12

Practical machine intelligence

E. D. Sacerdoti
Machine Intelligence Corporation
Palo Alto, California, USA

1. OVERVIEW

Machine intelligence, more commonly known by the misnomer artifical intelligence, is now about twenty-five years old as a scientific field. In contrast with early predictions, its practical applicability has been frustratingly slow to develop. It appears, however, that we are now (finally!) on the verge of practicality in a number of specialities within machine intelligence more or less simultaneously. This can be expected to result in the short term in a qualitative shift in the nature of the field itself, and to result in the longer term in a shift in the way certain industries go about their business.

Machine Intelligence Corporation (MIC) was founded in 1978 as a vehicle for bringing the more practical aspects of the field into widespread use. Its charter is to develop, produce and market products based on many aspects of Machine Intelligence technology.

This paper will discuss three specific areas of work in machine intelligence that MIC feels are ripe for commercial application: machine vision, natural-language access to computers, and expert systems. It will close with some observations on what makes these areas appropriate for application at this time, and on the difference between a technical solution to a problem and a product.

2. MACHINE VISION

Achieving computer-based vision has a high payoff, though it is technically difficult. In the short term, such tedious tasks now performed by humans as inspection of manufactured objects, analysis of aerial photographs, and visual screening of blood or tissue samples can be performed with greater speed, accuracy, and repeatability by machine. In the longer term, more ambitious tasks such as visual monitoring for auto safety, sophisticated aids for the blind, and space surveillance and other military applications are possible.

As with most areas of machine intelligence, performance in interpreting a perceived scene improves with how much the system knows in advance about the scene and the objects in it. If the perceived environment is sufficiently well controlled, it is now feasible to extract most of the desired features from the image in real time with relatively inexpensive hardware. Thus, a *vision module*, a complete system for interpreting images of a restricted type, can be developed for many kinds of specific applications. These include automatic reading of printed characters and insignia, identification and classification of microscopic particles for industrial quality control and air pollution monitoring, identification and counting of blood cells, recognition of military target signatures from radar and optical data, and recognition of diagnostic features from X-ray and other medical imagery. The factory constitutes an environment in which it is typically simple to control lighting to enable qualitative and quantitative inspection and measurement of workpieces and subassemblies, and visually-guided material handling and assembly operations. MIC has thus determined to pursue factory automation as its first commercial market for machine vision.

Applications for machine vision in manufacturing can be broken into two broad classes: inspection and manipulation.

A. Machine Vision for Visual Inspection

There are two kinds of inspection task. One consists of highly accurate quantitative measurements of critical dimensions of workpieces or placement of parts in assembly. Many applications of this type, extending from microscopic measurements of cells or fine particles to measurement of exterior and interior dimensions of key features of very large workpieces, are within the state of the art; ongoing research and development will further extend the range of application at a rapid pace. Integration of vision with an x-y stage allows for high-precision inspection by taking multiple views of an object.

The second class of inspection applications involves qualitative and semi-quantitative visual inspection as is done by humans without the aid of gauges or other measuring instruments. Examples of these tasks include sorting by shape or insignia, inspection for label registration, cosmetic properties and finish, burrs, cracks, and other defects, inspection for completeness of assembly, determining approximate size, location, and count of key features such as holes, shafts, mounting flanges, determination of handedness, and monitoring for safety. In these classes of application, the processed visual information is usually used to control mechanical separation and sorting, and/or maintain statistical records.

B. Machine Vision for Sensor-Controlled Manipulation

A machine vision module can be used to guide a robot in manipulating workpieces for material handling and assembly. Applications include sorting randomly oriented workpieces on belts and other conveyors; loading machine tools, furnaces, and test equipment; packing and palletizing workpieces and finished goods;

picking workpieces from semi-organized piles in bins and containers; guiding tools in manufacturing processes such as fastening, welding, sealing, gluing, deburring, cutting, polishing, sandblasting, painting, and so on; and assembly operations such as fitting parts together and fastening them.

C. 'Simple' Machine Vision

MIC's initial product, the VS-100, processes images that are extremely simple compared to natural scenes. The current state of the art permits only such 'simple' images to be processed in a small number of seconds by a microprocessor-based system. Yet even these so-called simple images embody a great deal of data; and elaborate data reduction techniques, drawn directly from experience in advanced vision research, must be employed to enable small systems to process the images at all.

