MACHINE VISION: A multi-disciplinary systems engineering problem.

Dr. Donald G. Bailey

Visiting Assistant Research Engineer
Department of Electrical and Computer Engineering,
University of California, Santa Barbara, California, 93106.

ABSTRACT

The successful application of image processing to industrial inspection and measurement requires the combining of a number of techniques from different disciplines including optics, electronics, computer hardware and software design, and mechanical engineering. The main subsystems of a typical machine vision system are the image capture subsystem, incorporating sensors and image digitization, the image processing subsystem, where the required information is extracted from the image, and the control subsystem, which uses the information obtained to control a task or activity. In recent years, the commercial availability of general purpose modules such as cameras, frame grabbers, robots, and their controllers has simplified system design considerably. The advent of very large scale integrated circuit technology is broadening the range of applications of machine vision by enabling fast hardware to be designed.

1. INTRODUCTION

From time immemorial, man has been fascinated by his environment. Sight was one of the most important senses used by man as he has sought to manipulate the things around him, and understand the laws governing their behaviour. As man developed implements and techniques to make it easier to manipulate his environment, he also developed machines and methods to assist and improve his vision. Now, man is been adding vision to his machines to increase their power and flexibility.

It was not until the advent of third generation computers in the 1960s that digital image processing techniques started becoming widespread. This was because the processing speed and storage required for practical image processing algorithms was unavailable before then. Since then, the field of digital image processing has grown considerably, being studied and applied in such diverse fields as geography, geology, resource management, forestry, agriculture, oceanography, medicine, space exploration, industrial automation, nondestructive evaluation, biology, physics, electronics and forensics. In fact just about any task requiring vision may be a candidate for image processing.

All of the applications of digital image processing may be split into three broad areas. The first is image coding, where the volume of data contained in an image is reduced for storage or transmission. Another area is the improvement of pictorial information for human interpretation. This is also called computer assisted vision. The final application area is computerized vision or machine vision.

Machine vision has been defined as "the automatic acquisition and analysis of images to obtain desired data for interpreting a scene or controlling an activity"[1]. In an early application, computerized vision was used in laboratory situations to automate routine analysis; now it is being applied to industrial and manufacturing tasks.

2. APPLICATIONS OF MACHINE VISION

The two main application areas of machine vision are inspection and manipulation. Of the two areas, inspection is by far the larger. A recent survey showed that 90% of the machine vision market is in inspection and quality control[1]. Inspection tasks may be split into two broad categories[2]: those in which decisions are based on highly quantitative measurements, and those in which the measurements are of a more qualitative nature.

In quantitative, or measurement tasks, machine vision provides a powerful general purpose tool. The use of solid state sensors enables accurate noncontact measurements to be made rapidly at potentially low cost. The type of information that may be of interest in a measurement task is the length, area, or position of an object.

In qualitative inspection tasks, machine vision is used to emulate the visual inspection performed by a human inspector. In manufacturing processes, there are many instances where workpieces of several types are packed together in containers, or randomly arranged on conveyors. Machine vision may be used to select a specific workpiece, to determine the number of pieces of each type, or to sort the pieces into their different types. Another application of inspection is the detection of functional or cosmetic defects. These defects may be flaws or errors incurred during part fabrication or handling. Functional defects are those which may cause malfunction or improper operation if the component is used. Cosmetic defects reduce the subjective quality, without impairing performance.

The majority of inspection tasks are menial and have a high boredom factor. In fact the very nature of the task that causes boredom is one of the prerequisites for the successful application of image processing techniques since such tasks tend to be well defined and relatively simple. Machine vision is able to perform these tasks consistently and reliably allowing 100% inspection of all components produced[3]. In fact, some companies which subcontract much of their work are also aiming for 100% inspection of all incoming components. This minimizes waste and increases production by ensuring that the final product does not incorporate any faulty components.

