THE ROLE OF MACHINE SENSING IN CIM

Robert U. Ayres
International Institute for Applied Systems Analysis
Laxenburg, Austria

Jeffrey L. Funk
Carnegie-Mellon University
Pittsburgh, PA 15213, USA

RR-89-13
December 1989

Reprinted from *Robotics and Computer-Integrated Manufacturing*,
Volume 5, No. 1, pp. 53–71, 1989.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
Laxenburg, Austria

Reprinted with permission from *Robotics and Computer-Integrated Manufacturing*, Volume 5, No. 1, pp. 53–71, 1989.

Printed by Novographic, Vienna, Austria

PREFACE

This paper has been written in several versions. The first version was written in 1985 based on joint research carried out by the two authors at Carnegie-Mellon University. Later, in summer 1986, Professor Ayres took over the project on Computer Integrated Manufacturing (CIM) at IIASA. On returning to Carnegie-Mellon in the fall of 1987 he rewrote the paper, to reflect some of the insights gained from the CIM project during 1986–87. Thus, it is appropriate to include it as part of the publications of the CIM project.

J. RANTA, *Leader*
Computer Integrated Manufacturing

F. SCHMIDT-BLEEK, *Leader*
Technology, Economy, and Society Program

Robotics & Computer-Integrated Manufacturing, Vol. 5, No. 1, pp. 53–71, 1989
Printed in Great Britain

0736-5845/89$3.00 + 0.00
Pergamon Press plc

• *Paper*

THE ROLE OF MACHINE SENSING IN CIM

ROBERT U. AYRES and JEFFREY L. FUNK*

Department of Engineering and Public Policy, Carnegie-Mellon University, Pittsburgh, PA 15213, U.S.A.

Computer-integrated manufacturing (CIM) is still an objective for the future rather than a present reality. Mechanical integration was introduced over 60 years ago by Ford to increase output and cut unit costs, but at the cost of making high volume production extremely inflexible. The answer is apparently "computer integration" with multi-purpose machines linked together by digital communications networks and adaptively controlled by computers.

In practice this means replacing flexible human workers with high quality sensory interpretative abilities as decision-makers by "smart sensors" with artificial intelligence. However, adaptive controls are severely constrained by the capabilities of existing sensors and interpretative computer software, especially the latter. Most existing sensors are narrow-band, producing at most a few bits of data per second for control purposes. This provides enough information for a small class of machine control decisions, but is insufficient for part recognition, part orientation, or quality inspection.

CIM also means that workpieces (components and work-in-progress) will have to communicate with machines, as machines will have to communicate with each other. Thus, the true economic significance of recent breakthroughs in machine vision/taction is that they will finally unlock the door to CIM, or "5th generation" automation. It is argued that the economics of machine vision/taction should not be assessed in the narrow context of specific tasks in direct competition with human workers but as the hitherto missing links that will permit all the elements of the factory of the future to communicate effectively with each other so as to function as an organism rather than as a set of independent cells.

1. INTRODUCTION

The main purpose of this paper is to present a technology forecast for flexible CIM in a broad economic framework. Technological developments do not occur in a vacuum, especially not the technology in question here. The adoption of technological improvement is clearly driven by economic forces. In turn these depend on progress in technology, though the feedbacks are slower and more diffuse.

An important feedback between technological and economic factors is the inflexibility with respect to product the changes of conventional hard-automated mass production facilities. The advanced mechanical integration introduced in the 1950s and 1960s by U.S. auto and appliance manufacturers sharply raised the cost of product design change.[1] This factor made it difficult for U.S.-based mass-producers to compete in a dynamic marketplace on the basis of introducing continuing improvements in product quality or performance. The problem became more acute as U.S.-based firms were displaced by Japanese and other east-Asian firms as low-cost international producers. U.S. manufacturers are no longer insulated from foreign competition in their domestic markets, as was largely the case until the 1970s. A more flexible manufacturing technology that would permit an accelerated rate of product innovation without making the production facility obsolete has therefore become increasingly necessary to survival.

2. THE PROBLEM OF COORDINATION AND CONTROL IN A FLEXIBLE MULTI-PRODUCT PLANT

The work plan in any factory can be schematically represented as a hierarchy of basic part manufacturing units feeding parts to subassembly stations

*Present address: Westinghouse R&D Center, Pittsburgh, PA, U.S.A.

Acknowledgement—this research was supported by the National Science Foundation under grant (MEA-8205437). We are particularly grateful to William Spurgeon of NSF for his support and guidance.

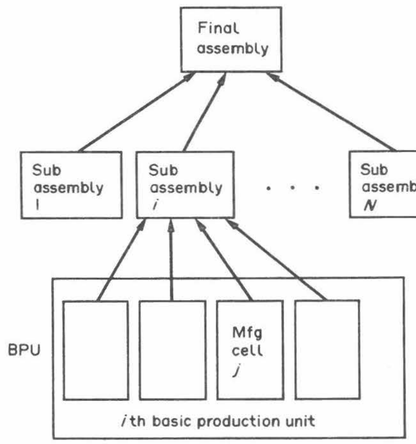


Fig. 1. Shop organization.

and thence to final assembly. The basic scheme is illustrated in Fig. 1. Suppose the "final" product (say an auto engine) consists of N different subassemblies, some of which are used singly while others are used in multiples. (For instance, a 4-cylinder engine requires only one crankshaft, but four piston subassemblies, four valve subassemblies, four spark plugs, etc.) The ith subassembly, in turn, consists of m^i different parts, some of which are used singly while others are needed in integer numbers. Assuming, for the moment, that no parts are common to more than one subassembly, the final product requires M different parts where

$$M = \sum_{i=1}^{N} M_i. \tag{1}$$

Each of these parts must be produced by a distinct group of machines. In large volume production plants, these are typically linked together in a rotary or linear sequence by an indexing workpiece transfer system of some sort. In lower volume facilities parts may be transferred by belts, carts or even by hand.

In the extreme case of a cell dedicated to one variant of one part, the effective rates of these machines must be identical since each machine does one and only one operation on each workpiece at a pre-specified rate before passing it to the next machine. If more than one variant of the part type is to be produced, a more flexible parts transfer system may be necessary. Such a group of machines, linked together and dedicated to making a single part or variants of a single part, can be called a manufacturing cell or a flexible machining system (FMS). Machine layout in a cell depends on the degree of specialization. For a very specialized cell for a single part, where machines are linked by an indexing transfer line, the layout is likely to be linearly sequential (parallel) as shown in Fig. 2(a). For less specialized cells with more flexible transfer systems, machines might be grouped for various other sequences of operations, as in Fig. 2(b).

For the ith subassembly, there will be M_i different cells operating in parallel to produce parts that must be brought together at an assembly station. Just as the effective rate of all machines in a given cell must be the same (for a given part), all the cells feeding parts to a given assembly station must do so at rates that are fixed in relation to each other and based on the number of copies of each part in the subassembly. The group of manufacturing cells producing parts destined for a given (sub)assembly station can be called a cell-cluster. For a complex product, such as an auto engine or the car itself, a number of cell clusters may feed sub-assemblies to a higher order assembly plant. Again, the product design dictates

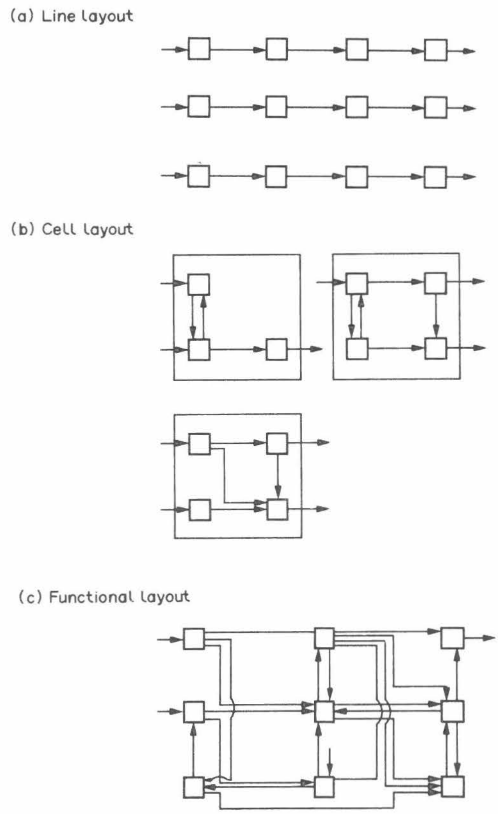


Fig. 2. Three types of machine layouts.

relationships between the operating rates of the various clusters, which, in turn, dictate the operating rates of the individual cells.

A key characteristic of the synchronous sequential mechanical transfer system is that if any machine must be stopped for any reason, all of the machines stop. Any tool change, machine breakdown, blockage or jamming problem halts the whole line. Even though such events are comparatively rare, when there is a large number (15–100) of machines linked serially, it is difficult to achieve a high overall utilization rate for the equipment. In addition, if any workpiece is damaged in transit through the cell, it must either be removed—thereby unbalancing the line—or continue through the later steps in the sequence resulting in a waste of material and machine time.

