7. DISCUSSION

This paper gave a discussion about nonsymbolic methods for artificial intelligence. A hypothesis was given that many tasks in artificial intelligence do not require symbolic data structures, and symbolic manipulations with them. Another solution to these problems is via numeric and subsumbolic computing, and the new trend in computer science, neural networks, seems be a way to do it. It may be too early to say if neural networks can really solve hard problems, like computer vision, but at least it's a way to overcome some of the barriers we are faced today.

References

Arbib, M.A. (1988). Neural computing: the challenge on the sixth generation. Reprint from the *EDUCOM Bulleting, 23*(1), pp. 2-12.

Ballard, D.H., Brown (1982), C.M., Computer vision, Englewood Cliffs, New Jersey:Prentice-Hall.

Drefus H.L. & Drefus S.E. (1986). *Mind over Machine.* New York: Free Press.

Jain A.K. (1989). *Fundamantals of digital image processing.* Englewood Cliffs, NJ: Prentice Hall.

Lindsay, P.H. & Norman D.A. (1977), *Human Information Processing, 2nd.* Orlando, Florida: Academic Press.

McCarthy, J. (1960). Recursive functions of symbolic expressions and their computation by machine. *Communications of the ACM,* vol. 3, pp. 184-195.

Minsky M. (1975). A Framework for representing knowledge. In *The Psychology of computer vision* P. Winston (Ed.). New York: McGraw-Hill.

Schank R. C. & Abelson R. (1977), *Scripts, plans, goals and understanding.* Hillsdale, N.J.: Lawrence Erlbaum Assoc.

Searle, J. (1980). Minds, Brains and Programs. *The Behavioral and Brain Sciences 3*(3), pp. 417-457.

Quilliam M.R. (1968). Semantic memory. In *Semantic information processing* M. Minsky (Ed.). Cambridge, Mass.: MIT-Press, pp. 227-270.

Smolensky, P. (1987). Connectionist AI, symbolic AI, and the brain. *Artificial Intelligence Review 1.* pp. 95-109.

Winston, P. (1977), *Artificial Intelligence.* Reading, Massachusetts: Addison-Wesley.

lower parts of the structure are perdefined and fixed; in the brain the synaptic placticy is frozen at early age and in computer vision standard feature extraction techniques and unsupervised processing are commonly used. Higher levels of information processing involve memory related processing where supervised learning is used. Also extracted features like lines, corners and edges are common in both structures. Segmentation and feature clustering finds evidence from human information processing; see discussion by Linsday and Norman (1977) for example. The operation types in the computer vision are probably similar but the actual functions and their implementation is different; for computer vision systems see references Ballard & Brown (1982); Jain (1989).

Assuming that the overall structure depicted in figure 4. is correct, the task is then to replace the symbol processing steps with subsymbolic ones and see that the lower processing levels provide enough information for the higher levels, where supervised learning is used. For most of these operations subsymbolic methods can be found by using neural networks. The problem is that the current knowlede is still on the network level, while pratical computer vision systems require system level approach. We need to find mechanisms that explain howto build systems from different types of neural networks.

