Evaluation of Image Processing Systems for Application in Research and Education

Axel Pinz, Walter G. Kropatsch, Alois Hinterleitner, Johann P. Petrak

Department for Pattern Recognition and Image Processing
Institute for Automation
Technical University of Vienna
Treitlstr. 3/1832, A-1040 Vienna, Austria

Abstract: *In this report we assess image processing hardware and software from a university institute's point of view. After giving a brief sketch of the overall situation and a review of related work on computer vision systems and image analysis benchmarks, we proceed with a description of typical image analysis tasks. There are basic-, mid- and high-level tasks in education, research and consulting. Three types of image processing systems are defined: general, hardware oriented, and software oriented. We conclude that "the appropriate system" does not exist, so that a heterogeneous network of several different image analysis components constitutes the best solution.*

1 Introduction

The framework for this report was set by the new department for Pattern Recognition and Image Processing (PRIP) at the Technical University of Vienna. This department has to be equipped with hardware and software for Digital Image Processing and several quite different requirements have to be satisfied by the equipment selected. We believe that our situation is a typical one for any university department involved in Digital Image Analysis, so that our evaluation criteria and consequences drawn may also concern a much larger interest group. The results of our investigations will demonstrate general conclusions about the actual market of such systems. Hence, our suggestions and selections may be useful also for others in the research community.

The initial situation is sketched in Figure 1, which contains three major components: the input to the system are the different image sources; the desired output constitutes a second part; and the engine that allows to process the input and to deliver the desired output under the given constraints is the unknown center part. Among the "external" sources, there is a scene which can be captured by a CCD-camera or by other sensors producing image data (e.g. radar, laser range finder), and there are several remote sensors that store the image data on intermediate media before they are read into the system. Such sensors are used

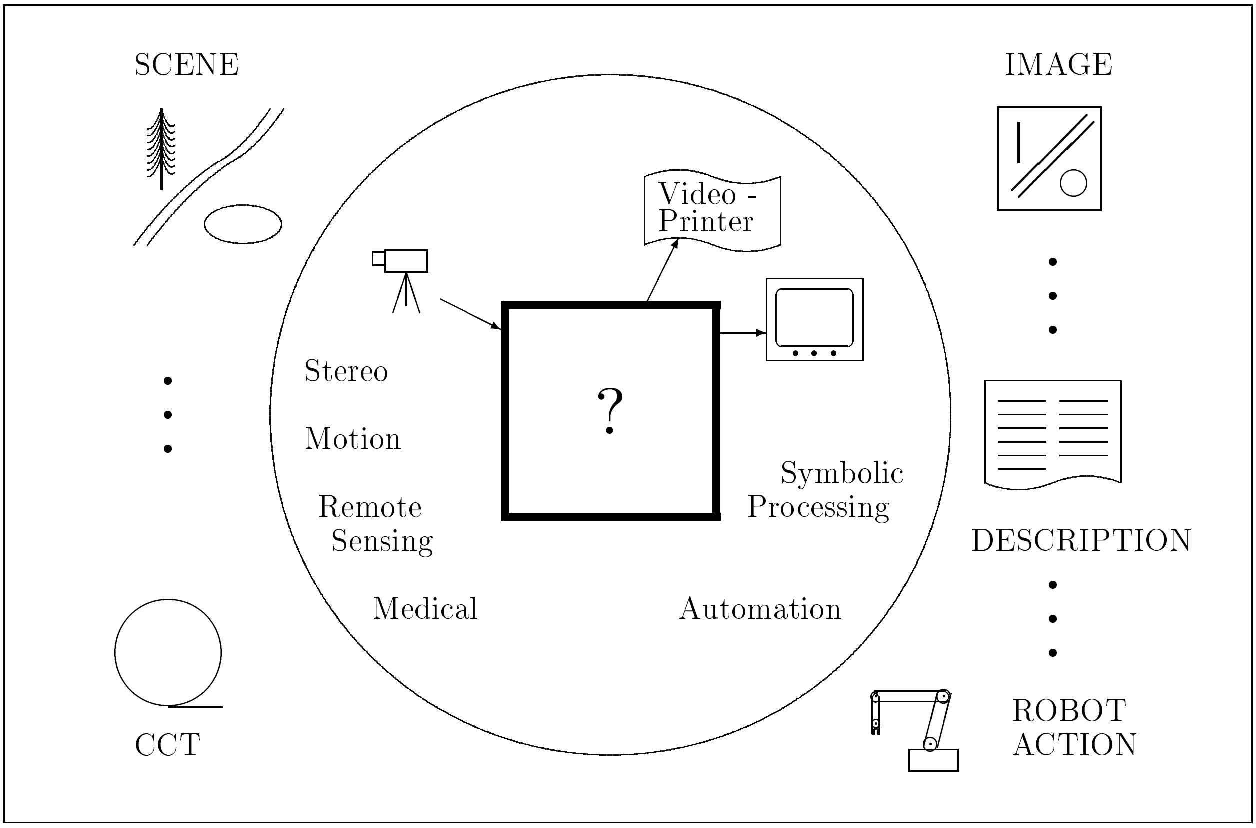

Figure 1: Sketch of the initial situation

in medicine (CT, PET, etc.) and in remote sensing (Landsat, SPOT, etc.). The format of the image data is highly application dependent. Some applications deal with multiple images: stereo requires two cameras taking two frames at the same moment; motion analysis processes need many consecutive frames; remotely sensed images consist of many spectral bands; and some medical datasets are three-dimensional.