The coordination problems of mechanically synchronizing a number of such transfer lines together with an automatic assembly system are obviously much greater. In practice, the individual manufacturing cells are usually 'buffered" by intermediate storage of finished or semi-finished parts; the more such storage buffers there are, the less a perturbation at one location propagates disturbances through the system. But parts storage is costly, both in terms of capital tied up in partially completed workpieces, and the need for investment in specialized storage devices (usually toploaded "towers" with a gravity-driven spiral track or chute). In some cases, pallets or magazines are needed for workpiece transport to retain their physical orientation while awaiting mechanical loading of the machine at the next work station. Evidently, there are significant costs to the use of intermediate storage in a factory. The "just-in-time" system pioneered by Toyota and widely implemented in Japan has proven to be highly cost-effective, partly because it imposes tighter quality control standards on the feeder input streams.[19]

The discussion of the last three paragraphs applies to the case of a factory dedicated to a single model of a single complex product. The coordination problem becomes much more complicated when the plant must produce a number of different models or variations of the product or a number of different products with variable and uncertain production runs. If the designs are allowed to evolve over time, the coordination problem is still more complex and difficult; dedicated machines linked by mechanically synchronous transfer systems are not applicable at all in such an environment because each manufacturing cell must be somewhat flexible in terms of its input requirements, operating rates, and output specifications.

In the case of a dedicated line (or cell), each machine tool is specialized to a single operation for which it can be optimized by design. Tool changing is thus minimized; tools are replaced at fixed intervals based on precalculated useful life. Specialized jigs and fixtures (holding devices) are permanently installed on each machine, corresponding exactly to the specified shape of the workpiece and specified requirements of the operation. Machines are designed to execute a fixed sequence of motions at a single optimum speed, to minimize the costs of each specific operation. On the other hand, in a "flexible" cell the individual machine tools must be capable of operating at various turning and cutting speeds, angles of attack, cutting depths, and so forth. Coordinating the machines in a flexible cell with the parts handling system to achieve maximum output becomes a formidable challenge. Indeed, realistic cases are mathematically intractable: exact solutions cannot be computed. Such cases can only be analyzed by means of simulation and crude approximations.[37]

In a traditional "job shop" producing custom prototypes or small batches, the coordination of workflows to maximize productivity is carried out primarily by the shop foreman using information gathered from the individual machine operators and his own accumulated experience. In Japan this technique is known as Seiban. In practice, however, the complexity of optimization results in little more than coordinating and avoiding major bottlenecks. Usually, each machine is independently set up to carry out one (or more) operations on a certain number of workpieces based on the number of items in the batch. To be "on the safe side", the foreman is likely to order extras of each part to accommodate mistakes or faults. The partially completed work pieces go into intermediate storage—most likely a bin—while the machine is set up again for another operation. As the overall workload permits, the foreman eventually assigns another machinist and machine to set up and run the next operation or operations in the sequence. Fairly high machine and worker utilization can be achieved by completely separating each operation from the next in sequence, but at the price of stretching out in-plant transit time for each batch and carrying a large inventory of unfinished parts. In most job shops, workpieces are actually being worked on a very small fraction of the turnaround time between receipt of order and delivery to the customer.

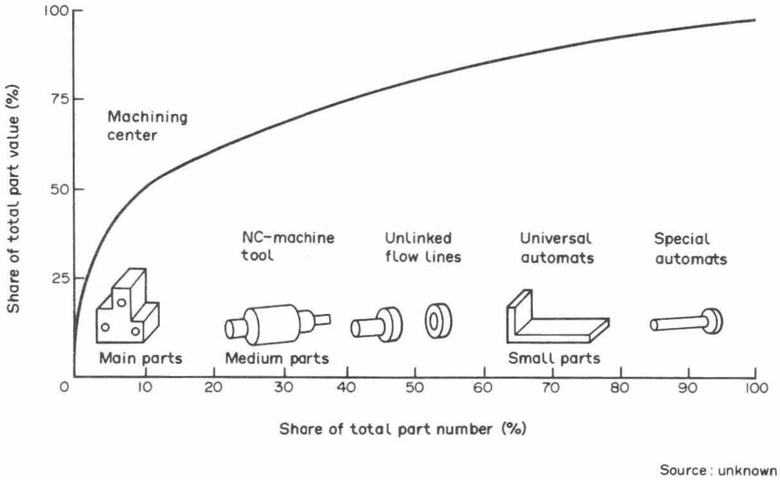


Fig. 3. Group technology and manufacturing systems.

In such a shop the machines are likely to be laid out by function (e.g. drill presses, punch presses, lathes, milling machines, etc.) as illustrated in Fig. 2(c). An important step forward in shop layout and scheduling was the classification of parts into families by shape, known as group technology (GT). Such classification systems are an important tool for deciding how many of each part type can best be produced (i.e. on what machine). Typical relationships between geometrical complexity and choice of machine-type are shown in Fig. 3.

For shops producing in larger and more predictable batches, it is possible to increase output rates by utilizing specialized jigs and fixtures (which must be custom-made for each job), bigger faster machines with numerical controls, and more elaborate parts handling systems for transfer loading and unloading the high speed machines. But these aspects of mechanization leave the above-mentioned co-ordination and control problem untouched. Indeed, as the investment in high performance machines and parts handling equipment rises, the importance of maximizing the joint output of the group of machines operating in sequence on a given part rises.

The small general-purpose manually controlled tool in a typical job shop can be idle much of the time because the machine is comparatively cheap; the major cost is the skilled labor. To utilize high cost skilled machinists efficiently, it is important that no machinist should ever have to wait for a machine to become available. But as individual machines have increased in power and precision, they have also risen sharply in cost. In a plant using expensive multi-axis high production rate tools, the machines must be utilized more efficiently. One of the first ways to achieve higher levels of machine usage is to reduce set-up times by adopting NC or CNC.

The role of the skilled machinist is thus shifted gradually from that of machine operator to that of general supervision and set-up. Once the program is prepared and calibrated,* an NC machine tool can machine complex shapes at a much higher rate than its manually controlled predecessor. In fact, CNC can increase output per machine by up to a factor of 5 although the average improvement is probably closer to half that much.

Since 1954, the trend in discrete parts manufacturing has been unmistakably toward increasing individual machine capability and internalizing more and more machine-control functions resulting in a further reduction of the role of the human operator. In effect, the machine tool has become more flexible and able to take over a greater number of decisions, beginning with automatic actuation and "stop" conditions. An obvious extension of capability is toward automatic tool-changing; whereas in a dedicated machining cell or hard-automated synchronous transfer line tool-changing can be scheduled in advance by the designer, this is not the case where a group of multi-purpose machines is producing a wide

*Unfortunately, NC machine programs are not, in general, interchangeable among machines due mainly to variations in the amount of "give" (or looseness) in the joints of each machine. This, in turn, depends on the age and history of the machine, as well as its initial characteristics.

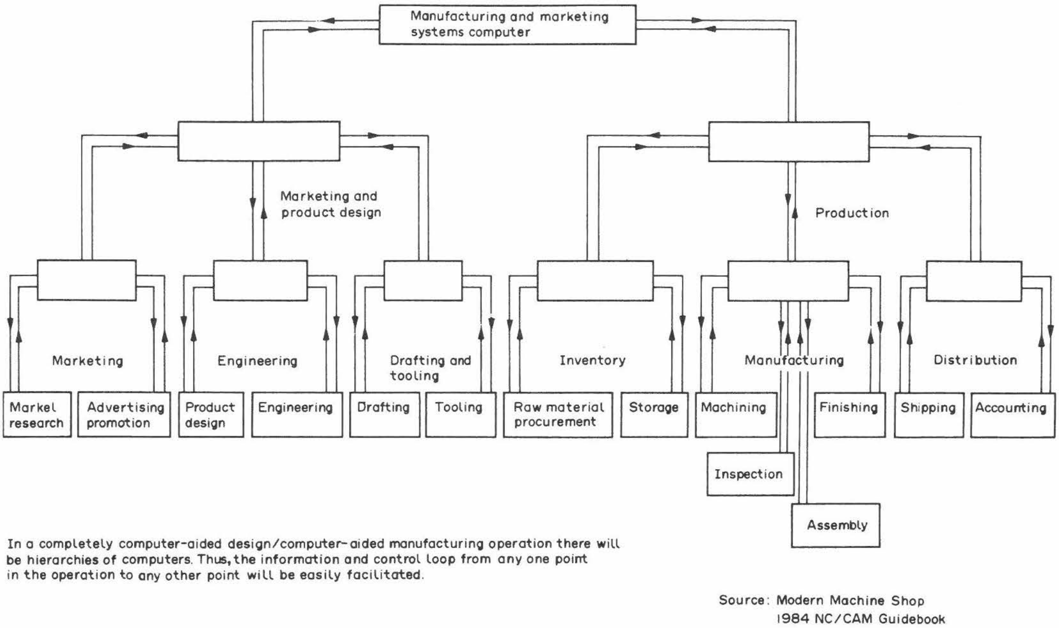


Fig. 4. Source: Modern Machine Shop 1984 NC/CAM Guidebook.

variety of different products. In the latter situation, it is essential that records of each tool's use (material being cut, speed, cutting time) be kept and stored in memory. Moreover, the information must be available at the scheduling level of a hierarchical control system—a level higher than the machine control level—so that individual machines are not shut down for tool-changing in the middle of production runs.