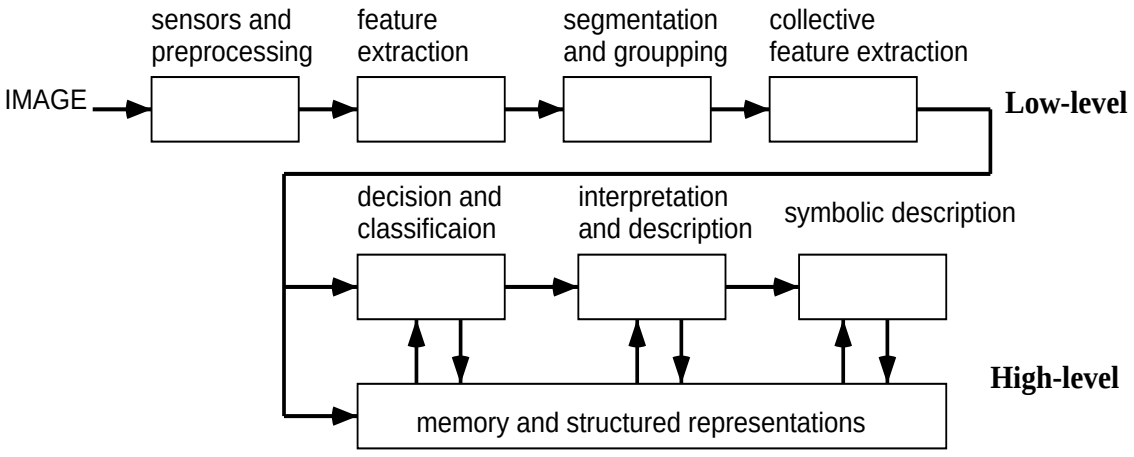


Figure 4. A typical computer vision system architecture.

parallel low level information into symbols too early. When the information is refined in the successing stages of recognition and image understanding, no total loss of parallelism should occur. Figure 3. illustrates in context of computer vision how the parallelism in the information flow and the information selectivity are related. It is suggested here that optimal performance is achieved when these two are in balance with respect to total time complexity of processing. In the lower levels of processing the algorithms should be more data-parallel than in the higher levels, where structural parallelism is to be used. The result should be a feed forward or a pipeline type of pattern recognizion system, where the total (parallel) processing time should be the same at and stage of the processing. It is obvious that the human brain is capable for this. This would also imply that symbolic processing techniques like backtracking search must not be used.

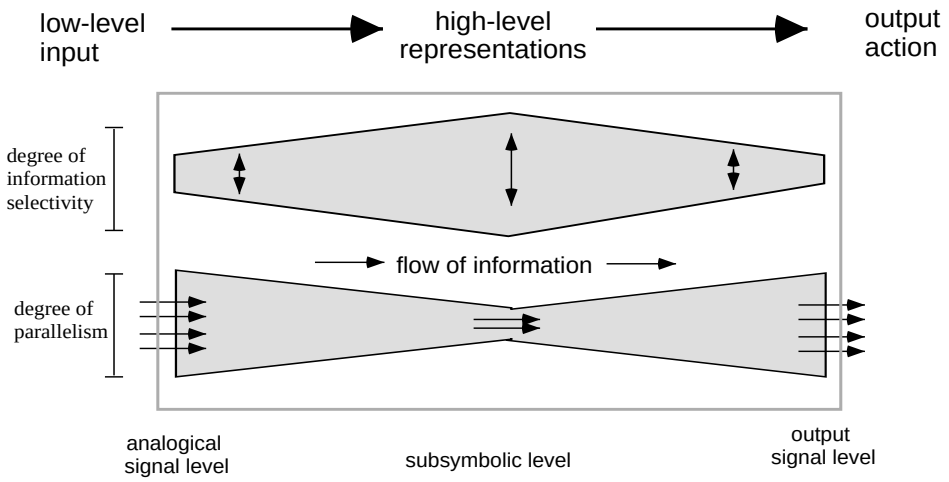


Figure 3. The relation between the data rate and the algorithm complexity in the vision problem.

6. A SUBSYMBOLIC COMPUTER VISION ARCHITECTURE

The structure of a computer vision system such as the one shown in figure 4. is composed of a succession of levels where the information is refined. Grounded on the current knowledge the anatomy of a biological visual system and the construction of it's the computerized counterpart offer many points of resemblance. The

user have their own internal representations of the world and the outputs from the computer and the man can be translated trough a man-made language of symbols.