Although there is a wide variety of techniques used among the different inspection systems that have been applied commercially, they all use a priori knowledge of the specific problem[3]. In industrial situations, where the inspected components are usually man made, regular geometry, such as straight lines and right angles, greatly simplifies the inspection task.

Machine vision is also playing an increasing role in manufacturing or assembly environments. The main areas for the application of machine vision to manipulative tasks are in component acquisition, fabrication, and assembly[2].

In the area of component acquisition, machine vision gives increased flexibility. The first generation of automatic machines required a very ordered environment: that is, the objects that were being manipulated had to be at the right place, in the right orientation, at the right time. With vision, this requirement can be relaxed, and expensive jigs or other handling equipment is no longer essential[4], often resulting in a considerable cost reduction.

In the past, a large number of important processes could not be automated because of the small number of components manufactured in each batch. The small batch size meant that the cost of specialized jigs or dedicated production machinery is out of proportion to the return obtained[2]. The use of vision overcomes this restriction by giving added flexibility in component and tool manipulation.

Visual feedback is useful in assembly tasks for alignment or mating of components, since it can be used to compensate for a buildup of positional error in the components being manipulated. When an assembly task is performed by humans, there is an implicit inspection of the components. If a component which has an obvious defect is encountered, it is usually placed to the side and not incorporated into the final product. Inspection is being incorporated into the manufacturing and assembly stages in order to maintain, and improve the overall quality of the end product[2].

3. AREAS OF EXPERTISE REQUIRED

Techniques from a wide range of disciplines are involved in the construction of an image processing system for a particular task. The areas of expertise are discussed in detail in the following sections.

3.1. OBJECT ILLUMINATION

Correct lighting is essential to simplify the processing of the captured image. It is difficult, and often impossible to compensate for poor image quality caused by poor object illumination. Some of the problems that may be encountered are: a nonuniformity of illumination across the image; shadows or glints; and sometimes poor contrast. All of these problems can introduce false features, or cause others not to be detected[5]. This may result in the acceptance of a defective component, or the rejection of an acceptable component.

In some circumstances, special lighting techniques, sometimes called structured lighting,[6] may be used to simplify the resultant image processing algorithm[4]. Some of the commonly used lighting methods will be described.

The simplest form of lighting is back lighting. This is most useful when only shape information is required. If the objects being imaged are opaque and not overlapping, they may be located by a simple global threshold. Back lighting is also useful when imaging transparent objects. Related to back lighting is dark field illumination. The light is behind the object, but not in the field of view of the camera. When there is no object present, the image is dark. When a translucent object is placed in the field of view, it is seen as light against the dark background.

Direct front lighting may be used to avoid shadows. This is very useful when the object being imaged has a rough surface. When imaging the surface of integrated circuits, direct lighting is required to show the different layers as different colors through interference. Diffuse lighting has properties similar to direct lighting, with the added advantage of minimizing glints caused by highly reflective surfaces.

Side or glancing illumination is useful for highlighting surface texture or irregularities, especially on highly reflective surfaces. An example would be detecting scratches and dents on sheet metal[7].

The most commonly used structured lighting method is probably striped light. It may be used to detect the presence of a part[8]. Through the use of triangulation between the light source and sensor, information about the range may be inferred.[9]

It is important to tailor the lighting scheme to the specific object or range of objects being imaged. One advantage that industrial applications have over general scene analysis is that there is considerable control over the environment. This means that extraneous objects may be kept out of the field of view, and the lighting may be controlled more effectively.

3.2. IMAGE CAPTURE

The process of obtaining a computer representation of the scene or object is image capture. There are two stages involved in image capture: imaging and digitization.

