THE DICHOTOMY BETWEEN MARKET NEEDS AND WHAT COMPUTERIZED IMAGE PROCESSING CAN PROVIDE

T. Kasvand
National Research Council of Canada, Ottawa

Abstract

Computer based picture processing or machine vision is beginning to move from the laboratory into a variety of practical application areas. Some of the reasons why this is still a trickle rather than a torrent are discussed.

DICHOTOMIE ENTRE LES EXIGENCES DU MARCHÉ ET LES POSSIBILITÉS OFFERTES PAR LES TECHNIQUES DE TRAITEMENT DES IMAGES PAR ORDINATEURS

Résumé

Les méthodes de traitement d'images par ordinateur ont commencé à sortir des laboratoires et sont utilisées dans toute une variété de domaines pratiques. On présente certaines des raisons qui font que ce mouvement est à peine amorcé.

INTRODUCTION

At present there is no mass consumer market for the products of computerized picture processing. A picture processing system in every home may be imagined only for far more advanced systems than now available. Unfortunately, we are unable to design a vision system for driving a car, weeding the garden or playing bridge. Such problems are unsolved due to lack of basic knowhow rather than computer hardware.

Even on the far simpler level there is a dichotomy between what the market needs and what our present knowhow and systems are capable of delivering. The market is "used to" employing the capabilities of the human visual system. These capabilities are very sophisticated, but most importantly of all, the job breakdowns have assumed these capabilities. The most efficient assembly line employing computerized vision systems will not be of similar construction to an assembly line employing humans. In certain of the "manual labour type" operations computerized vision systems could be employed now, in others the process needs redesigning. In situation where the worker is mobile, the "scene analysis" problem is usually beyond our knowhow.

On the so called "white collar level", the market wants results which are equivalent or better than those produced by an expert analyzing the same pictures or data. As yet, computers basically cannot deliver results of such quality and reliability. Even though experts frequently disagree or do not notice details of importance in a picture, such lapses are considered "human". Equivalent mistakes on the part of the computer, however, are not forgiven. Furthermore, human expectations are seldom stationary. Thus, even if the initial demands were met, the demands increase with the availability of new results, and may even reach a state where the expert wants his thinking to be done by the computer.

Divided we stand...

Nowadays it is considered acceptable and even "being with it" if one is in the position to push computer buttons, type messages, see character strings marching by on a CRT or orchestrate the performance by waving a light pen. With computer time sharing the individual operator can be replaced by many operators, allowing task simplification. The highly trained individual operator is no longer needed, being replacable by a less trained multitude. If the task breakdown is properly performed, throughput can be increased allowing mass production on the timehonored principles developed for assembly lines.

Possibly the only reason why this senario has not materialized is not its infeasibility but rather the lack of a market for its product. In computer-aided picture processing a market is developing, but it is a market for "speciality products" and not a mass market for a few standardized items. The techniques applied are equally diverse.

In very general terms, the products of computerized picture processing are of the following types:

a) Pictures produced for human study and analysis.
b) Picture context is analysed automatically and only the results are used or presented for human study.
c) The scene to be analyzed is under careful control.

The most familiar examples of pictures meant for human study are earth satellite photographs and scans, deep space probe photos, two dimensional reconstructions of x-ray absorbtion or isotope emission etc. These pictures are corrected to remove a variety of distortions, enhanced for better visibility of particular details and are often displayed in false colour or "perspective" for possible additional clarity (1). Some of these presentations may remain of lasting value, with others, however, when the initial "wonder" has faded, there is a need for further automatic analysis.

The automatic analysis of pictures, however, whenever feasible, is specialized for each particular problem. Three basic types of hardware are generally used to realize a solution:

i) Special purpose hardware built for the task
ii) General purpose computers with special hardware, programmed for the given task.
iii) General purpose computers without task specific hardware.

All these systems have their drawbacks as well as their advantages.

The special purpose hardware carries out its task quickly, or as quickly as is feasible at present, but at the expense of rigidity. The repertoir of available operations is severely limited. New processing algorithms, new modes of interaction etc are not easily realized. However, if the task can be sufficiently simplified, and if there is a need for many identical systems, the hardware cost can be minimized. Many such systems are already in use (2).