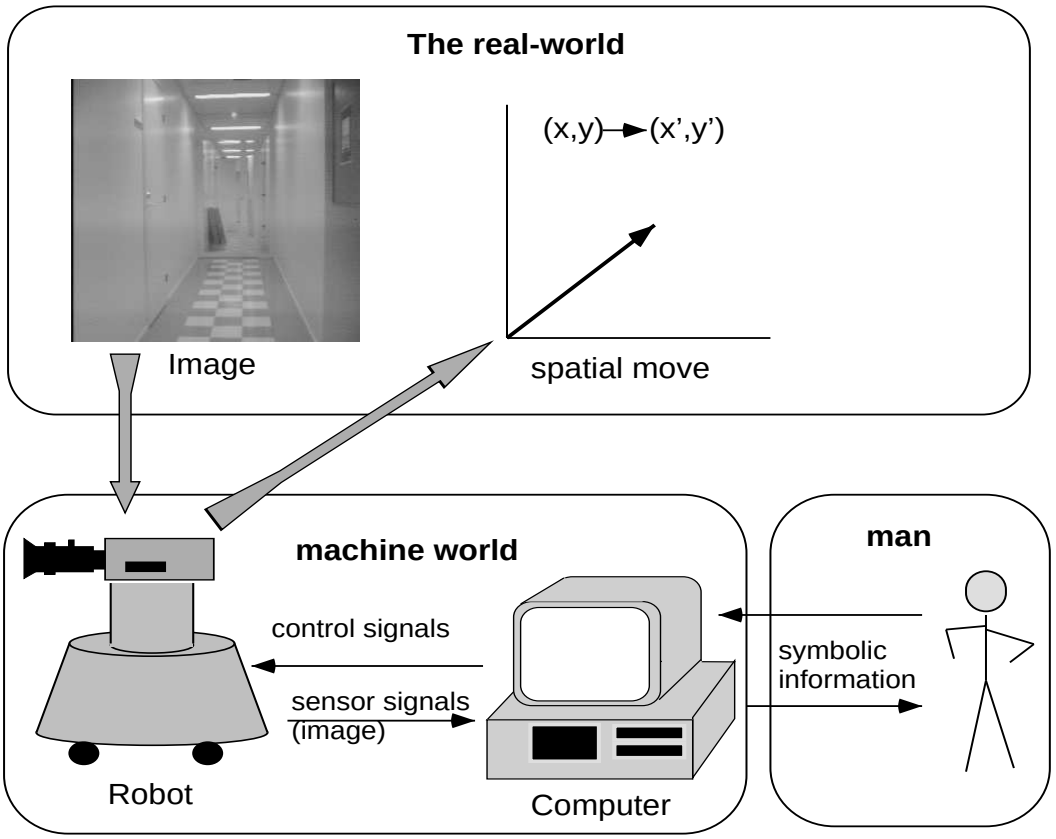


Figure 2. The world, the computer and the man. A robot makes a move in a corridor from position (x, y) to position (x', y'). Control signals for the movement are computed by using information that is extracted from the image without symbolic processing and symbolic description of actions is given for the user.

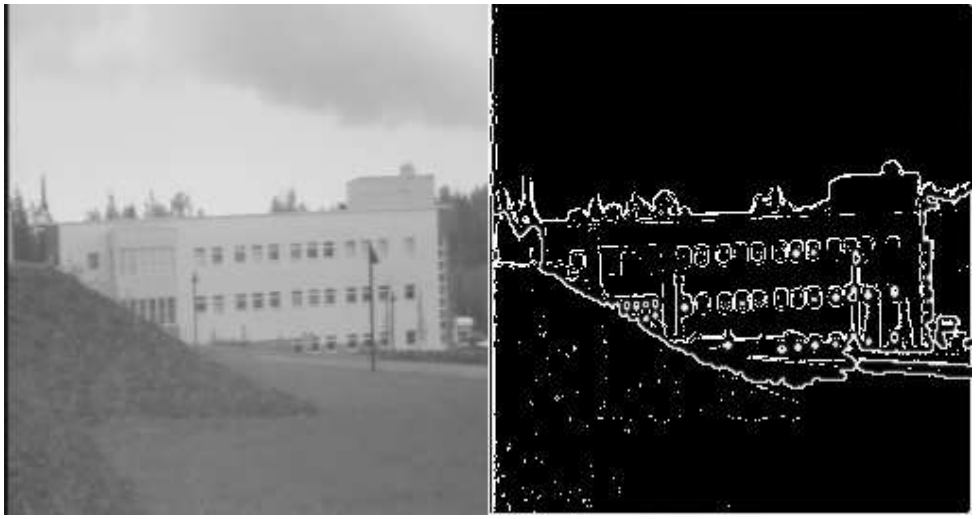
A source of inspiration for the recent models in computer vision has been the biological visual system. Similar motivations in the development of other fields of applied intelligence can be observed simultaneously. The resulting boom of *neural networks* have given new insights into how the connections between the real, physical, world and conceptual processing should be combined.

