Coming to Terms with Machine Vision and Computer Vision:

And it’s not just a matter of word choice. The way these concepts are perceived can make a big difference for everything from practical project implementation to university funding.

By Bruce G. Batchelor

There is often a natural assumption among non-specialists that Machine Vision (MV) and Computer Vision (CV) are synonymous because they both refer to Artificial Vision. Indeed, the distinction between MV and CV terms is not accepted by all researchers in Artificial Vision. I recently had the opportunity to discuss the meanings of these terms with another researcher. He is a very able and active worker and spends his days devising new mathematical and algorithmic techniques for analyzing and understanding images. He works in what I shall refer to here as Computer Vision, and regards the terms Machine Vision and Computer Vision as conveying exactly the same meaning.

I strongly disagree with this point of view, and in order to convince him that is a fundamental difference between these subjects, I challenged my debating partner to design a vision system that is able to decide whether a bright “silver” coin is lying with its obverse or reverse side uppermost. By his own admission, he could not even begin to do so! The reason for this is simple: he had not been trained in an appropriate practical subject (e.g., engineering or physics) and therefore felt unable to design a lighting-viewing system. Since the starting point for this challenge is a coin, not a computer file, the relevant mechanical, lighting, optical and camera sub-systems must be designed first—before any work on developing or implementing image processing algorithms is even contemplated.

Coming to terms

Before I begin, I should explain that I am an engineer by both natural inclination and training. Until six years ago, I worked in a university engineering department, although I am on the faculty of a Computer Science department now. I am therefore in a good position to observe the different ways that engineers and computer scientists approach working on artificial vision.

I contend that MV requires skills and knowledge that CV ignores. And I do not mean to belittle the intellectual requirements of Computer Vision in any way. On the contrary, it is my firm belief that CV is an intellectually “deep” subject, while Machine Vision requires a broader range of both knowledge and skills. My aim, in writing this article, is to plead for mutual respect between the practitioners of what I perceive as two disparate subjects.

Unfortunately, respect is lacking in some quarters at the moment. I assert that MV and CV should be regarded as distinct academic and commercial subjects, that each should have its own institutions for organizing meetings, publishing results and honoring its leading exponents and possessing its own funding program for research. I also maintain that anybody who attempts to apply vision systems to “real-world” industrial applications in a casual way is at best naïve and at worst intellectually dishonest; MV deserves full-time specialist attention.

As further evidence of the lack of a distinct identity for MV, many researchers blur the boundary separating it from the related subjects of Artificial Intelligence (AI), Digital Image Processing (DIP) and Pattern Recognition (PR), sometimes referred to as Neural Networks nowadays. MV is often confused with DIP, because the latter is perceived by some workers as being the key technology for Machine Vision. I would express matters a different way: DIP is just one of many technologies, all of which are essential to produce an effective and successful MV system. DIP is also an essential component of Computer Vision.

Pattern Recognition (and, to a smaller extent, Artificial Intelligence) is sometimes perceived as encompassing aspects of MV, although the distinction is clearer than for CV. Practitioners in one of these subjects should not expect to be able to undertake Machine Vision work on a casual part-time basis.

Several other terms need to be mentioned here: Automated Visual Inspection and Visual Process Control/Monitoring are certainly legitimate terms, referring to industrial MV. However, the term Robot Vision is used ambiguously, since it is often applied to the theoretical aspects of robot control (i.e., as a subset of AI), as well as to the study of practical systems. Image Analysis, another term in widespread use, refers to deriving and manipulating quantitative measurements from pictures, including, of course, in industrial applications. Many people refer instead to Image Understanding. However, these are nuances and certainly distract us from the main theme of this article.

Machine Vision requires an engineering to system design.

All of the terms mentioned above are used in technical literature, and by speakers at numerous business, meetings and conferences, in an imprecise way. There is no doubt that the sloppy use of these terms has been the cause of a considerable amount of confusion. But that’s not the crux of it. The confusion caused has often resulted in unsatisfactory service being provided to customers of industrial vision systems. It has also hampered research, and has probably hindered career progression of engineering staff.

Learned societies and companies

The problem of nomenclature arises because MV, CV, and DIP (and sometimes AI and PR) are all concerned with the processing and analysis of pictures within electronic equipment. Notice that I did not mention computers explicitly, since not all processing of pictures takes place within a device that is recognizable as a computer. By definition, the technical definition of CV assumes that these activities take place inside a computer.

