PERGAMON

Engineering Applications of Artificial Intelligence 11 (1998) 161–162

Engineering Applications of
ARTIFICIAL INTELLIGENCE

# Preface to the special section on machine vision for intelligent vehicles and autonomous robots

Alberto Broggi *

*Dipartimento di Ingegneria dell'Informazione, Università di Parma, I-43100 Parma, Italy*

Received 1 December 1997

Machine Vision, as one of many ways of sensing the surrounding environment, has several advantages over radio, acoustic and tactile sensors (such as radars, lasers, sonars, or bumpers), such as the capability to acquire data in a non-invasive way, thus not altering the environment.

The use of active sensors, on the other hand, involves a method of measuring the alterations in signals emitted by the sensors themselves, and has some specific characteristics:

- Active sensors can measure quantities in a more direct way than vision can. As an example, a Doppler radar can directly measure the relative movements between an object and the viewer, while vision can detect movement only as a result of complex processing of image sequences.
- Active sensors require lower capability in the computing resources employed, as they use a considerably smaller amount of acquired data.

In many indoor applications, such as the navigation of autonomous robots in unknown settings, vision and active sensors can perform complementary tasks in the recognition of objects, detection of the free space, or checking for some specific object characteristics. Unfortunately, when several robots are moving within the same environment, active sensors may interfere with each other, thus decreasing their reliability and usefulness. This problem becomes even harder in an unstructured outdoor environment, in which a large number of vehicles could be moving simultaneously, such as – for example – when autonomous vehicles move on (intelligent) highways.

Hence, in cases in which a massive and widespread use of autonomous sensing agents is envisaged, the use of passive sensors such as cameras has definite advantages over active sensors. These are the cases in which vision becomes of paramount importance.

Furthermore, the recent advances in computational hardware, such as a higher degree of integration and a reduction in power supply voltages, allows users to employ machines that can deliver high computational power, with fast networking facilities, at an affordable price. Since the early stages of vision (low-level image processing) are computationally demanding, the availability of low-cost engines has helped to solve some basic bottlenecks. Current technology allows the use of SIMD-like processing paradigms, even in general-purpose processors such as the new generation of processors that include multimedia extensions (MMX).

In addition to this, current cameras include new, important features that allow some basic problems to be addressed and solved directly at the sensor level. For example, image stabilisation can now be performed during image acquisition, while extension of the dynamics of cameras can eliminate the processing required to adapt the acquisition parameters to specific light conditions. The resolution of sensors has also been drastically enhanced. In order to decrease the acquisition and transfer times, new technological solutions can be found in CMOS cameras, with the important advantages that pixels can be addressed independently (as in traditional memories), and that their future integration on the processing chip seems to be straightforward.

These advances in the technology have not only promoted improved hardware devices, but have also triggered a renewed interest in the use of techniques for the processing of iconic information, generally addressed by the field of artificial intelligence, and dealing with image interpretation and – more generally, when the fusion with data from other sensors is also integrated – with perception. The problem of understanding the mechanisms underlying the process of human visual perception, with the objective of its

* Corresponding author. Tel.: +39 521 905 707; fax: +39 521 905 723; e-mail: broggi@ce.unipr.it.

0952-1976/98/$19.00 
PII: S0952-1976(97)00072-9

emulation on a computer system, is one of the great challenges of artificial intelligence. The main purpose is to obtain high-level descriptions of scenes, through the interpretation of visual data. This complex process involves different steps, all related to different aspects of perception:

- from the processing of iconic data at different resolutions, to the use of a fovea-like focus-of-attention mechanism, at the lowest level;
- from the identification of the features of interest, to their selection and extraction, based on their significance, at the medium level;
- from the high-level grouping of information, to a synthesis of the description of the scene.

The success of computational approaches to perception is demonstrated by the increasing number of autonomous systems that are now being used in structured and controlled industrial environments, and are being studied and implemented to work in more-complex and unknown settings.

This special section presents eight papers dealing with different aspects of, and proposing different solutions to, the problems of perception and movement in a surrounding environment. They involve both the theory and various applications of AI techniques, in different fields, ranging from indoor settings to automotive applications, and from cross-country surroundings to underwater environments.

I would like to thank the Executive Editor, Rob A. Vingerhoeds, for giving me the possibility of presenting this extremely up-to-date field of research in this journal. I would also like to thank Alessandra Fascioli for her outstanding help for the management of the special section.