The VS-100 was developed to provide a broad range of image processing capabilities, rather than to perform optimally on a specific limited task. It recognizes and inspects images of complex objects against a constrasting background in ½ second to several seconds, depending on the complexity of the image. The objects can be anywhere in the field of view, in any orientation.

The VS-100 accepts grey-scale data from a range of camera types, which it thresholds to produce a binary (black-and-white) image. This is a significant data reduction in its own right; thresholding an image with 64 levels of grey to a binary one reduces the data per pixel from six bits to one. Yet there remain, for the typical resolution of 256 $\times$ 256 pixels, over 65 000 bits of data in a single image, allowing only 15 microseconds per pixel to process an image in one second, or 1 microsecond per pixel for a more widely acceptable rate of fifteen images per second. Since a microcomputer such as the DEC LSI-11/02 incorporated in the VS-100 requires 6–10 microseconds *per instruction*, it is obvious that special-purpose hardware and carefully crafted software must be employed.

The images input to the VS-100 are run-length encoded for data compression and subsequent processing speed. Efficient algorithms operating in the LSI-11 perform a complete connectivity analysis of the encoded images, building data structures that represent essential features of each contiguous region. Up to 13 distinguishing features such as area, perimeter, centre of gravity, number of holes, and minimum and maximum radius can be extracted for each region. Additional features provide information on each region's position and orientation.

Object recognition of each region in the scene can be performed using a nearest neighbour classifier operating on a user-selectable subset of the features. Precise numerical measurements are computed to indicate the degree of confidence in the system's recognition of the object. If the degree of match of the selected features is below a user-settable threshold, the object is rejected as unknown or defective.

While the VS-100 is perhaps the most advanced image processing system that is commercially available today, it must evolve and adapt to provide a range of products capable of opening the large potential market for industrial vision.

Of primary importance, it must operate faster. It must be specialized for particular functions, since the full power of the VS-100 is rarely needed for any specific industrial operation. It must be ruggedized to operate in hostile environments. It must be coupled with structured light, patterned lighting projected on the objects being viewed, so that 3-dimensional information can be derived from the 2-dimensional images of the light patterns. Finally, it must be extended to handle grey-scale and colour images.

MIC is extending its product line beyond integral vision systems, providing 'production processors' of which image processing is a component. A first example of this is 'Univision', a result of a joint effort with Unimation, Inc. A Univision system consists of a Puma manipulator coupled with a VS-100. Unimation's VAL language for robot control has been extended to include image-processing commands, enabling robot activities such as material handling and assembly to occur with visual feedback and control. Thus, an early workbench environment for machine intelligence research, the 'hand-eye system', is now a commercially available product.

3. NATURAL-LANGUAGE ACCESS TO COMPUTER SYSTEMS

Fluent communication with computer systems is a major focus of work in machine intelligence. For over two decades, researchers have attempted to solve the problem of communicating with machines in natural languages, such as English. Within the last few years, systems have been developed in several laboratories for real-time analysis by computer of typed input regarding a very limited subject area. The most extensive of these, LADDER [1], developed at SRI International, employs a vocabulary of over 2200 words and a grammar of over 130 rules to answer questions regarding a navy management database. Such systems run today on expensive, large-scale computers.

Under partial support from the National Science Foundation, MIC has developed the core of a microcomputer-based language processing system, which can provide most of the capabilities of the large systems in a much more cost-effective fashion. Completion of a practical system based on this development should permit computer-naive individuals, who must currently be trained to interact with computers using highly stilted, artificial languages, to use much more natural languages.

Constraints of processing speed (for efficient search through a grammar) and available random-access storage (for large grammars and vocabularies) both pose limitations on what is feasible for a language-understanding system operating in a microcomputer environment. The availability of the latest generation of relatively fast processors capable of supporting large address spaces, together with the continuing decline in the cost of random-access memory, are expected to make such systems technically practical in the near term.

Of course, a technically practical system is not the same as a product. The technology must be adapted to a particular application in which

- it is not cost-effective or not practical to train people to use an artificial language to interact with a computer,
- the range of interactions with the computer is rather wide
- introduction of a language-understanding capability does not require changing other aspects of the overall system.

4. EXPERT SYSTEMS

In every professional field there are large bodies of information acquired through study and experience by practitioners. In many fields, individuals can be identified whose performance consistently approaches the best. The goal of expert systems technology is to embody the experts' knowledge in some field within a computer. Then, the computer can act as an expert consultant for non-expert professionals or laymen. Existing systems, such as MYCIN [2], for diagnosing blood infections, or PROSPECTOR [3], for evaluating field sites for minable mineral deposits, can perform at a level exceeding that of the average practitioner in the field. These systems typically run on large, time-shared computers.