On the other hand, MV does not impose such a limitation, although computer-based processing of images is encompassed by this term. MV allows the processing of images to take place in specialized digital networks, arrays of field-programmable gate arrays (FPGAs), multiprocessor configurations, optical/opto-electronic computers, analog electronics, or hybrid systems.
Table 1: Comparing Machine Vision and Computer Vision. Entries in the Machine Vision column relate to the factory-floor target machine, unless otherwise stated.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Machine Vision</th>
<th>Computer Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic/practical motivation</td>
<td>Practical</td>
<td>Academic</td>
</tr>
<tr>
<td>Advanced in theoretical sense</td>
<td>Unlikely: (Practical issues are likely to dominate over academic matters.)</td>
<td>Yes: Academic papers often contain a lot of “deep” mathematics</td>
</tr>
<tr>
<td>Cost</td>
<td>Critical</td>
<td>Likely to be of secondary importance</td>
</tr>
<tr>
<td>Dedicated electronic hardware for high-speed processing</td>
<td>Very likely</td>
<td>No (by definition)</td>
</tr>
<tr>
<td>Designers willing to use non-algorithmic solutions to solve problems</td>
<td>Yes (e.g., systems are likely to benefit from careful lighting)</td>
<td>No</td>
</tr>
<tr>
<td>In situ programming</td>
<td>Possible</td>
<td>Unlikely</td>
</tr>
<tr>
<td>Input data</td>
<td>A machine part, piece of metal, plastic, glass, wood, etc.</td>
<td>Computer file</td>
</tr>
<tr>
<td>Knowledge of human vision influences system design</td>
<td>Most unlikely</td>
<td>Very likely</td>
</tr>
<tr>
<td>Most important criteria by which a vision system is judged</td>
<td>(a) Ease of use (b) Cost-effectiveness (c) Consistent &amp; reliable operation</td>
<td>Performance</td>
</tr>
<tr>
<td>Multi-disciplinary</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Nature of an acceptable solution</td>
<td>Satisfactory performance</td>
<td>Optimal performance</td>
</tr>
<tr>
<td>Nature of subject</td>
<td>Systems Engineering, practical</td>
<td>Computer Science, academic (i.e., theoretical)</td>
</tr>
<tr>
<td>Operates free-standing</td>
<td>(a) Interactive prototyping system must be able to interact with its human operator (b) Factory floor (target machine) must be able to operate free-standing</td>
<td>May rely on human interaction</td>
</tr>
<tr>
<td>Operator skill level required</td>
<td>(a) Interactive prototyping system: medium / high. (b) Factory floor (target machine) must be able to cope with low skill level</td>
<td>May be very high</td>
</tr>
<tr>
<td>Output data</td>
<td>Simple signal to control external equipment</td>
<td>Complex signal for human being</td>
</tr>
<tr>
<td>Speed of processing</td>
<td>(a) ISP must be fast enough for effective interaction (b) Real-time operation is very important for TM.</td>
<td>Not of prime importance</td>
</tr>
<tr>
<td>User interface</td>
<td>Critical feature for ITP and TM</td>
<td>May be able to tolerate weak interface</td>
</tr>
</tbody>
</table>

It must be understood that MV, CV and DIP share a great many terms, concepts and algorithmic techniques, but they have a completely different set of attitudes and priorities. (See Table 1 above.) To summarize, the division between MV and CV reflects the separation that exists between engineering and science. Just as engineers have long struggled to establish the distinction between their studies and science, I am trying to do the same for MV and its scientific/mathematical counterpart (i.e., CV).

Should know better?

Confusion of terms is even manifest in the names of learned societies—and this is true worldwide. For example, here in the U.K., the British Machine Vision Association holds a series of conferences in which almost all of the papers are on what I maintain is Computer Vision. There is certainly a great difference between the articles included in the proceedings of the BMVA on the one hand and those included in conferences organized by the Machine Vision Association of the Society of Manufacturing Engineers (SME) on the other. Indeed, it is doubtful whether many of the papers that are acceptable to a BMVA audience would be welcomed at an MVA conference, or vice versa!

Commercial organizations sometimes confuse matters too. For example, some companies employ phrases of their own invention, such as: “image recognition,” “vision analysis,” “scene processing,” “optical verification,” etc. The use of non-standard terms such as these makes it difficult for would-be customers to compare competitors’ products and services.

Some general computer companies, including one very large and influential software supplier, are lately using the term “Computer Vision” to refer specifically to gesture recognition. (This is used to assist communication between a desktop computer and its user.) This does concern image recognition of a sort, but has nothing to do with Machine Vision, since very little effective control can be exercised over the operating environment. (Would you sit in a darkened room, under ultra-violet lamps, wearing fluorescent gloves, to operate your computer?) However, there now exists a distinct danger that the term “Computer Vision” will become widely used through association with this particular application, and from there become adopted as the umbrella term for all types of artificial vision systems in the public’s mind. It would be a shame if a late arrival such as this redefined the terminology, causing even greater confusion. (And some would say that we’ve already been through a similar adventure with the document processing firms who announced that they were the “image processing industry”.)