One strategy to minimize such difficulties in multi-product shops is to incorporate into a single machine tool all the machining operations needed for a given part (e.g. milling, drilling, boring, facing, threading and tapping). Extremely general-purpose machines called "machining centers" were introduced in the 1960s explicitly to exploit the capabilities of NC. Such machines may have as many as 90 different tools and programmable tool-changing capability. A typical comparison is shown in Table 1. But the strategy of increasing the flexibility of individual machines is limited because only one of 90 tools can be utilized at a time. Thus, the more tools and degrees of freedom the machine has in reserve, the more its maximum capability is likely to be under-utilized during any particular operation. Machining centers are therefore mainly used for

Table 1. Comparison between methods for machining an electric motor housing[8]

Measure	Conventional machines	Machining center with automatic tool changer
Setups	5	1
Tools	22	18
Operations	111	111
Machines	4	1
Time (min)	99	41

small batch production of very complex prismatic parts like the motor housing cited in Table 1.

Another major trend in multi-product shops since the 1950s is computerized integration of individual multipurpose machines into linked sets—manufacturing cells or FMS—to produce families of parts. Such families are classified by geometry and size using "group technology" or GT, as noted earlier.

To integrate a number of machines or cells, a hierarchical top-down control system is often advocated* (as illustrated in Fig. 4) because detailed instructions for the actions to be performed by each particular machine must be converted into "machine language" specialized to a specific interface. However, for efficiency, it is necessary for the production

*This is primarily due to the rapid development of microprocessors, which permits the logical partitioning of software.[8] Distributed control is also far less vulnerable to massive breakdowns of the system as a whole.

engineers to be able to program virtually all design, coordination and scheduling functions in a single higher-level language. Thus a supervisory computer should be able to translate from the high level language used by the human engineers to the detailed machine-level languages understood by each machine. It seems increasingly likely that the control functions will be distributed around the system rather than centralized in a single supervisory computer. Some implications of this will be considered later.

A third and related trend in factory automation is toward permitting instructions to be given in truly functional terms. Ideally, it should be possible for a machine or machining cell to be instructed in natural language: e.g. "make 10 copies of model #XYZ 123". The supervisory computer would then consult an on-line database to determine what parts are required and how many of each. It then consults stored files of existing programs, and, based on a scheduling algorithm, decides which machine or cell is to make the part, then calls for the needed bar-stock or other material requirements from inventory, downloads the instruction program for each desired part number to the microprocessor controlling the designated machine (or machining cell), and activates it.

The use of high-level languages imposes substantial requirements on the computer hardware, machine interfaces, and software. One limitation on the utility of non-adapative numerical control at present is that feed-rate and cutting-speed should be based on tool hardness and condition (wear). An efficient and flexible programming system should allow for a range of each parameter. Since the cumulative amount of wear depends on the initial tool hardness and its cutting history, no sensory feedback from the tool itself would be necessary if the supervisory computer had an adequate mathematical model for tool wear as a function of type of use in storage. It could also compute economically optimum feed rates and cutting speeds. However, this mode of (non-adaptive) control is inherently inflexible and intolerant of unexpected deviations.

In practice, too, the size of the required database and the large computational requirements of model-based controls seem to preclude doing these optimization calculations on-line in real-time. A more practical approach might be to use semi-empirical models, e.g. the Taylor tool-life equations. However, the optimal condition as determined by such methods may possibly violate physical conditions for safe (non-abusive) tool operation[38] since empirical tool-wear equations only reflect physical wear mechanisms (such as abrasion, adhesion, cratering, fracture, and plastic deformation) in an average sense. Again, deviations from the expected cause difficulty.

Yen and Wright[38] propose adaptive strategy. This requires a unified consideration of applicable physical constraints and economic optimization, the latter being carried out only after the former are satisfied. The applicable physical constraints are functions of dominant wear mechanisms but these can generally be expressed as localized cutting tool temperature and force limits. The adaptive control strategy, then, is to continuously monitor these variables and adaptively vary the feed-rates and cutting speeds to stay within the "safe" operating regime. This element of the overall control function can best be carried out at the individual machine level, giving rise to the notion of distributed adaptative control. The higher level supervisory computer needs only to be informed when a worn tool is actually replaced, so it can call for another one from stock or from the supplier.

The use of feedback information, generated by sensors within the machine to provide decision information on the state-of-the-machine or the state-of-the-workpiece for the control computer, requires further hardware/software interfaces. The sensor data must be "read"—usually as an analog signal—and converted into digital form. In order to be used in a decision algorithm, this signal must then be interpreted by comparing it with a stored or model-generated value or norm. For instance, if the actual metal removal rate is too low, as compared to the acceptable range of values, the MCU should inform the control computer and call for a tool-change.

The vast majority of sensors currently used in industry produce low-grade, narrow-spectrum, yes–no signals. Such signals can only convey very simple messages and trigger correspondingly simple binary decisions. But there are many inspection situations where it is desirable to make complex adjustments or modifications in the instruction program in order to meet part specifications. For instance, if the force feedback from the workhead increases beyond a certain point, the tool may be jammed. An "unjamming" routine must then be initiated. Sensory information must be adequate to enable the computer to correctly interpret the situation and issue instructions (to a robot arm, for instance) to re-orient a part if necessary. Or, suppose a milling operation called for by the initial program results in a part with incorrect physical dimensions. The computer must sense the situation and decide whether

Table 2. Five generations of automation

	Pre-manual control	First (1300) fixed mechanical stored program (clockwork)	Second (1800) variable sequence mechanical program (punched card/tape)	Third (1950) variable sequence electro-mechanical (analog/digital)	Fourth (1975) variable sequence digital (CNC) (computer control)	Fifth (1990?) adaptive intelligent control (AIC) (systems integration)
Source of instructions for machine (How is message sent?)	Human operator	Machine designer/ builder	Off-line programmer/ operator records sequences of instructions manually	On-line operator "teaches" machine manually Off-line programmer prepares instructions	Generated by computer, based on machine level stored program instructions modified by feedback	Generated by micro compter based on high-level language instructions, modified by sensory feedback; or generated by algorithm from instructions carried by workpiece
Mode of storage (How is message stored?)	NA	Built-in (e.g. as patterns of cams, gears)	Serial: patterns as coded, holes in cards/tape or as pre	Serial: as mechanical (analog) record (e.g. on wax vinyl disc) Serial: as purely electrical impulses (e.g. on magnetic tape)	In computer memory as program, with branching possibilities	In computer memory as program with interpretive/ adaptive capability and/or as coded instuctions on "escort memory" carried by workpiece
Interface with controller (How is message received?)	Mechanical linkage to power source	Mechanical: machine is self-controlled by direct mechanical links to drive shaft or power source	Mechanical: machine is controlled by mechanical linkage actuated by cards via peg-in-hole mechanism	Electro-mechanical: controlled by valves, switches, etc. that are activated by transducers — turn, controlled by playback of recording	Electronic: machine reproduces motions computed by program, based on feedback information	Electronic: telecommunications linking from other machines/ computers; or by magnetic scanner from "escort memory" carried with workpiece
Sensors providing feedback	NA	NA	NA	NA	Narrow Analog (converted to digital) (e.g. voltmeter/ strain gauge) Spectrum Digital (e.g. optical encoders)	Analog or digital, wide-spectrum complete descriptions visual, tactile, requiring computerized interpretation

the flaw can be eliminated by additional milling; if so, the instruction program must be revised appropriately; if not, the part must be sent back for rework or be discarded. If the procedure fails, the machine must then stop automatically and call for help. Then, too, the basic program should be permanently modified to ensure a problem is not repeated endlessly. Capabilities of this sort are at or beyond the present state-of-the art because they require (i) sophisticated machine or tactile systems, (ii) complex decision algorithms, and ultimately (iii) an ability for the supervisory system to "learn" from experience.

The capabilities described above must be a generalized part of the system software, not specific to any specific machine or cell, because individual machines as well as the computers may be replaced or regrouped at any time. The adjustment and modification capabilities must therefore be capable of interpreting sensory inputs that are also somewhat generalized in nature. This implies that the introduction of artificial intelligence (AI) into factory operation must be preceded or accompanied by the availability of machine vision and sophisticated wide-spectrum sensor (e.g. sensory information) processing capabilities. All of these capabilities evidently belong to the next (emerging) state of factory automation, which we have denoted "5th generation" automation in Table 2.

The assembly task deserves further discussion at this point, since it is an essential and hitherto somewhat neglected component of the integrated production system now evolving. Assembly tasks are significantly less automated at present than parts manufacturing tasks for the same volume of output, as shown in Table 3. In fact, humans continue to be needed in the vast bulk of all assembly operations, especially insertions. According to Boothroyd,[6] 77% of subassemblies and 86% of final assemblies are done manually (*ca* 1980) either on benches or progressive assembly lines; while only 6% of subassemblies and 4% of final assemblies could be classed as mechanized (*ca* 1980). Alternative assembly systems are shown schematically in Fig. 5.