One keyword seems to be parallel processing. We should not squeeze the massively

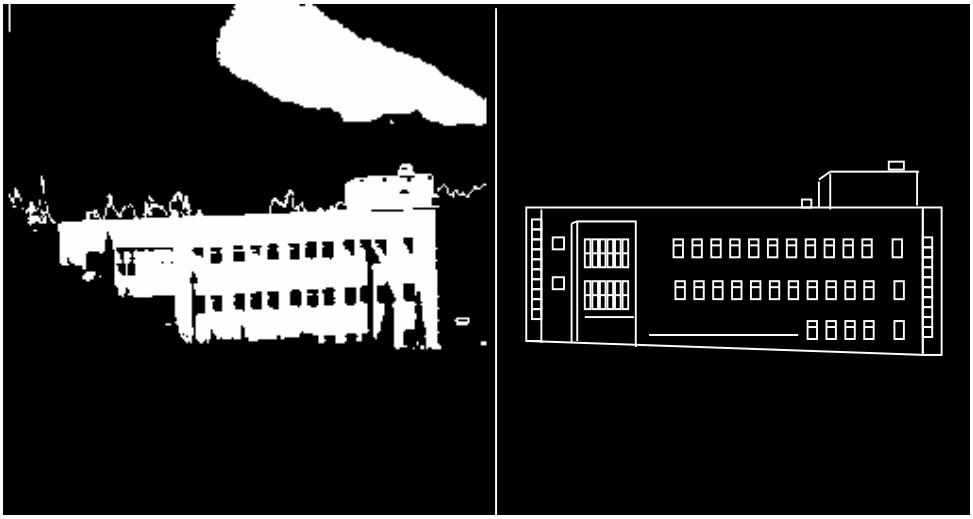
rules or anything that can be used for knowledge representation, matching, reasoning, planning and control operations. In pratice high-level vision takes advantage of symbolic processing and AI to do make desicions based on the information of the image. Problems in AI have become problems in computer vision. In the case of matching for example, a graph or similar structure derived from the image is put into correspondence with existing models. This is a typical combinatorically explosive problem where the search space of answer grows in orders faster than the dimension of the problem variables, i.e. big problems can't be solved with any computer. Since automated symbolic representations from images are usually incomplete and complex, many pratical applications that are simple for humans prove to be difficult for the computer.

5. THE HYPHOTHESIS

The hypothesis is that symbols and symbolic representations are not required for many pratical computer vision problems. Even stronger claim could be given with some applications, saying that symbolic computing in computer vision can be harmful. If one looks a computer vision system as a black box that takes sensory inputs, the output of the system is then a real-world action or an information for the user. Only in the second case the system must give the user symbolic description of the image, "there is a red ball and blue cube" or "50 per cent of the image is classified as forest", for example. Traditionally this conversion from numeric signals to symbols takes place between the low-level and the high-level processing but there is no reason why the translation into symbols could not be done as a very last and final step of the image analysis. The analysis itself would then be numerical or subsymbolic. In the case of real-world actions no symbolic interpretion would be required since the output is numeric as well. Symbolic representations are required only as a tool to transfer information between the man and the computer; information flow between the computer and the world is numerical. Also interaction between the world and the human brain is numerical (or analog) up to levels where brain behaviour has been resolved. These relations suggest something like figure 2. where both the computer based system and the



(a) Raw image (b) Edges



(c) Thresholded (d) Drawing

Figure 1. Example representation of an image; (a) raw, (analogical); (b) detected edges; (c) thresholded; (d) human drawing.