Applications

Industrial applications of Machine Vision include such tasks as inspecting, measuring, verifying, identifying, recognizing, grading, sorting, counting, locating and manipulating parts, monitoring/controlling production processes, etc. Vision systems may be applied to raw or processed materials, piece parts, packing, packaging materials and structures, continuous-flow products (webs, extrusions, etc.), simple and complex assemblies, or to tools and manufacturing processes. MV systems can also be applied to complex systems (e.g. complicated physical structures, wiring looms, pipework, transport mechanisms, etc.) and to a wide range of factory-management tasks (e.g. monitoring material stocks, checking the floor for debris, controlling cleaning, ensuring safe working conditions for human operators, etc.).

“Machine Vision,” as we are defining it here, is also applied to a variety of other areas with considerable significant commercial importance and academic interest—and of obvious ongoing interest to readers of Advanced Imaging:

• Aerial/satellite image analysis
• Agriculture
• Document processing
• Forensic science, including fingerprint recognition
• Health screening
• Medicine
To summarize, there is no excuse for fooling a customer who has no detailed knowledge of either CV or MV into thinking that a specialist in CV is competent to design a system for use in a factory. Unfortunately, intellectual honesty is not universal, and the consequences of misleading customers in this way can have a major impact on the wider, long-term public perception of a developing technology such as MV. There are subtle ways in which erroneous views can be propagated and perpetuated. Every experienced MV engineer knows that very subtle points can sometimes lead the undoing of an otherwise successful application.

### Systems Engineering

A person working in MV needs to be a Systems Engineer, not a scientist. (See Table 2 at left.) There is a very marked difference in attitude and culture between Computer Science and Systems Engineering. Indeed, there is a deplorable lack of understanding in the public's mind about the relation between the broader aspects of engineering and science. This is very often reflected in the misuse and over-use of the word "scientist" in newspapers, television and radio news programs. The misuse of the still less-common terms "Machine Vision" and "Computer Vision" is therefore hardly surprising! It is important for the Machine Vision specialist to educate the customer at the first and every other opportunity.

Now, some tasks are intermediate in having features of both MV and CV, as I have described them above. Indeed, the situation that I have described is oversimplified to make a point. In reality, there is a continuum between the extremes summarized in Table 1. However, this does not in any way diminish the strength of the general argument that MV and CV should be regarded as being distinct subjects.

My motivation for writing this article is to plead for a greater degree of respect for Machine Vision, particularly among our colleagues working in Computer Vision. But we must always be aware that real respect requires a two-way process of acknowledging each other's merits and using their methods and talents when it is appropriate to do so. Machine Vision engineers, for their part, must keep abreast of developments in Computer Vision, because many of them will, sooner or later, be useful in solving significant practical problems in industry.

---

To sum up, the basic tenet of all work in MV is that it is unacceptable to limit our concern to the mathematics/algorithms, or to any other single aspect of the task in hand. MV is fundamentally multidisciplinary in nature. Striking the right balance is the skill of a good MV system designer.

### Just Words?

Does it matter what we call a subject? The answer, most definitely, is, "Yes!" Technologists have stopped using phrases that were once very popular. Perhaps the most notable example is Cybernetics, although this name is discredited in many people's minds now. There are several reasons why MV deserves to be identified as a subject in its own right.

- To ensure proper refereeing of academic papers.
- To promote appropriate funding of research.
- To protect customers of vision systems from claims of expertise in MV by workers who have a narrow field of expertise and are therefore not competent to design a complete system.
- To protect the careers of its practitioners.

Of these, the third is perhaps the single most important reason for the need to establish the identity of MV, since a great deal of harm has already been done by specialists in CV, DIP, PR (and, to a lesser extent, AI) dabbling in Machine Vision who are not competent to do so.

Ignoring the image acquisition sub-system is a sure way to generate an over-complicated image processing solution to a problem that could be solved better by a simple adjustment to the lighting or optics!

The classic problem of this type is that of thresholding an image. There are many optical ways to achieve high-contrast images and thereby circumvent the problem. However, many vision systems have been installed that use excessively complicated algorithms to compensate for poor lighting. Many other vision systems have been less reliable than they might have been had they been designed from the outset with a good appreciation of all of the "systems issues" involved.

---

*Machine and (Computer!) Vision veteran Bruce Batchelor is with the Department of Computer Science, Cardiff University, P.O. Box 916, Cardiff, Wales, U.K. He can be e-mailed at bruce@cs.cf.ac.uk, and has much additional useful material on these and related matters at URL: http://bruce.cs.cf.ac.uk/bruce/index.html.*