Table 4. Typical assembly unit operations[26]

Operation		Frequency (%)
A	Peg-in-hole	34.5
B	Push-and-twist	12.8
C	Multiple peg hole	6.5
D	Peg and retainer insert	5.0
E	Screw and/or bolt insertion	26.8
F	Force fit	7.3
G	Remove location pin	1.0
H	Flip over	2.0
I	Provide temporary support (fixture)	1.5
J	Crimp sheet metal	0.5
K	Remove temporary support (fixture)	1.5
L	Weld/solder	0.5

The basic assembly operations have been classified by Kondoleon,[22] as shown in Table 4. By experiments with a number of typical products (a toaster-oven, a bicycle brake, and an electric jigsaw) Kondoleon was able to ascertain a frequency distribution for these operations. For this group of products, at least, simple peg-in-hole insertion outnumbered all others, followed by insertion of screws or bolts. Given a set of unit operations, such as the foregoing, the importance of external sensory feedback can be determined by experiment. For instance, peg removal, twisting, or metalcrimping do not require sensory feedback (in most cases), whereas positioning for insertion of pegs or screws is inherently sense-dependent, as will be seen later.

Some of the more important factors affecting positioning for insertion are as follows.[26]

- The amount of clearance (for free space) between parts after assembly.
- The degree to which they are misaligned when they first touch.
- The friction force between parts when they slide together.

A typical positioning and insertion problem is illustrated in Fig. 6. Holes are usually chamfered to aid in insertion. As the peg enters the hole it touches one side of the inside channel first. If the angular misalignment is too large, the peg will subsequently

Table 3. Mechanization vs. scale of production

Task category	Custom	Batch	Mass
Parts recognition and sorting	Manual	Manual	Not applicable (NA)
Parts transfer	Manual	Transitional (e.g. belt machine)	Mechanized (e.g. transfer machine)
Machine loading and unloading	Manual	Mostly manual	Mechanized (e.g. feeders)
Tool-welding (including machine operation)	Semi-mechanized (manual control)	Mostly mechanized (NC) except for supervisors	Mechanized fixed sequence
Parts inspection	Manual	Manual	Transitional
Parts mating and assembly	Manual	Mostly manual	Transitional

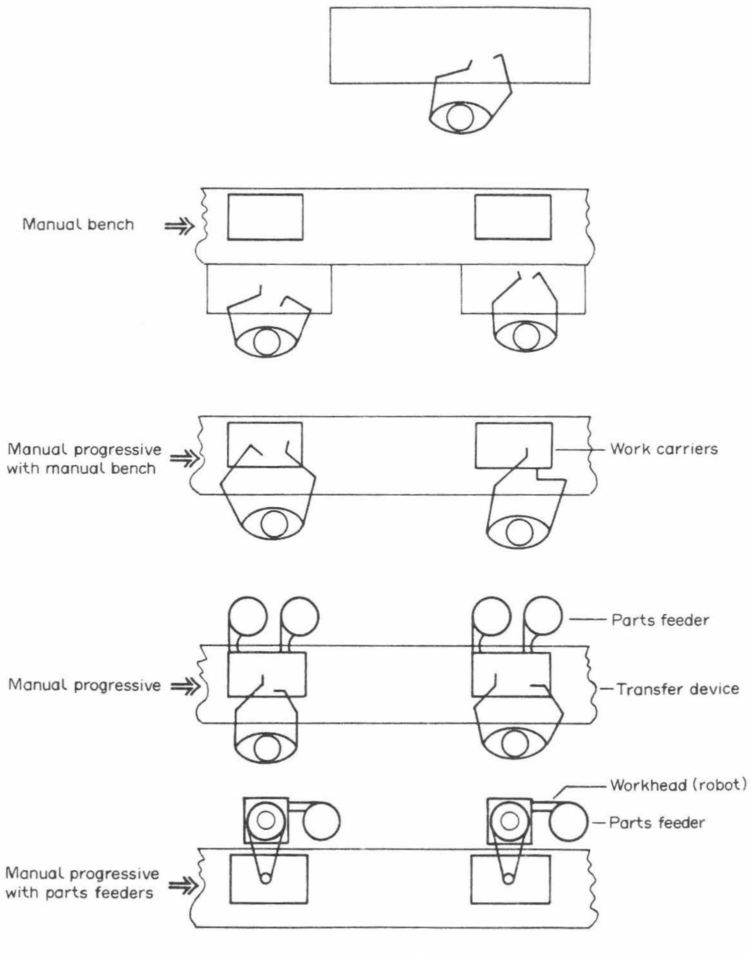
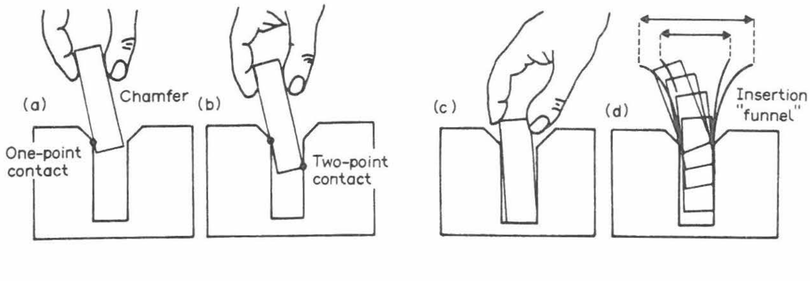


Fig. 5. Alternative assembly methods.



Source: Nevins & Whitney 78

Fig. 6. Insertion of peg in hole.[26]

touch the opposite side of the hole. Whether parts will mate successfully or not depends on the relative error between the actual and desired alignment as they touch for the first time. If the relative misalignment is small, mating can usually proceed without difficulty, but if the misalignment is larger, the insertion will jam, causing an interruption at least and possibly damage.

To repeat one key point: the control function is inseparable from sensory feedback capability. This is true whether control is distributed or centralized. In the early days of computers it seemed obvious that the digital computer, with its ability to perform complex calculations at ultra-high speeds, should be able to take over this control function quite easily.* In fact, however, this goal has proven to be far more elusive than was originally believed. The problem—at least since the mid-1970s—is not lack of raw computational power, or the cost of computation *per se*; the core of the problem lies with other elements of the control system, notably the sensors and the associated sensory-interpretation software and hardware.

3. CHARACTERIZATION OF SENSOR TECHNOLOGY FOR CIM

In continuous flow manufacturing processes (especially in the chemical industry), pressure, temperature, flow-rate, and other sensors have long been integrated into computerized control systems. But the integration of machines for processing discrete parts into such systems has been much slower, partly due to the relatively high cost or poor performance of appropriate sensors.

Some kinds of information required by machine tool controllers include:

- linear position (of workpiece or tool) on an axis
- angle between two axes
- linear motion (speed) along an axis
- angular motion (rotational speed) around an axis
- linear acceleration or force exerted along an axis
- torque or torque around an axis
- power
- temperature at a specified point.

Commercial sensors may be *passive* or *active*. Passive detectors react to a motion or condition of the tool or workpiece and generate a signal that can be interpreted. A well-defined mapping from the "state-space" of the tool/workpiece to the "signal-space" is needed to calibrate the detector. Examples include pressure transducers, thermocouples, IR detectors and photo-detectors. However, passive detectors tend to be limited in resolution by the inevitable presence of uncontrollable noise in the system. Clever instrument design and signal processing can reduce the importance of this factor, but only at significantly higher cost. Examples of passive sensors include tachometers, transformers, Hall-effect piezo-electric accelerometers, strain gauges, photo-detectors and IR detectors.

When passive detectors are still inherently too limited, an active detector system can often be devised to increase the signal-to-noise ratio. The sensor itself generates a precisely defined signal which is then propagated, absorbed, reflected, scattered, frequency-shifted or otherwise modified by the state (or changing state) of the object being sensed. Examples of active sensor systems include photo-optical encoders, inductive probes, lasers, radar, sonar (ultrasonic ranging), "structured" light proximity ranging, and electronic imaging (machine vision).

Obviously, sensors are used for a variety of different purposes, which makes direct comparison difficult and potentially misleading. A list of common types is shown in Table 5. One important purpose in discrete part manufacturing (which does not arise at all in most continuous process applications) is *identification*; vision is only one approach to identification. The entire question is dealt with later.

Sensors of the above types are appropriate when used for monitoring some aspects of machine performance and tool/workpiece orientation with respect to a single degree of freedom. They are generally inadequate for tool/workpiece monitoring in multiple degrees of freedom or for post-process tool/workpiece inspection. Surface integrity and finish (e.g. smoothness) are very difficult to measure, although the intensity of reflected laser light is increasingly used for this purpose. For more precise inspection tasks, especially where complex part shapes are involved, human vision is normally required. Machine vision offers an eventual substitute for human eyes in surface inspection.