High-level vision and image understanding relates structured representations that are extracted from the image with internal structures describing the world in the scope of the application. Internal structures may be models of the objects, decision

4. COMPUTER VISION: HIGH-LEVEL AND LOW-LEVEL

In computer vision the problem solving has traditionally been separated into low-level, dealing with physical data, and into high-level, doing AI type of processing for image understanding. A good introdutory to low-level versus high-level computer vision is given for example by Ballard and Brown (1982). Low-level and high-level paradigms, however, are quite different and many attempts to fill the gap by using different types of intermediate processing have been proposed. So far there are no good solutions howto combine the world of numbers and the world of symbols. In his well known criticism Smolensky (1987) uses the word abyss to express how deep the difference between these two types of information is.

Some low-level operations for a real world image (1a) are illustrated in the figures 1b and 1c. Figure 1d is a human drawing of that is seen in figure 1a. For human eye an image is not only a thresholded shape or a boundary line but entireness given in the context of many kinds of knowledge. The drawing in figure 1d is good example since there is information of the building that the original image 1a does not contain. If one should make a computer vision system that is capable for human type of information processing then one should use high-level knowledge during the recognition and image analysis process. Traditionally this high-level information has been stored by using symbolic data structures that have proven so useful for the man-machine interaction. This symbolic information is then matched with the symbolic representation extracted from the image to make object recognition, classification etc. A typical way to make these representations is to segment the image into regions that have similar semantic meanings: sky, building and tree. Segmentation is possible if low-level information introduces enough evidence for the task. In the given example the image could be simply thresholded for black, medium gray and light gray levels of intensity resulting tree segmented ares. In most practical situations more sophisticated methods are required. After segmentation one should name the segments and make a syntactic representation, a relational data structure of segments with connectivity rules.

Winston's approach is now known as "weak AI" (Searle 1980) that sees the computer as a powerful tool in the study of the mind. Strong AI takes more rash hypothesis: "appropriately programmed computer can really be a mind".

A striking characteristic of AI type of computing has been that it deals mainly with non-numerical information, symbols. Second important feature in almost every AI system is a data base for knowledge representations. There are various suggestions how these data structures should be construced, for example: semantic networks (Quilliam 1968), frames (Minsky 1975) and scripts (Schank & Ableson 1977). In order to solve pratical problems many types of heuristics have been developed to extract useful information from man-made representations of knowledge; see Winston (1977) for examples.

In reality AI was never able to realize all promises that were made; requirements like fault tolerance and fast search of answers proved to be difficult when using symbolic representations. The most significant failture of AI, according Drefus & Drefus (1986), was however so called common-sense problem: how to store and access facts about everyday world human beings seem to know. As a result AI has not been able to make good connection with the real, physical and numerical world. It also seems unlikely that learning in biological systems could be a search process using symbolic (and serial) data structures and it could be feasible that symbols are not used in human information processing. Arbib (1987) emphasizes that in biological systems input perceptions and output actions are closely related and this relation is not trough symbolic processing:

Because our interest in action and perception, I stress that an instantiation is not a program that is running off in some abstract symbol manipulation mode, but is a program within an organism or robot that is interacting with the world around it. The gathering of information through perception and the emission of behaviour through action must be integrated.

or what is the effect of physical environment to man; think weather forecasting for example. Also today, industrial products are more sophisticated and with higher quality than before. New aids for everyday life are emerging and machinery replaces humans in unconfortable positions. This development has been made possible via numerical computation in engineering problems. Computers that are designed specially for numerical problems are fast in calculations such as addition, substraction, multiplication and dividion with floating point numbers. The current state-of-the-art in parallel supercomputing is billion floating point operations per second.