There are two components to an expert system: the expert knowledge itself, and a 'core' system for manipulating that knowledge and interacting with the user. General methodologies have been developed for encoding expert knowledge; the encoding is typically done by a computer scientist in close collaboration with an expert or experts from the field of specialization. A core system, usable for systems in a range of subject areas, has been implemented on a microcomputer in the PASCAL programming language.

There are two distinct types of users for expert systems: professionals who require access to specialized expertise (e.g., a geologist using a model of uranium deposition in sandstone), and laymen in need of summary-level expertise in a general area (e.g., a dog owner whose pet has certain symptoms and wants to know if he should call the vet). The professional may access this expertise either though a timesharing service to a large, central computer, or through a personal computer. The layman will, in the next few years, likely employ a personal computer for this activity.

The market for professional-quality expert systems is highly dependent on the particular subject, and is very intensive in its use of technical experts, both specialists in the field of application, and 'knowledge engineers' with expertise in that aspect of machine intelligence. Thus, MIC expects this business to grow very slowly in the short term.

The market for layman-quality expert systems could potentially extend to a significant fraction of the home computer market over the next five years.

5. OBSERVATIONS AND CONCLUSIONS

This section might be subtitled 'the difference between a rock-solid software technology and a product'. It is an attempt to generalize from MIC's experience in the market for machine vision systems to what might be expected for other applications of machine intelligence.

PROBLEMS OF ROBOTICS

There are a number of barriers standing between development of a successful laboratory prototype and producing a marketable product. Unfortunately for those of us whose background is primarily technical, these barriers are generally not of a scientific nature, nor even of an engineering nature, but tend to be sociological and economic.

The primary barrier is the lack of a machine-intelligence infrastructure in the marketplace. Products based on machine intelligence are going to have the common characteristic that they perform operations in a more sophisticated fashion than those typically performed by the computer systems they replace or augment. Selling such a product requires being able to help a wide range of potential customers to *understand* what the product does; it requires those customers to be able to *apply* the product to their particular problems without a great deal of assistance; and it requires those customers to be able to *operate* the product, for the most part with existing personnel without excessive training. If any of these conditions cannot be met by a machine intelligence system, it probably is not the basis for a saleable product. For example, vision systems that require someone with expertise at vision programming to set up for a new run, or database query systems that require someone with at least a first course in computational linguistics to adapt them to a new data base, are not going to be widely accepted products. A core of an expert system that requires domain experts to work with computer experts to develop workable systems defines a potential consulting business but not a software or hardware business. A second barrier derives from the nature of the problems chosen by practical researchers in machine intelligence. We tend to choose problems to work on in which the range of acceptable inputs is small enough so that the system can cover all of it, and yet for which there is enough latitude so that precompiling responses to all inputs is not feasible. However, a potential product must compete against alternative approaches which are less general, but are more powerful within the accepted range of inputs. If the solving of any specific problem is of such general value that it seems to define a market for a machine-intelligence-based system, it has probably been faced by other, less general approaches with at least some success. A machine intelligence approach may have been shown to be feasible in the laboratory, and yet a collection of more standard approaches to subsets of the problem may turn out to be more practical.

A third barrier relates to people's natural resitance to change. The computer-based systems in use today were designed around limitations existing at the time of their design. Among these limitations was certainly a lack of machine-intelligence capability. The introduction of such a capability into one part of a system cannot typically be predicated on rationalizing the overall system to take best advantage of the new capability. Introduction of new technology must be evolutionary, not revolutionary.

In summary, then, if you wish to create a product embodying machine intelligence, it does not suffice to have engineered a workable machine-intelligence-based solution to a perceived problem. The overall environment into which the new solution is to be placed, including existing hardware, software, and procedures, and, most of all, the existing people involved with the problem, must be taken into account.

REFERENCES

[1] Hendrix, G. G., Sacerdoti, E. D., Sagalowicz, D., & Slocum, J., (1978). Developing a natural language interface to complex data, *ACM Transactions on Database Systems*, 3, No. 2, 105-147.

[2] Shortliffe, E. H., (1976). *MYCIN: Computer-Based Medical Consultations*. New York: American Elsevier.

[3] Duda, R. O., Hart, P. E., Konolige, K., & Reboh, R., (1979). A computer-based consultant for mineral exploration, *Final Report, SRI Project 6415*. Menlo Park: SRI International.