Before embarking on a detailed discussion of machine vision *per se*, it is important to note that for one of its potential roles—parts identification (or tool identification)—it has competition from other

*In the introduction to *Cybernetics*, written in 1946, M.I.T.'s famous mathematician and control theorist, Norbert Wiener, commented, "The automatic factory, the assembly line without human agents, are only so far ahead of us as is limited by our willingness to put such a degree of effort into their engineering as was spent, for example, in the development of radar in the Second World War".[34]

Table 5. Sensor types

Data required	Type of sensor
Linear position (e.g. slide on bed)	Photo-optic grating fixed to the bed illuminated by LED and reflections sensed by photo detector
	Linear inductive resolver, i.e. consisting of a fixed coil in the bed and moving coil mounted to the slide (Inductosyn)
	Linear magnetic grating fixed on bed and scanned by magnetic pickup on slide (+50 μm)
	Laser interferometer (+12.5 μm) split beam; one fixed to slide, one fixed to bed
Angular position (e.g. lead screws)	Angular industrial resolver consisting of rotating and fixed transformer
	Photo-optic encoder consisting of rotating disk of alternating clear and opaque sections spinning between an LED and a detector*
Linear position (spindle deflection)	Inductive probe
	Deflectometer (electro-optical)
Linear position (workpiece on bone diameter hole depth)	Contact probe (caliper) with pneumatic or electric sensor
	Contact probe (caliper), with linear variable differential transducer (LVDT)
	Magnetic grating
	Laser interferometer
Linear position (proximity)	Eddy currents probe (ferrous surface objects)
	Inductive probe (ferrous objects)
	Hall effect, sensor, DC current is proportional to magnetic field strength
	Capacitance gauges (conducting/non-conducting)
	Electropneumatic proximity probe (change in orifice pressure in presence of surface)
	Ultrasonic probe
	"Structured light"
Angular velocity (e.g. spindle RPM)	DEC tachometer (DC generator)
	Optical encoder*
	Inductive probe in proximity to slotted ring
Power (e.g. spindle HP)	Watts transducer consisting of multiplication of motor voltage and current; voltage sensed by a voltage transformer (AC) or voltmeter (DC); amperage sensed by current transformer (AC) or voltage drop across a shunt in series with power; multiplication by IC chip
Angular torque (e.g. spindle torque or tool)	Power (above) divided by spindle RPM (above) #3%
	Strain gauge dynamometer is most accurate to expensive method
Linear acceleration (e.g. spindle vibration or deflection force)	Piezo-electric accelerometer (quality crystal), voltage proportional to force applied
Temperature	Thermocouple (voltage proportional to temperature)
	Thermistor (resistance inversely proportional to temperature)
Pressure (e.g. tactile force feedback from robot gripper or fingers)	Silicon (MOS) strain gauge
	Piezoresistive transducer
	Conductive elastomer strain gauge
	Resistive potentiometer
	Inductive potentiometer
	Deflectometer (electro-optical)

*Signal can be directly digitized.

emerging technologies. The simplest of these involves the use of bar codes and optical scanners (VCRs). This technology is now quite highly developed (and widely used in supermarkets and drugstores) but has its drawbacks. The major drawback is that bar codes can only be read from a particular position and orientation. They are inherently inappropriate for small parts or parts lacking flat exposed surfaces that also remain unaffected as the process continues. In practice, this eliminates many, if not most, possible applications.

A more sophisticated technology is the reusable electronic tag or "escort memory" which is a programmable semiconductor memory that can be read by microwave at a short distance (~1 m). As an example, Philips markets such a device (Programmable Remote Identification) with 8 kbyte storage capacity, including a miniature battery powered transceiver. This is currently being used by BMW, VW and Pontiac to track car parts. Volvo and SAAB use a slightly different system known as COTAG (64 kbytes, 400 khz). Magnetic inductive tags are also now available up to 2 kbytes capacity. To be sure, these devices still cost too much (~$25) for very widespread use, but as costs drop, they may find more applications, especially in assembly.

Machine vision systems are dropping rapidly in cost as well and have the great inherent advantage of not requiring direct contact with the part.

4. MACHINE VISION

Vision has been characterized by Barrow and Tenenbaum[3] as that information processing task which accepts inputs in the form of two dimensional arrays of intensity/color values and produces a concise symbolic description of the "scene" in terms of a prior menu of possible objects, conditions and relationships. In fact, the key attribute of sensor technology in 5th generation automation will be a transition from measuring operating variables in a narrow bandwidth or single dimension to detecting and analyzing multi-dimensional, multi-colored "scenes" in much the same way as do human eyes. Ultimately, scene detection/analysis is a necessary step toward building a dynamic simulation model of the physical elements in a production system to be used by the supervisory computers. Thus, machine vision technology is critical to the success of computer-integrated manufacturing.

Detailed discussions of the current applications of machine vision are available in the published literature.[5,7,11,20,32] Machine vision technology can be characterized in terms of its image input hardware (i.e. cameras and illumination) or the types of visual information it can handle (spectral, spatial, temporal) or the image processing system (i.e. feature extraction and decision making).

An image input system consists primarily of cameras (passive photo sensors) and a method of illumination. Cameras, digitized and organized into rectangular array, typically 256 × 256 (more or less) picture cells or "pixel" elements, convert a visual scene into electronic signals. Vidicons have been widely used for vision systems because of their wide availability and low price. However, they have many drawbacks; fragile, their tube lives are limited. They suffer from electronic noise, blooming, parabolic distortion, signal drift, frequency distortion, and require frequent adjustment. Solid state cameras, which consist of an array of photosensitive electronic elements, solve many of these problems. These devices use CCD (charge-coupled device), CID (charge-injection device), CPD (charge-priming device) or MOS (metal oxide silicon) technology. They are smaller, lighter, less fragile, and have a longer operating life than vidicons. A comparison of some camera types as of 1979 is given in Table 6.[35] Since then, solid state sensing devices and image processors have progressed rapidly.[4,20,31] The popular 256 bit linear array CCD (or MOS) chip is rapidly being replaced by 2K or 4K bit devices; other recent noteworthy advances include the single-chip image processor[28] and the successful introduction of sub-pixel techniques.[13]

Table 6. Typical passive-sensor parameters (1979)[35]

Parameter	Image dissector	Image orthicon	Sb_2S_3 target vidicon	PbO target vidicon	Si target vidicon	Linear array	CID area array	CCD area array
Energy-transfer limits								
Spectral response, nm								
Lower	350	320	400	350	350	500	420	400
Upper	650	580	680	700	1000	980	1100	950
Sensitivity, 1 μm/m	650	0.001	11	3	0.1	1	1	1
Transfer function								
Horizontal resolution (10% MTF)								
Line pairs/mm	80	12	72	78	65	80	43	17
TVL	2000	500	700	750	620	1024	380	188
Vertical resolution, TVL						1	480	244
Contract transfer, @ 400 TVL (MFT)	0.6	0.5	0.5	0.6	0.4			
Gamma	1	1	0.65	0.95	1	1	1	1
Typical signal characteristics								
Signal-to-noise ratio, dB	45	34	40	50	40	50	50	50
Bandwidth, pixels/sec	50 k	6 m	8.5 m	9 m	7.5 m	1 m	3.5 m	2.7 m
Geometric fidelity								
Linearity, %	0.2	2	2	2	2	0.1	0.1	0.1

The performance numbers represent a possible operating mode for each sensor. For solid-state arrays, "resolution" means the number of active detectors. Gamma expresses the conversion characteristic between surface irradiance, E, and signal current, I, in the approximation, $I = KE\gamma$. The approximation to the actual performance applies only for full-field irradiance at nominal values for the specific sensor. For display and detection, $\gamma < 1$ may be desirable.

The method of illumination reflects the type of features which are most important to the particular application. The four general methods of illumination are diffuse lighting, backlighting, direction lighting, and spatially modulated or "structured" lighting. Diffuse lighting is generally used for objects characterized by smooth regular features and is the most common method utilized. Backlighting is ideally suited for applications in which the object's silhouette is of primary interest. Narrowly angled directional lighting is ideally suited for the inspection of surfaces for flaws because very little light is scattered from a flaw-free surface, and a defect, such as a scratch, can be detected from its shadow or from scattering. Spatially modulated light involves projecting points, stripes, or grids, onto an object. General Motors first utilized this approach with the Consight-1 system[21] in order to detect an object's silhouette. However the primary application now appears to be in the detection of object curvature or to infer 3-D information through triangulation.

A machine vision system can often be characterized in terms of the information it can process; the three types of information usually distinguished are spectral, spatial and temporal. The best approach to processing depends on the purpose of the exercise.

Spectral information can be divided into frequency (i.e. color, X-rays) and intensity. Current machine vision systems cannot differentiate between wavelengths and thus cannot handle color data. Applied Intelligent System's Pixie 5000, introduced in 1984, differentiates between red, blue, and green. At present, optical filters can extract information about specific wavelengths, but this is a crude and slow process. X-Rays are useful for detecting internal cracks in castings and other formed or machined parts through identifying changes in density. Penn Video offers a system which can be used for manual or automatic inspection of an X-ray image.