In systems using general purpose computers with special hardware, the hardware is used to speed up the processing and to simplify operator interactions. Such systems are quite expensive and, of course, the special hardware can inhibit the introduction of new ideas. Only in very large volume application are these systems warranted (3).

Picture processing systems based on general purpose computers without tasks specific hardware are expensive and slow, but they are as flexible as the ingenuity of the programmers can make them. Practically always, these systems are meant for research and development rather than for mass production. Feasibility studies are frequently done on such systems (4).

In some situations the designer of the system is in the position to specify the scene to be analysed. Of course, the temptation is to produce an economical system by minimizing scene complexity. Once such a system has established itself, the development tends to stop since more sophisticated systems cannot compete on an equal footing. Optical character reading (OCR) is a good illustration.

Market Needs

The market for vision systems is too diverse to be easily summarizable. However, from the viewpoint of how automatic vision systems are employed, the following operational areas exist:

i) The vision system is autonomous, ie not under direct operator control.

ii) The vision system is part of an instrument in daily use. The device is interactive allowing operator intervention.

iii) The vision system is part of a computer installation in a laboratory. The system is highly interactive and versatile but basically meant for teaching, research and development.

Examples of all these categories have already been mentioned.

In the long term, the greatest need is expected to be for autonomous systems. These, however, are the most difficult ones to design unless the problem areas can be reduced to sequences of relatively simple operations. This is often possible. For example, all our present production machinery is "blind". Devices to capture the necessary images in computer-processable form are available commercially. Manipulators to handle objects are also available. However, the configuration of the assembly line relative to the sensor and the manipulator as well as the illumination to obtain proper image contrast is a novel design problem. In ordinary situations these systems will start to compete with the assembly line worker. This complication will not exist in hostile environments. The beginnings of autonomous visions systems are in the laboratory stage (5).

The design of vision systems for instruments or devices which rely on the human operator tends to be much simpler than the design of autonomous visions systems. The operator and the machine have complementary characteristics. Pattern recognition is easy for man, measuring and calculating is easy for the machine. By proper interfacing, the strong points of each can be used to best advantage. In a well designed system the operator can check the machine's results at a glance. Such systems can be very successful. There is a temptation, however, to use such systems also on problems for which they were not designed.

Quite a variety of interactive systems is already available. Baring certain social, legal and cost problems, many more devices will come into common use. For example, automatic prescreening of medical images or measurements where only the "special cases" are brought to the doctor's attention. Extraction of the needed data from aerial photographs, satellite sensed data, technical photographs and surveys, etc. etc.

The vision systems laboratory should, of course, be the source and testing ground of new ideas rather than a consumer. In some simpler vision problems it may be possible to minimize or even bypass the laboratory development and testing stages. However, since the intricacies of the vision problem are easy to overlook, it is not known how many design attempts have failed.

<u>The next step</u>

Large amounts of literature are available on pattern recognition and picture processing. Even though most of the results are not directly usable for practical problems, the comparison of market needs with prior results tends to define new directions for further research and development.

The basic impediments to quicker progress for totally autonomous vision systems are:

i) the cost effectiveness of the devices.
ii) the need to reformulate commonly accepted procedures.

iii) the lack of operational similarity between human and computer vision.

The first point refers to the design and hardware costs of presently realizable computer vision systems. The hardware costs, however, are decreasing rapidly. This leaves the design costs and the need to redesign the procedures which have evolved from our own visual abilities. The feasibility of redesigning a process to accomodate a computer vision system can only be judged in each individual case. Generalizations are difficult. To try to bring the operational capabilities of computer vision closer to those of human vision is at present only a research goal.

The situation, however, is much simpler with the so called interactive vision systems. The cost effectiveness problem still remains, but the machine is only an assistant to the operator. Problems which the machine cannot handle are simply referred to the operator for a decision. The opposite is also true, problems which the operator cannot handle are referred to the machine for decision. For proper design of interactive systems, it is imperative to recognize the weaknesses and the strong points of the human and the machine's capabilities.