In the first, a sensor converts an optical intensity pattern into an electrical signal. Consider, for example, a two dimensional solid state sensor. A lens casts an image onto the sensor which converts the image to a two dimensional charge pattern. This is achieved by the incoming photons liberating electrons, which are then trapped or accumulated in special areas on the sensor[10]. The charge that accumulates in each area is therefore proportional to the light falling on that area, and the integration time, or the duration of the incident illumination. A typical integration time for sensors used in industrial applications is 20 ms (the main reason for this is that readily available cameras are television compatible). After the image has accumulated on the sensor, the two dimensional charge pattern is read out. This involves converting the charge associated with each picture element, or pixel, into a voltage. Since with most devices it is not possible to read out individual pixels at random, the device generates a serial train of discrete voltage pulses.

A minor knowledge of optics is necessary for the selection of appropriate lens systems. This involves determining the correct focal length and aperture to obtain sufficient resolution and depth of field to adequately image the object of interest. With very small objects or when very fine resolution is required, the use of a microscope may be appropriate.

There is a wide range of sensor types, both one and two dimensional, that may be used for machine vision. Some form of mechanical scanning is required with one dimensional sensors to form the second dimension. This may be in the form of a scanning mirror or lens, or by using the uniform movement of the object being imaged (for example along a conveyor[8]). In applications involving measurements, solid state sensors are preferable to tube sensors. With solid state sensors, the individual sensing elements are in fixed precisely known positions allowing the relationship between object points and sensor elements and pixels to be known exactly. This is not the case with tube sensors where the image is obtained by analog scanning. Nonlinearities, and drifts caused by temperature changes or component aging mean the scanning cannot be precise. In industrial environments, where the sensor may be subjected to vibration or being knocked, solid state sensors tend to be more reliable.

This discussion so far has concentrated on optical sensors. Similar principles apply to other types of sensors such as acoustic transducers. Recently, laser range finders have been used to make it easier to obtain three dimensional information. A topic currently being researched is how to efficiently and effectively combine the images formed by different types of sensors.

In the second stage of image capture, the electrical signal from the sensor is digitized and transferred into computer memory, usually using some form of direct memory access (DMA) device. A DMA device which captures an image when requested by the processor is known as a frame grabber. The frame grabber ensures that the correspondence between the pixel on the sensor and the location in memory is known, making the system suitable for mensuration tasks. The resultant array of numbers representing the scene or object is known as a digital image. Depending on the type of sensor used, the numbers may represent object reflectivity, range, or some other property.

If a television compatible sensor is desired, there are many frame grabber boards commercially available for most types of computer systems. However, for nonstandard sensors, either hardware modifications must be made to such boards, or the image capture system must be developed from scratch, requiring extensive electronic design - both analog and digital.

3.3. ALGORITHM DEVELOPMENT

Once an image has been loaded into the computer, it may then be processed. This is the application of a series of actions or operations to an image which leads to a desired result. This result may be the measurement of a critical dimension; a description of the object; or the classification of the object into one of several categories (for example grading). There are three main requirements to successfully develop an algorithm for a particular machine vision task.

The first is an interactive system for image processing which has available a large number of operations, and the ability to display intermediate results. This is necessary since algorithm development is largely a heuristic process[11]. The task may be broken down into a sequence of broad steps but there is little or no underlying theory that may be used to determine exactly when a particular operation should be used. In practice the algorithm is developed by trying out an operation. If the result of the operation is not satisfactory, another is tried in its place. This procedure is repeated until the desired result is obtained. Thus algorithm generation requires much backtracking and much interaction in order to determine a solution for the task being examined.

The second requirement is that the person developing the algorithm has extensive background and experience in image processing, particularly in the type of application for which the new algorithm is being developed. The trial and error approach described in the previous paragraph is simplified and in practice, can only be performed by an appropriate specialist. The specialist must have detailed knowledge of the properties of each of the available operations, experience in determining exactly which operations would be beneficial in a given situation, and the ability to choose the operation best suited to the task at hand.