The use of intensity information is usually characterized in terms of the number of "gray levels" which a system can differentiate. Many commercially available systems can only differentiate between two gray levels in one image. In these so-called binary systems, all gray levels below a specified threshold are converted to "white" and those above are converted to "black". The problem with this method is in selecting the optimum threshold. There are pseudo "gray scale systems" which process a number of images at different thresholds in order to extract gray scale information. Experimental systems, and a few recent commercial ones, can actually process multiple shades of gray, although this is not yet working well due to the lack of consistent contrast within most applications.

Spatial information refers to the position, orientation, and shape of an object. Most systems reduce three-dimensional scenes into two-dimensional images. This can provide information about part position and orientation if the third dimension (e.g. distance from the camera) is accurately known. In many applications, such as part recognition (i.e. "bin picking") and inspection (dimensional measurement), this is not the case. Three-dimensional information would be extremely valuable in these and other applications.

The three major techniques for obtaining three-dimensional information are stereo imaging, structural illumination, and "time-of-flight". Both stereo imaging and structural illumination utilize triangulation to calculate the location of an object in three-dimensional space. Stereo imaging is accomplished by correlating the two-dimensional image of two or more cameras. Spatial coordinates are derived using the small discrepancies in images from different viewpoints, as superimposed on the image plane; only one camera is needed; the other is replaced by an active light source which can scan the spatial volume of interest. While one direction vector is still obtained from the passive sensor, the second is defined by the known position and intensity of the illumination source and the laws of optics.

Stereo imaging relies on identifying identical features in the images of two or more cameras. The calculation of the coordinates in three-dimensional space is based on triangulation. There are two major problems associated with the method: First, the cameras may not necessarily be exposed to exactly the same region of a complex object since obscuration and shadowing may make it impossible to compute position coordinates for certain areas. Second, it is difficult to identify identical points or regions in several two-dimensional images at a reasonable speed. Numerous techniques, such as gray-scale mapping, edge enhancement, and vertex identification, are being explored. Gray-scale mapping may prove particularly useful in industrial applications where the images considered have sharp discontinuous surfaces emphasized either by shading or by differences in angular reflectivity. Color may also prove to be useful since the human eye makes use of color as a discriminant. Thus, the real problem in image correlation becomes one of automating the selection of a sufficient number of interest or target points and interpolating the remainder. Interest points tend to be the images of points associated with edges, faces or spatial discontinuities of the object.[10]

The triangulation computation itself is not difficult or time consuming but may be limited as to the accuracy of the image coordinates used in the calculation in two ways: by the inherent resolution of the image sensor, and by the accuracy with which a point can be uniquely identified in the two stereoscopic images. Ultimately, the latter constraint is the key element.

Structural illumination uses a controlled light source to simplify the problem and thereby reduce the computational complexity. As opposed to stereo imaging, a correlation between image points is not required; the illumination source defines one direction vector while the second is obtained from the passive sensor. Both the form of the source and the manner in which it is used are controlled to maximize the data acquired and to minimize computational complexity.

Possible light sources include light sheets, swept laser beams, laser spots, and other patterned formats. These sources would have to be scanned over the surface of objects in the workplace and such scanning can be quite slow. The use of sheets and grids speeds up the process but also decreases the ability to correctly identify appropriate illumination features after they fall onto a surface.[10] The most serious drawbacks to this technique appear to be array coding and the large quantity of data (although less than stereo imaging) which must be stored and manipulated.[16]

Direct ranging can be accomplished with collinear sources and detectors to measure the time it takes a signal to propagate from source to target and back. Distance is then calculated using the known signal transport velocity. The most familiar use of this technique is sonar, in which the echoes of acoustic pulses are recorded.

The "time of flight" approach can be implemented in two ways. The first is to measure elapsed time using an impulse source. The second utilizes a modulated continuous wave signal which is compared with a return signal in order to measure phase differences. These phase differences are interpreted as range measurements. Range sensing can be done point-by-point, line-by-line, or in some cases by an area-based scheme.

No single approach seems to offer adequate stability, speed, and resolution concurrently. Ultrasonic signals tend to have poor resolution, but modern signal processing techniques with integrated approaches should eventually result in a practical ultrasonic range sensor.[10] Ultrasonics also suffers from two other limitations; first, ultrasonic signals are severely attenuated in air as opposed to fluid media (i.e. sonar); second, the propagation of an ultrasonic signal is a physical molecule-to-molecule or atom-to-atom process. Random thermal motion of atmosphere species assures a significant beam spread resulting in a loss of resolution.

Scanned laser systems have been the basis for most direct point range measurement approaches. Current laboratory systems can take 5–10 min to synthesize an area range image and can involve much adjustment and calibration.[10] The development of higher-powered semiconductor lasers could improve data acquisition rates and signal-to-noise ratios. Semiconductor lasers are emphasized because of their size and inherent ruggedness. Improved means for nonmechanical laser beam detection must also be developed. Rotating/oscillating mirrors and/or prisms can perform the function, but they lack the ruggedness required for field or factory use.[16]

Temporal information refers to motion. This can be derived from successive static "snapshots" of a moving scene using the same basic techniques as stereo-imaging, except that only a single camera is needed and the discrepancies between successive images determine the relative motion of objects in the field of view. The problems are similar to those of stereo imaging.

5. IMAGE SEGMENTATION

In the simplest terms, image processing involves interpreting or making "sense" of images, which are two-dimensional arrays of brightness values. While all current approaches can be reduced at some level of abstraction to the stages of "preprocessing", segmentation, recognition and representation, these terms are themselves ambiguous. Preprocessing using dedicated hardware is often done in order to reformat the information from an array of pixels in a more useful manner. Segmentation is the process of breaking the visual scene into its constituent objects or features. Recognition is the process of matching at least one object in a scene with an archetype of that object stored in memory. Finally, description is a symbolic representation of the structure and geometrical relationships of all identified objects in the scene.

The motivation for preprocessing is that certain data reduction, enhancement and smoothing operations have been recognized as being useful for large classes of applications. This suggests the use of standardized firmware solutions rather than software. Time and position-invariant operations can be used to "filter" an image in order to detect a given pattern in it, to "restore" a pattern that has been degraded by approximation or other such operations, to "smooth", or "enhance" a pattern, or to suppress noise. These operations can not only be done by mathematical operations using conventional digital computers, but also by special-purpose electro-optical devices.[9] The cross-correlation between two pattern functions provides a simple method of determining whether or not they are identical except for translation and multiplication by a constant (this will be discussed in more detail later). Pattern restoration can be performed through the use of Fourier transforms[14] while a pattern can be smoothed through replacing its value at each point by the average of the values of neighboring points.

Segmentation is the process of breaking a visual scene into its constituent objects of features. Segmentation can be accomplished by searching for gradients and discontinuities (i.e. edges) in the brightness patterns. Gradients in brightness can be sharpened by averaging over small regions or neighborhoods of pixels, making dark areas darker and light areas lighter. The simplest technique is to reduce the image (if possible) to a silhouette: this is known as a binary image, or BI. The next step in sophistication is to generate a simple "line drawing", based mainly on well-defined edges. This techniques is referred to as "edge detection" or ED. These are the two chief methods used by currently available machine vision systems.

Methods for refining edge detection include the application of gradient or Laplacian operations, convolution with masks representing "ideal" step edges, and neighborhood processing or cellular logic. The latter is the most popular method. In neighborhood processing, the gray scale value of each pixel is transformed to a new value based on their present value and the present values of their eight nearest neighbors. Commercial systems which utilize this technique are available from many firms. These companies use circuitry (firmware) rather than software for neighborhood processing. This has had profound effects on the speed/cost ratio of these systems, and the trend should continue in the future as LSI and VLSI technologies are more fully exploited. Figure 7 shows aggregated cost estimates made by one of the authors (J. Funk) based on interviews with industry experts. Presently the cost-effectiveness of these systems is improving very rapidly due to the application of already existing LSI and VLSI technology. The announcement of a single chip image processor by Intel (1987) confirms this trend. Subsequent improvements in cost effective-

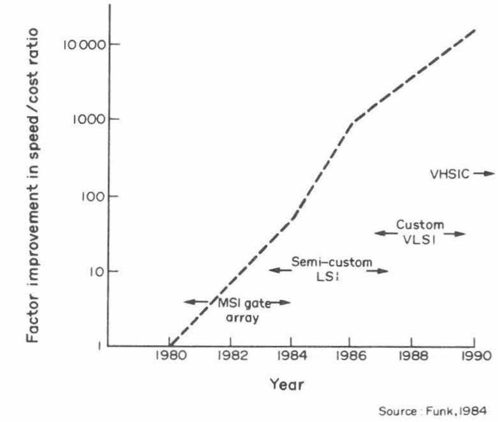


Fig. 7. Estimated improvements in speed/cost ratio for neighborhood processing.[15]

ness may be slower because they will be dependent on advances in basic microelectronic technology such as VHSIC.[36]

6. IMAGE DESCRIPTION AND RECOGNITION

Description is the process of characterizing the (segmented) scene in terms of topological geometrical or spectral descriptors. Such descriptors include the number of holes, or "blobs", the dimensions of bounding rectangles or circles, and so forth. Feature analysis can be "global" (i.e. spatial relations between features are not considered) or "local" (they are considered). Global feature analysis (GFA) is thus less sophisticated than local feature analysis (LFA). It can be used only for the simplest tasks, such as counting or locating non-overlapping discrete parts.

