower and had one transducer each.

For safety reasons, a radio control (RC) unit was used to permit the operator to shut down the system at any time. This system also permitted the operator to remotely drive the mower. A standard RC controller was modified to interface with the digital computer controller and was used in compliance with the Federal Communications Commission (FCC) rules concerning radio control equipment.

Since small terrain variations may often be encountered, a local control device consisting of a line scan camera was used to provide detection of the grass cut/uncut boundary and corresponding guidance signals to the mower. A 256-element camera was used. The camera was positioned to view the front pathway. The signals from the camera were converted to binary, averaged, and used to determine a sectored cut line vector. The basic detection depended on the lighting conditions and height and texture of the grass. Therefore, a processing step to determine the validity of the derived vector was implemented by analyzing the composite histogram of the image. The magnitude of the vector represented the remaining distance along an orientation of the grass cut. The direction of the vector represented the mower's direction with respect to the field.

A global positioning system was used to provide periodic updates to the system position. The absolute global positioning system used an omnidirectional imaging system implemented with a horizontally mounted fish-eye lens. The system was gimbal mounted to provide a constant horizontal orientation, which simplified the computation of position. The position was determined by imaging two or more fixed location markers. The algorithm used is similar to that previously developed by Fukui (1981). His demonstration used a conventional field-of-view lens to view a standard target with a fixed-height camera. The position of the camera was determined from the view angles of the target in the horizontal and vertical directions. This system's arrangement required the horizontal view angles of two targets to determine the position of the camera and, consequently, of the mower. Targets were selected to provide high contrast with the background and distinguishability from each other.

Imaging from a mobile robot involves several special requirements. First, since it is desirable to image while moving, motion blurring must be avoided. Since it is impractical to use a strobe illuminator during daylight, a shutter mechanism with speeds of 250 to 1000 frames per second is required. Also, since a very short illumination time is required, a camera with high sensitivity is necessary. Finally, high resolution is desirable since the overall accuracy is determined to a large extent by the digital resolution.

The gimbal-mounting system in this example also provides sensors that measured the steepness of the grade. If a grade steeper than the drive grade of the mower was encountered, the mower could stop, back away, and attempt another route. The pitch and roll sensors, mounted on the gimbal rotation axes, provided the level information. This feature prevented the mower from overturning. These sensors also provided information as to the validity of the global position information, since global updates were valid only when the camera was horizontal.

The computer control for the PEGASUS system was a hierarchical architecture. A 16-bit 8088/8087 MPX-16 single-board computer was used as the supervisor. This computer was programmed in the FORTH language. The 8087 math microprocessor supported 80-bit floating-point operations and increased the numerical throughput by a factor of 500. Seven Z-8 microprocessors operated as dedicated controllers for the individual subsystems.

The main computer processing unit (CPU) was used to control the overall strategy as well as serve as the global positioning system interface. The sonar sensors, linear array, and remote control system were all controlled by separate Z-8 microprocessors. Each of the two drive motors and shaft encoders was also controlled by a Z-8 microprocessor. One other Z-8 was used as a steering coordinator and was connected to both the motor controllers. A final Z-8 was used as a serial traffic controller. A total of eight microprocessors were used in the controller.

The computer language selected for the master control program was FORTH. This is a high-level language; however, due to the nature of its compilation and execution, it is faster than most high-level languages and can operate at speeds approaching that of assembly language. A FORTH program is made up of strings of commands, most of which are one-word commands. The feature that makes this language unique and so attractive is that the user can create new commands as needed. The building block nature of this language permits the programmer to define new commands by stringing together a series of previously defined words. After each new word is defined, it is compiled and added to the dictionary of previously defined words. The new word is then available for use, just like the words provided with the system. Words are built of lower level, error-free words, and debugging consists of determining whether the lower level words operate together as expected. In essence, a FORTH program is simply a word that has been defined as a series of lower level words.

5.8 Sensor and Control Integration

Since a variety of sensors may be required for an intelligent robot, the problem of integrating the sensory information with the stored information to develop a control strategy is also important. In some cases a single computer may be powerful enough to control the robot. In more complex systems, a hierarchical, distributed computer can be used by the mobile robot or flexible manufacturing system. An executive controller may be used to implement the overall strategy. It communicates to a series of dedicated processors that control the robot functions and receive input from the sensor systems. Sublevels in the hierarchy may also be used for related tasks. A large central microprocessor with high-level language capabilities, connected to smaller microprocessors on a common bus, provides an implementation method for the hierarchical control. The software strategy may then be contained in the master controller, and high-speed actions are controlled by the distributed microprocessors.

The distributed sensor and control system is similar in many ways to the central nervous system in the human. Many actions are controlled by neural networks in the spinal cord without conscious control. These local reflexes and autonomous functions are vital to human survival and are important in discovering how they may be simulated in a robot. The study of these mechanisms in robots may ultimately lead to a greater understanding of how we function as humans.

Questions

1. Discuss the concept of common sense and how it relates to humans and robots. Is common sense desirable? When and why?
2. Discuss the question of liability for the actions and possible consequences arising from the use of mobile robots.
3. Compare the use of visual and tactile sensors for a variety of tasks, such as recognizing an object by its shape, color, material, or texture.