Finally, it essential that the speciaiist has detailed knowledge of the task to which the image processing is being applied[12]. This knowledge is important since it increases the probability that the right measurements will be made and that the algorithm will be robust. Without such "problem knowledge", invalid assumptions may be made about the nature of the task. There is always a danger that a resultant algorithm will work well for the specific cases that were used to develop it, but not be sufficiently general or robust to be of practical use.

One of the difficulties of developing algorithms is the problem of data reduction. An image contains a large amount of information, a lot of which is irrelevant to the problem being investigated. Consider for example the task of grading an object for acceptance or rejection. The input image may be 128 x 128 pixels of 8 bits each, a total of 16 kilobytes. From this, a single bit needs to be determined; whether the object is acceptable or not. This corresponds to a data reduction of greater than 130,000.

Another problem encountered when developing algorithms is that images are in fact two dimensional representations of three dimensional objects. Information about the third dimension is important when tracking moving objects, picking up an object from an arbitrary position and orientation (for example from a bin), or inspecting a truly three dimensional object (one that is not flat). This information may be obtained through triangulation by using multiple sensors, or the use of special lighting techniques, or some other range finding technique (laser or sonar)[8].

The development of algorithms for industrial inspection and measurement can be broken down into six steps[13].

1) DESIGNATE THE TYPE OF RESULT REQUIRED. The type of result required determines the kind of information that needs to be extracted from the image and thus governs the formulation of the overall strategy. Broad areas of application include:
- making quantitative measurements (size, length, area, position, intensity)
- checking for integrity or completeness (parts, features, burrs, cracks, etc)
- inspecting cosmetic or surface properties (stains, colors, scratches, etc)
- sorting.

2) OBTAIN A REPRESENTATIVE IMAGE. During the initial algorithm development process, the image should be as simple as possible while still being representative of the task being examined. As the algorithm is refined, a wider variety of images can be used to test the algorithm and determine its limitations. The most important aspect of the image capture step is the use of appropriate lighting.

3) PREPROCESS THE IMAGE. Preprocessing is often used to enhance the information required from the image while suppressing information that is irrelevant. Preprocessing can frequently ensure that the more complicated operations of later stages are successful. Which of the following preprocessing operations are necessary depends strongly on the broad strategy selected at the outset, and the images that are being preprocessed. Typical preprocessing operations are used for:

Compensating for deficiencies in the image capture process. Due to various physical constraints, the image captured may not be ideal. In compensating for deficiencies, we may use various preprocessing operations to
- remove background or thermal noise
- correct for nonlinearities in the image transfer function
- correct for distortion caused by optics or camera angle
- compensate for deficiencies in illumination.

Normalizing the image. This step permits making accurate comparisons between images. To obtain a properly normalized image we may need to
- normalize intensity (contrast expansion, histogram equalization) to compensate for variations in the lighting or reflectivity of the objects being imaged
- normalize position (centering, orientation) for registration with a template or model.

Filtering. Local filters may be used to
- smooth noise
- enhance or detect edges or other features
- classify objects through matched filtering.

Thresholding. This step classifies pixels according to their intensity and usually results in a binary image. Thresholding is used to
- separate object and background
- detect features.

Segmenting. The image is divided into meaningful parts or regions with common properties. The four main segmentation methods use operations to
- detect and link edges
- track boundaries
- threshold the image
- grow regions.

4) EXTRACT THE REQUIRED INFORMATION. The information extraction stage transforms the data from a series of fixed intensities to a more descriptive symbolic form. The image data consists of a large number of pixel values, each carrying very little information. This is converted to a form where there are few numbers, but each number carries significant information about the object or scene being imaged. The types of operations used for information extraction include:

Feature extraction. Features of the object are extracted that contain information useful for discriminating against other objects that may be encountered. This information is often a small set of numbers and may represent size, position, shape, etc.

Measurement. In some applications, the information required is a measurement, while in others the measurements made are used for classification purposes. Quantities that may be measured include length, area, and intensity. When absolute measurements are required, the system should be calibrated by executing the algorithm on a similar object of known dimensions. These measurements may then be used to calibrate the result by scaling.