Customized circuitry (firmware) can be applied to the description step and still retain a large amount of flexibility in the system. The key point is whether it is necessary to retain information from previous steps (e.g. segmented image or pixel array) in order to perform the subsequent recognition step. Many pattern recognition techniques (e.g. binary and gray level statistics) must refer to information stored in the memory several times before reaching a decision. These applications can only utilize a minimal amount of firmware. Other pattern recognition techniques (e.g. edge detection) extract features pertaining to pixel intensity changes in order to verify the quality of the part. These applications (e.g. flaw inspection) can use specialized circuitry for the edge detection and description stage.

Recognition is the process of correlating each segmented object in a scene with a known symbol or label. This is the one step in which firmware cannot be applied if any flexibility is desired. Microprocessors are utilized for the recognition steps. The three principal approaches to recognition are (1) template matching, (2) decision theoretic, and (3) structural pattern approaches.

Template matching compares the input pattern pixel-by-pixel to one or more model templates stored in memory. (It does not make use of feature analysis.) A decision is made based on preselected matching criteria or similarity measures (i.e. correlation). Part recognition applications would classify the input pattern according to its closest match (correct part and orientation) while inspection applications would make a decision as to the quality of the part based on how good the match was. This can be done through image subtraction or using cross-correlation techniques with either binary or gray scale data.

The cross-correlation between the sensed and remembered pattern functions provides a simple method of determining whether or not they are identical except for translation and multiplication by a constant. This will produce a correlation coefficient at each point in the image which will take on a value between zero and one. The closer the value is to one the greater the similarity between the patterns.

The major problem with template matching lies in its pixel-by-pixel approach. Such processes can yield imprecise matches, making it difficult to decide when a match has been detected; they are sensitive to distortions in both gray scale and geometry, and are computationally expensive.[30] Illumination and scanning conditions can use inconsistent gray scale information. Geometrical distortions are primarily the result of inconsistent scenes and part appearance. Computational complexity all but precludes template matching in part recognition applications due to the generally large number of possible orientations, locations, and stable states a part can exhibit. However, it is more useful in inspection applications due to the generally better engineering of scenes.

The speed of template matching can be significantly improved through the use of special array-processing circuitry. Pixel-by-pixel comparisons can be performed using shift registers and comparators. Even higher speeds can be obtained by optical pattern recognition techniques utilizing the wave-front reflected from the scene as a whole.[9]

This device works by optically obtaining the correlation function of the reference pattern and the input scene. The scene is illuminated with coherent laser light forming a diffraction pattern. The pattern is then focused by a lens, producing the (spatial Fourier transform) on a plane. An optical filter, which is the Fourier transform of the reference pattern, is located at the focal plane. The superposition of the two patterns yields the product of the Fourier transforms. This resultant pattern is then refocused onto the output plane by a second lens, producing the inverse Fourier transform of the product. This is the correlation function. If the reference pattern is present in the input scene, the output contains a peak of light whose location indicates the location of the reference pattern in the scene.

The advantage of the optical pattern recognition (OPR) system is that the input laser light beam operates in parallel on all pixels within the input plane. The processing time is the time it takes light to travel from input to output. For a $10^3 \times 10^3$ pixel image with 8 bits intensity level (8×10^6 bits of data) the time would be trivial. The three major shortcomings of optical processors so far have been their accuracy, rigid stability requirements, and inflexibility.

So called "decision theoretic" approaches compare scenes in terms of their features, rather than pixel-by-pixel. The decision is based on a similarity rule between the extracted features and the stored model features. Features based upon geometrical or intensity information are extracted and represented in terms of a statistical distance or discriminant function. The major techniques available are binary or gray-level scale statistics or "blob analysis". Commercial vision systems which utilize decision theoretic techniques usually "learn". Learning refers to the use of training patterns to obtain the coefficients of the decisicn function via the utilization of training algorithms. A defect-free object is shown to the vision systems which calculates a feature vector. This technique is very easy to use.

A major advantage of the decision theoretic approach is that learning techniques have reached a higher level of development than structural pattern recognition techniques. Their principal disadvantage is that they ignore useful structural information in the pattern, such as the geometrical relationships between objects or components of objects. In other words, they utilize global feature analysis (GFA) but not (yet) local feature analysis (LFA). Structural information is generally important when the problem requires both classification and description as opposed to just classification of a scene. Because part recognition applications in manufacturing require information about a parts location and orientation, it appears that at least some structural relationship information (based on LFA) may be required. On the other hand, many inspection applications (e.g. component verification and flaw detection) only required decisions about the presence or absence of specified features of the scene or parts, and hence can be accomplished with GFA.

So-called structural methods of computer pattern recognition attempt to describe fundamental relationships among pattern primitives by extensive use of stored mathematical models embodying prior knowledge that can be used to "decode" the scene. The approach is based on a general theory of vision developed by Marr and Nishihara[24] and Marr.[23] "Syntactic" pattern recognition is the most widely favored method. Patterns are usually represented as a string, a tree, or a graph of pattern primitives and their relations. Describing complex patterns in terms of a hierarchical composition of similar subpatterns is especially attractive.

The language that provides the structural description of patterns in terms of a set of pattern primitives and their composition can be called a "pattern description language". The rules governing the composition of primitives into patterns are usually specified by the grammar of the pattern description language. Recognition is performed with a syntax analysis, or parsing, of the "sentence" describing the given pattern to determine whether or not it is syntactically (i.e. grammatically) correct with respect to the specified grammar. This approach probably has the greatest generality, and can be expected to more frequently used in the future.

To summarize the last three sections: sensor technology—especially "machine vision"—is in the midst of a period of radical, as yet largely unrecognized, change. From a rather pedestrian role on the periphery of discrete part manufacturing, this technology is about to take center-stage. Machine vision is the missing key to flexible parts recognition without direct contact. It is also critical for robotic manipulation with multiple degrees of freedom whenever rigid parts pre-orientation is unfeasible or unreasonably expensive. Vision is also clearly essential for automated parts inspection. Finally, vision and/or taction are the *sine qua non* of automated parts assembly, as is clearly indicated by the work of Kondoleon,[22] Nevins and Whitney,[26,27] Harman,[18] Sanderson and Perry,[32] Miller[25] and Funk.[15]

7. ECONOMIC IMPLICATIONS OF 5TH GENERATION AUTOMATION

First, and most important, although flexible CIM automation is arriving gradually, its cumulative impact will be revolutionary, not just evolutionary. The reason is, quite simply, that over the next three to four decades it will eliminate virtually all routine and repetitive jobs where the human worker now acts as a machine controller or "baby-sitter". By the second or third decade of the next century, most factories will be virtually unmanned for two (or three) shifts. The only jobs remaining on the factory "floor" will be repair and non-routine maintenance, security, equipment installation and special fabrication (e.g. model prototypes, fixtures, tools and dies, and patterns). Engineering, design, planning and computer programming will also continue to be human functions, although largely carried out with computer assistance in remote locations electronically linked to the production facility.

In fact, it is noteworthy that most of the remaining non-routine jobs can be accomplished, in principle, on an occasional—even irregular—basis. One important implication of this fact is that *unmanned factories need not be located on the surface of the earth*. In fact, many future factories may well be located in earth orbit, or on the lunar surface. The purpose would be to take advantage of special conditions that might be unfavorable for humans but favorable for production purposes, e.g. to exploit available raw materails, energy, low gravity, low temperature, vacuum, or ease of waste disposal. Such a plant could in many cases be maintained by robots and/or remote-controlled teleoperators with human visitations at relatively long intervals.

The unmanned "5th generation" factory of 2015 (give or take a few years) will operate around the clock, 24 hours a day, 365 days a year, to maximize the utilization of costly computers and capital equipment. Because of this much more intensive use, mechanical equipment will not last as long as it does in traditional plants. A typical plant might have a useful life of 3–5 years at most before the physical machines and equipment would have to be replaced.

However, even 3–5 years useful life may be too long for most products; one of the major driving forces toward flexible CIM is the need to reduce barriers to innovation. In particular, competitive survival in world markets is increasingly essential to bring new product ideas from blackboard to marketplace in a much shorter time than has been possible in the recent past. Indeed, the long-sought direct computer link between CAD and CAM is primarily needed for this purpose. A truly flexible CIM system should be able to shorten the lag from several years in some cases at present, to a few months or less.

The sophisticated future CAD system we can now begin to envision will incorporate a number of so-called "expert systems" to help the product designer in three important ways. First, it will carry out many "housekeeping" functions for the designer semi-automatically; for example, it will calculate minimum dimensions of stress-bearing members in complex systems such as gear trains based on *functional* specifications (horsepower, speed, etc.). Second, it will interact with the designer to permit testing and heuristic optimization of many design variants with respect to performance, reliability, and manufacturability.* Third, having (tentatively) selected a satisfactory design (including materials), it will assist the manufacturing engineer to develop an efficient process sequence. Here again, there will be testing and heuristic optimization of many variants, giving joint consideration to minimizing wastage of materials, process energy, the number of different machines that must be used, and the tolerances required at each step.