The human system or operator has during his evolution developed a very sophisticated pattern recognition capability. This ability the computers cannot match. However, in terms of calculating power or number crunching, human intuition is no match at all to our computers. The powers are complementary at present. However, human abilities cannot be improved while no corresponding barrier exists for computers. Consequently, for versatile communication between man and computer, the computer input and output has to match human capabilities and not vice versa.

Stated somewhat unkindly, the interaction with a computer occurs under two basic premises:

i) the uncomprehending computer
ii) the uncomprehending operator.

The first normally occurs in production, the other in research environment.

In pattern recognition tasks involving "normal" images the computer is the uncomprehending partner compared to ourselves. The interaction consists of outlining to the computer the picture areas to be processed, extracting it from difficulties and correcting its results. Both the details of the interaction and the form in which the results are presented will be highly problem specific.

In all but the simplest classification tasks, for example, the "images" are beyond human comprehension, while the mathematical representation of the problem is formally simple. Our "visual understanding" fails as soon as the problem is "out of this world". (Example: given 3 points in 4 dimensional space, which two are closest?). In such situations the interface equipment is used to display comprehensible transformations of the original problem or its solution.

Conclusion

The field of computer analysis of pictures or pattern recognition and scene analysis is approximately 25 years old. Attempts at character recognition started about 40 years ago. There are very few completely autonomous vision systems in operation. A variety of interactive systems are commercially available. Many diverse vision systems exist in the form of computer programs in laboratories. There is a profusion of literature. This is indicative of a field of endeavour that is just beginning to yield commercially feasible products.

However, in Canada at least, as far as the author is aware, there is a gap between those with a problem to be solved and those with facilities and some prior experience. Practical applications of computer vision are experimented with in some companies. Presumably due to the novelty of the field and the very modest funding, much of the effort is expended on setting up elementary laboratory facilities (rotating drums as scanners, photocell arrays for sensors, microcomputers with minimal memory and special circuits for analysis of the signals, etc). Little if any attention seems to be paid to prior experiments (in literature) and experiences gained, resulting only in rediscovery of known procedures and effects. At the same time, laboratories with better facilities for studying their specific problems appear not to be interested.

References

I) Examples of "picture production systems".

a) Deep space probe photographs, (Jet Propulsion Laboratory, see Ref. B).

b) Multispectral images of earth, (Canada Center for Remote Sensing).

c) X-ray based reconstruction of two dimensional images, (EMI, Ohio Nuclear).

d) Radio-isotope based reconstructions, (J & P Engineering, New Scientist, Jan. 27, 1977, p. 203).

e) General purpose displays: (Comtal, Norpak).

2) Examples of special purpose (interactive) instruments.
   a) Metallographic analysis:
      (Quantimet, Wild-Leitz system, Zeiss system).
   b) Cell analysis:
      (Perkon-Elmer Corning Electronics, Smith Kline Inc.) see: K. Preston Jr. Digital Picture Analysis in Cytology, Ref. A.

3) Examples of special purpose large scale systems:
   a) Applications in high energy physics: (System at CERN (Geneva), University of Toronto). See: R.L. McIlwain Jr., Image Processing in High Energy Physics, Ref. A.
   b) Applications in remote sensing: (Image 100, CDC system, Bendix system).
   c) More general applications: (Wild-Leitz TAS, Toshiba system).
   d) JPL Chromosome Karyotyping System. (K. R. Castleman)

4) Picture processing laboratory systems. These exist at every major research institution and university. (SRI,JPL,USC,MIT,Purdue,McGill,Lawerence Livermore, Electrotechnical Laboratory Tokyo,...)

   For a good illustration of how image processing problems are approached, see: C.A. Harlow, On Radiographic Image Analysis, ref. A.

5) Examples of autonomous systems in special application areas: See: W. A. Perkins: Multilevel Vision Recognition System, Ref. C. Industrial Applications: Ref. C.

A) Topics in Applied Physics, vol. 11, Springer Verlag 1976, editor A. Rosenfeld.

B) Scientific Findings from Mariner 6 & 7 Pictures of Mars, Journal of Geophysical Research, vol. 76, no.2, Jan. 10, 1971.

C) The Third International Joint Conference on Pattern Recognition, Coronado, Nov. 8-11, 1976.