Pattern identification and classification. At this stage, the object being imaged is classified into one of two or more categories based on the information extracted. This step is used for defect rejection, sorting, etc. Two types of classification involve
- partitioning the various parts of the image according to feature vectors
- comparing the image or image features with a template.

These information extraction steps are very application dependent. Appropriate features must be selected for the discrimination indicated during the initial stages of the algorithm development process.

5) POSTPROCESS THE IMAGE OR RESULTS EXTRACTED. In some circumstances, postprocessing may be required to convert the extracted information into the desired form. This may require
- displaying resultant images
- manipulating data for robot control
- presenting special information
- recording a tally of prodution or defect rates.

6) REFINE THE ALGORITHM. Since algorithms are initially developed using a single image, or a small range of representative images, they require testing. For this a large number of images is necessary, especially images which are marginal as far as any classifications are concerned. Various complicating factors and special cases may be introduced and the algorithm refined to accommodate them. In applications where speed is critical, the processing bottlenecks may be identified and alternative operations selected.

Recently there has been an influx of artificial intelligence techniques into the feature extraction and classification stages. This is particularly the case when outdoor scenes are imaged where uncontrolled lighting and a large number of different objects are encountered. Symbolic processing may be used to infer what an object is, based on positional or similar relationships. Research is also underway on using similar techniques to simplify the algorithm development process too.

3.4. MANIPULATION

In industrial situations, the purpose of machine vision is to control an activity. This requires using some form of mechanical hardware to act as a result of the information extracted from the image. The actual form of the hardware depends on the application. In a sorting or inspection environment, it may be a simple manipulator. If an inventory control system is incorporated, the data extracted from each image may be sent to another computer. When the machine vision system is used for a manufacturing or assembly task, the control hardware may be a robot arm or computer controlled machine. The common factor in all these cases is that the data extracted from the image is acted upon automatically.

The design of the manipulation hardware, and the interfacing of it to a vision system is quite a difficult mechanical engineering problem[14]. This process has been simplified to a certain extent in recent years by the commercial availability of robot arms and controllers.

3.5. IMPLEMENTATION

One of the problems in implementing the algorithm, particularly in an industrial environment, is the limited time available to perform the task. A typical time requirement for an industrial inspection task is one second per object. This limits the range of applications that may be handled by machine vision at the present time. Another problem encountered when designing a machine vision system for a specific task is determining the optimum hardware and software combination required.

The image processing system that is used to develop an algorithm for a particular task is very rarely the same system used to implement the resultant algorithm. This is because the algorithm development system is usually implemented on a general purpose computer for increased flexibility. While developing the algorithm, speed is not important provided the system is sufficiently fast to be interactive. However, when the algorithm is implemented as a machine vision system, speed is critical. Most algorithms can be made faster through the use of dedicated hardware.

Implementing individual operations in hardware, either discretely or through the use of very large scale integrated circuits (VLSI) can give speed increases of up to three orders of magnitude[13]. This speed results from exploiting the parallelism contained within the operation. For example local filtering applies the same operation to each point in an image The disadvantage of using dedicated hardware is limited flexibility. If the algorithm needs to be changed because of changing needs, the system may have to be scrapped and a completely new system constructed. For this reason, a machine vision system should be very modular. This gives systems using dedicated hardware a pipeline type of architecture.

Pipelined systems exploit the fact that most algorithms consist of a temporal sequence of independent operations (temporal parallelism). The algorithm is implemented by a corresponding sequence of independent processing units, with the output of one processing unit connecting directly to the input of the succeeding processor. When processing images, such an architecture is most efficient when each of the operations in the algorithm requires input only from a small local neighbourhood in an image for each pixel produced in the output image.