One or more prototypes of the final product then must be made and physically tested. Computer-assisted methods of rapidly simulating wear and deterioration under conditions of actual use will have been highly developed, therefore drastically shortening the time required for testing to a few weeks in most cases.

The last step will be creating the machine control software for the new product. Assuming the factory is intended as a multi-product plant, this means modifying the instructions for the supervisory computer which processes orders and controls materials flow, inventory, and machine (or cell) scheduling. It also means providing appropriate library programs to be downloaded to each of the lower levels in the control hierarchy, including individual machines.†

The new internal software components needed to manufacture a new product will itself be largely generated by a manufacturing engineer assisted by a computer—not the operating system in the factory, but an off-line computer system that contains in its memory a model of the entire factory, including controls and software. This off-line computer will have to be addressable in a high-level language closely resembling ordinary English. The computer-aided software development process will be quite similar to CAD; the computer will carry out most housekeeping functions, including the final process of generating and testing of code (and possibly even reducing it to hardware terms, i.e. a "chip" design).

Obviously, the computer software required to facilitate product design, product testing, manufacturing systems engineering and operating systems software generation will have to be very sophisticated by present standards. Indeed, it is now fairly clear that software availability is the real technological bottleneck in creating a 5th generation automated factory. (This has already been seen in regard to vision-processing.)

An unsuspected consequence of the nature of the future manufacturing system is that each unit of output—say, a car—will require a much smaller physical quantity of "metal-bending" equipment (because each machine will operate around the clock) than is now the case. On the other hand, it will require a much larger information processing capability, including a major investment in software.

Factories built before 1980 incorporate relatively little in the way of electronic sensors and information processing equipment (including computers)—probably much less than 5% of total investment in dollar value. By the end of this decade the percentage could very well double or even treble its present level. By the year 2015 it is not implausible that electronics and information processing capital equipment will constitute well over 50% of the total investment in a new plant. (Surprising as this may seem at first, it may be noted that a similar trend is observable at a much more advanced state in the case of military hardware.)

It is also important to note that, while the electronic and information processing component of the capital goods industry will rise sharply, the overall cost of capital equipment per unit of output may not rise significantly at all due to (1) sharp increases in the productivity (i.e. utilization) of machinery and equipment and (2) expected productivity gains in manufacturing which should be reflected in the declining unit cost of capital hardgoods. (All of this is good news for the electronics and computer industries, but more bad news for the primary metal producers and machinery manufacturers.)

A final point needs to be made as emphatically as possible. It is probably not possible for a U.S.-based firm to justify a major investment in retrofitting

*This involves such factors as minimizing the total number of parts, especially connectors; minimizing the number of parts that have to be machined to very high tolerances; minimizing or reducing the use of wiring harnesses or "soft" parts, etc.

†The problem is only slightly different in the case of a non-hierarchical distributed control system.

advanced manufacturing technologies on the basis of expected labor savings.* Unskilled labor is still much too cheap. But for a U.S. firm to seek such a labor-savings justification before investing almost certainly would be a mistake of the first magnitude. In the long run, no manufacturer will be able to survive in the U.S.A. without adopting flexible CIM; those who come late to the party are likely to be left out in the cold. This is the real and important justification for federally supported research in sensor-based robotics and related technologies.

REFERENCES

1. Abernathy, W.J.: *The Productivity Dilemma*. Baltimore, Johns Hopkins University Press, 1978.
2. Aerospace Industries Association of America, Inc, Washington, DC, 1985.
3. Barrow, H.G., Tenenbaum, J.M.: Configurational vision. *Proc. IEEE* **69**: 573, May 1981.
4. Bicking, R.E.: Evolution of silicon-based sensing. *Sensors* **2**: April 1985.
5. Bolles, R.C.: An overview of the applications of image understanding to industrial automation. In *Technical Symposium on Techniques and Applications of Image Understanding*, SPIE, Washington, DC, April 1981 (SRI Technical Note 242).
6. Boothroyd, G.: *Design For Assembly—A Designer Handbook*. Department of Mechanical Engineering, University of Massachusetts, Amherst MA, 1980.
7. Brady, M.: Seeing machines: current industrial applications. *Mech. Engng* 51–59, 1981.
8. Carter, C.F.: Towards flexible automation. *Mfg Engng* Aug. 1982.
9. Casasent, D.: Pattern recognition: a review. *IEEE Spectrum* 1981.
10. Corby, N.R.: Machine vision for robotics. *IEEE Trans. Ind. Electr.* **30**: 282–290, Aug. 1984.
11. Crowley, J.L.: *Machine Vision: Three Generations of Commercial Systems* (CMU-RI-TR-84–1). Robotics Institute, Carnegie-Mellon University, Pittsburgh, 1984.
12. Factory of the future. *The Economist* 30 May 1987.
13. Englander, D.: Expanding machine vision ganging with sub-pixel techniques. *Sensors* **4**: June 1987.
14. Fu, King-Sun: Pattern recognition for automatic inspection. *Computer* 34–40, Dec. 1982.
15. Funk, J.L.: The economics of machine vision and flexible assembly, Ph.D. Thesis Carnegie-Mellon University, Pittsburgh, PA, May 1984.
16. Anonymous: *A Review of Three-Dimensional Vision For Robotics*. GEO Center, May 1982.
17. Gonzalez, R.C., Safabakhsh, R.: Computer vision technique for industrial applications and robot control. *Computer* **15**: Dec. 1982.
18. Harman, L.D.: *Automated Tactile Sensing*. Robot IV Conference, Detroit MI, 2–4 Mar. 1983.
19. Hartley, J.: *Fighting The Recession In Manufacture*. IFS/McGraw-Hill, 1986.
20. Hofmeister, R.: Solid state imagers. *Sensors* **3**: June 1986.
21. Holland, S.W., Rossol, L., Ward, M.R.: Consiglet-1: a vision-controlled robot system for transferring parts from belt conveyors. In *Computer Vision and Sensor-Based Robots*, Dodd, G.G., Rossol, L. (Eds). New York, Plenum Press, 1979.
22. Kondoleon, A.S.: Application of technology-economic model of assembly techniques to programmable assembly machine configuration. Masters Thesis, M.I.T. Cambridge, MA, 1976.
23. Marr, D.C.: *Vision*. San Francisco, W.H. Freeman, 1982.
24. Marr, D.C., Nishihara, H.: Visual information processing artificial intelligence and the sensorium of sight. *Technol. Rev*. 42, Oct. 1978.
25. Miller, S.M.: Recent developments in robotics and flexible manufacturing systems. In *An Exploratory Assessment Of Second Generation Robotics And Sensor Based Systems*, Ayres, R.U., Miller, S.M. (Eds). Carnegie-Mellon University, Pittsburgh PA, Mar. 1984. Report to NSF.
26. Nevins, J.L., Whitney, D.E.: Computer-controlled assembly. *Scient. Amer*. **238**: 62–74, Feb. 1978.
27. Nevins, J.L., Whitney, D.E.: Robot assembly research and its future applications. In *Computer Vision And Sensor-Based Robots*, Dodd, G.G., Rossol, L. (Eds). New York, Plenum Press, 1982.
28. Pai, A.L., Lin, S.H., Ryan, D.P.: A single chip image processor. *Sensors* **4**: June 1987.
29. Rosen, C.A.: Machine vision and robotics: industrial requirements. In *Computer Vision and Sensor-Based Robots*, Dodd, G.G., Rossol, L. (Eds). New York, Plenum Press, 1979.
30. Rosenfield, *Computer Vision*. Computer Science Center, University of Maryland, College Park, MD, 1982.
31. Sanderson, R.J.: Electronic vision systems. *Electron. Imaging* **2**: April 1983.
32. Sanderson, A.G., Perry, G.: Sensor-based robotic assembly systems: research and applications in electronic manufacturing. *Proc. IEEE* **71**: July 1983.
33. U.S. Congress Joint Economic Committee, Washington, DC, 1984.
34. Wiener, N.: *Cybernetics: Control and Communications In The Animal And the Machine*. New York, Wiley, 1948.
35. Williams, W., Cunningham, X., Eskenazi, X.: Robotics vision. *Astronaut. Aeronaut*. May 1979.
36. Wilson, S.: Personal communication. 1984.
37. Wright, P.K., Englert, P.J.: Sensor based robotic manipulation and computer vision in flexible manufacturing cells. In *Winter Annual Meeting (Production Engineering Divison)*; 53–69. American Society of Mechanical Engineering, 1984.
38. Yen, D.W., Wright, K.: Adaptive control in machining—a new approach based on the physical constraints of tool wear mechanisms. *J. Engng Ind*. (submitted).

*For detailed analysis of the economics of sensor-based robotics for assembly, see Funk.[15] This statement in no way suggests that labor savings are not possible in many particular cases.