Symbols and data structures are used for representations that are readable by humans and that many times are based on simplified priciples of the human language. Text processing, spreadsheet applications, data bases and bussiness graphics are good examples of such everyday symbolic processing. Symbolic computing with simple numeric operations makes it easier to do planing, reporting and controlling within the human society. The human information processing lacks exactness in memory, recognition and planing tasks which symbolic computating can replace. It obvious that todays commonwealth with it's complicated hierarcial structures would be impossible to govern without symbolic processing machinery such as the computer.

3. ARTICIAL INTELLIGENCE: PROMISES AND CRITICISM

The roots of Aritficial Intelligence (AI) are in 1960's when the symbolic computation language LISP was developed as a tool for man-machine connections (McCarthy 1960) . In the beginning of 1970's the progress in computer science and especially in symbolic processing made enthusiasts to belive that within twenty years machines will be capable of doing any work a man can do. Patric Winston (1977) described his goals of Artificial Intelligence more cautiously:

The central goals of artificial intelligence are to make computers more usefully and to understand the principles which make intelligence possible.

Biological systems (brains) are able to construct meaningful high level representations from physical sensory data. Even very simple animals can do operations that are beoynd understanding for computer experts; think about fly for example, the way how it is controlling and targeting it's operations via sensory information proves to be more intelligent than any modern computer system can do. It is evident that the first step in perception systems must be analog or numeric computing but when can we move to symbolic representations ? The first attempts of artificial intelligence, AI, were to convert physical signals into symbols as early as possible. In the scope of current knowledge this approach has widely been recognized as a failure.

In classical computer vision the problem solving is separated into low level image processing, dealing with physical data, and into high level processing, doing AI type of symbol processing for image understanding. These two ways, however, are quite different and many attempts to fill the gap by using intermediate processing have been proposed, but one of the proposed systems have been widely accepted. This paper discusses about a possibility to use numeric and subsymbolic computing in all levels of the computer vision problem, and use symbols only as an interface that is readable for humans.

2. NUMERIC VERSUS SYMBOLIC INFORMATION

The evolution of automatic data processing has been a development in both numeric computation and symbolic computation. For better understanding of the limitations of the current computer science of a look at numeric computing versus symbolic computing is taken.

Numeric or analogical information is physical data that may be received via sensory measurements or as a solution for the partial differential equations describing the physical laws of the nature. The requirement to process sensory data or to simulate numerical equations is evident in every branch of engineering. Numeric processing is a tool to find out how man-made objects behave under laws of physics

ARTIFICIAL INTELLIGENCE WITHOUT SYMBOLS

Pasi Koikkalainen
Lappeenranta University of Technology
Laboratory of Information Processing
P.O.Box 20,SF-53851 Lappeenranta, Finland

ABSTRACT

This paper presents a hypothesis that many tasks in artificial intelligence can be performed without symbolic processing. An example case is computer vision where traditional, symbolic, processing proves to be inefficient. The suggested solution is to use subsymbolic computation and massively parallel networks where in contrast to traditional computer vision systems physical signals are converted to symbols as late as possible.

1. INTRODUCTION

A major goal of computer science has been to introduce information processing capabilities that are comparable to those of humans. One of the most obvious of these capabilities is vision. Visual perception, the recognition of objects such as a table or a human face seems self-evident for us but proves to be difficult to program for a computer. The reason is that the human information processing differs from the information processing of computers. Computers are build for numeric and symbolic computing. Numeric information is close to the physical, analog world while symbolic information is often easier for man. When using computers one type of information is good for some tasks and the other for different ones but there is a set of open problems, like the vision, where neither seems to be suitable.