An alternative architecture is to use parallel multiprocessor arrays. Such arrays exploit the fact that most operations perform the same function independently on a large number pixels in an image, or elements in a data structure (spatial parallelism). The image or other data structure is split among a large number of processors each of which perform the same operation in parallel with the others. Parallel processing is also most efficient when the operations require data only from a local region since this minimizes any communication overhead. It is also more suitable than a pipelined architecture for some of the high level information extraction operations. The individual processors in a multiprocessor array may be general purpose microprocessors or processors designed especially for image processing.

4. CONCLUSIONS

The range of applications of image processing has been developing rapidly as advances in both hardware and software have been made. Problems which were impractical in the past because of time constraints are now becoming practical. These advances will continue to be made as faster and cheaper computing becomes available. The advent of VLSI in recent years is providing a significant boost to machine vision by enabling fast hardware to be constructed. The commercial availability of such modules as imaging cameras, frame stores, robots and their controllers simplifies system design and development. Since machine vision (as opposed to image processing in general) is an emergent technology, similar advances in other areas may be expected in the near future.

Until now, the development of algorithms for image processing has been as much of an art as a science. This state of affairs is likely to continue for some time, although it may change as a result of continuing research into human vision. The application of artificial intelligence techniques may simplify the algorithm generation process by reducing the amount of specialized expertise required.

From this discussion, it should be evident that the design of a machine vision system for a particular application is not a task that can be undertaken lightly. It requires the combination of experience from a range of disciplines including lighting, optics, electronic design (both analog and digital), image processing, software engineering, artificial intelligence, computer design, and mechanical engineering. Applying image processing techniques to any task is really a complex systems engineering problem.

5. REFERENCES

1. Schaffer G., "Machine vision: A sense for computer integrated manufacturing", American Machinist, 128(6):101-129 (1984)
2. Rosen D., "Machine vision and robotics: industrial requirements", in "Computer vision and sensor-based robots", Editors Dodd G.G. and Rossol L., Plenum Press, New York, pp 3-22 (1979)
3. Kruger R.P. and Thompson W.B., "A technical and economic assessment of computer vision for industrial inspection and robotic assembly", Proceedings of the IEEE, 69:1524-1538 (1981)
4. Gevarter W.B., "An overview of artificial intelligence and robotics", Vol 2: Robotics, US National Bureau of Standards, Washington (1982)
5. Hodgson R.M., McNeill S.J., and Bailey D.G., "Image processing goes to work", Proceedings of IPENZ Conference, Napier, New Zealand (1984)
6. West P.C., "Machine vision in practise", IEEE Transactions on Industrial Applications, 19:794-801 (1983)
7. Lippincott H. and Stark H. "Optical - digital detection of dents and scratches on specular metal surfaces", Applied Optics, 21:2875-2881 (1982)
8. Holland S.W., Rossol L., and Ward M.R., "Consight-1: a vision controlled robot system for transferring parts from belt conveyors" in "Computer vision and sensor-based robots", Editors Dodd G.G. and Rossol L., Plenum Press, New York, pp 81-100 (1979)
9. Jarvis R.A., "A perspective on rangefinding techniques for computer vision", IEEE Transactions on Pattern Analysis and Machine Intelligence, 5:122-139 (1983)
10. Hobson D.F., "Charge Transfer Devices", Edward Arnold, London (1978)
11. Rosenfeld A., "Image analysis: problems, prospects and progress", Pattern Recognition 13(1):3-12 (1984)
12. Vogt R.C., "Formalized approaches to image algorithm development using mathematical morphology", Vision '86 Conference Proceedings, pp 5-17 to 5-37 (1986)
13. Bailey D.G., "Hardware and software developments for applied digital image processing", PhD thesis, Department of Electrical and Electronic Engineering, University of Canterbury, New Zealand (1985)
14. Luh J.Y.S., "An anatomy of industrial robots and their controls", IEEE Transactions on Automatic Control, 28:133-153 (1983)