A few examples illustrate the possible results of image analysis processes. Simple image processing tasks, like smoothing or registration, output a digital image; image understanding generates symbolic descriptions of the scene; and in industrial applications, the image analysis process may result in a robot action.

Starting from this situation, and having limited resources in terms of time and funds, we chose the following procedure:

- First, we defined our requirements more precisely, which lead to a definition of typical image analysis tasks.
- Then, we collected information about systems available in the market and
- finally, we evaluated them.

The above procedure is reflected in the structure of this article, which consists of three sections: *tasks, systems* and *results.*

1.1 How to acquire the necessary information ?

One source of information about image analysis systems are conferences, fairs, and advertising material. The large international conferences in our field are ICPR (International Conference on Pattern Recognition), ICCV (International Conference on Computer Vision), CVPR (Computer Vision and Pattern Recognition) and SCIA (Scandinavian Conference on Image Analysis). The largest conference of German speaking countries is the DAGM-Symposium. Finally, we wish to mention the annual conferences of the ÖAGM, the Austrian Working Group for Pattern Recognition. At these conferences there are industrial exhibitions of image analysis equipment. Two important fairs in Germany are the SYSTEMS in Munich and the IdentVision in Stuttgart.

Another source of information, which we believe to be most important, is personal contact with colleagues and manufacturers. We asked our colleagues for information about the kind of equipment they were using. Since we want to actively communicate with the international research community (which involves not only electronic mail and exchange of papers and data, but also the sharing of methods and code), a certain amount of standardization or, at least, compatibility is required. A future source of information will be the IAPR Technical Committee TC5 which will work on a survey of image processing and computer vision software on workstations [7]. Concerning the manufacturers, it turned out that it is absolutely necessary to see the system in action and to have a detailed demonstration and discussion with current users of the system. Since this is a rather time-consuming procedure, only a very limited number of systems can be inspected so closely.

Several authors describe specific software systems for image processing [17], [12], [8], and, in an early work, Hanson and Riseman compiled a complete book entitled “Computer Vision Systems” [3]. Winkler describes image analysis systems for industrial inspection tasks [18]. Jähne [5] discusses several PC add-on boards for image processing in the appendix of his book on digital image processing.

1.2 Benchmarks

While the design and application of benchmarks bears several difficulties as pointed out in [2], [1], and [11], they can give important clues. Conventional benchmarks [16], [1], [6] are tailored to measure either just one subsystem of a computer system (CPU, I/O-subsystem etc.), or they try to reflect the overall performance of the system by performing a typical application mix. There are two well-known image processing benchmarks of this kind: The Abingdon Cross and the DARPA Image Understanding Benchmark. The Abingdon Cross [9], [10], [13] was devised at the Abingdon Workshop in 1982 and - due to its easy implementation - has been run on a wide range of image processing systems. In [10], in particular, Preston reports interesting cost and performance comparisons for about 40 commercially available systems. The DARPA Image Understanding Benchmark [15] has been designed to cover a standardized set of “typical image processing tasks” and aims especially at parallel computer systems.

2 Typical Image Analysis tasks

We identify *three different levels of typical image analysis tasks*: basic-level, mid-level and high-level.

- At the *basic-level*, fundamental educational tasks are performed. The student learns about sampling and quantization to understand the nature of a digital image. These basic-level tasks include filter operations (smoothing, edge detection, convolution with user-defined kernels, Fourier-transform, etc.) and interactive manipulations (zoom, contrast enhancement). While a good performance is required for the transmission of images (fast distribution of an image to many workstations, printing of images), the student should experience the problem of dealing with a huge amount of pixel data. It is not necessary, therefore, to have high performance for the filter-operations, long execution times may give the student an intuitive understanding of the special requirements for image processing systems.
- *Mid-level* tasks involve the use of special purpose hardware for image I/O and for acceleration of processing (e.g. CCD-cameras, frame grabbers, pipeline processors and video printers). Sometimes there is a lack of sufficiently comfortable software to make use of these special hardware components. Mid-level tasks, therefore, include the programming of image processing systems to properly embed special purpose hardware. There are also mid-level educational tasks in programming like courses on “Programming of Image Processing Systems”.
- At the *high-level*, the tasks are more complex, so that we give a few examples only. High-level tasks are based on a large pool of known operations which can be executed effectively. The pool must be extendable by adding new operations and methods of any degree of complexity. Existing tools should be easily combinable (“macros”). In knowledge based image analysis we require tools for knowledge representation and the integration of a symbolic programming environment. We are quite sure that every researcher in the field will identify additional activities which he would also classify to be high level tasks.

Having sketched the three different levels of image analysis tasks, we estimate the means required. The diagram in Figure 2 shows an estimation of system prices in relation to the task level. Prices are given in Austrian Schillings (ATS, 1 US\$ $\simeq$ 12 ATS) and we used a logarithmic scale. While it is possible to get basic-level performance for far less than ATS 100,000.- (there are pure software products for standard PCs and workstations), mid-level tasks require special purpose hardware. They start around ATS 150,000.- and range up to ATS 700,000.-. A dedicated image processing system, capable of many of the above mentioned high-level tasks, will start well above ATS 1,000,000.-. Note the interesting gap between mid- and high-level tasks ($\simeq$ ATS 700,000.- to $\simeq$ ATS 1,200,000.-, see Figure 2).

Another possibility divides image analysis into *three types of tasks*: education, research and consulting. Figure 3 shows the correspondence between these task types and the task levels defined previously. In fundamental educational tasks, a basic task level will normally be

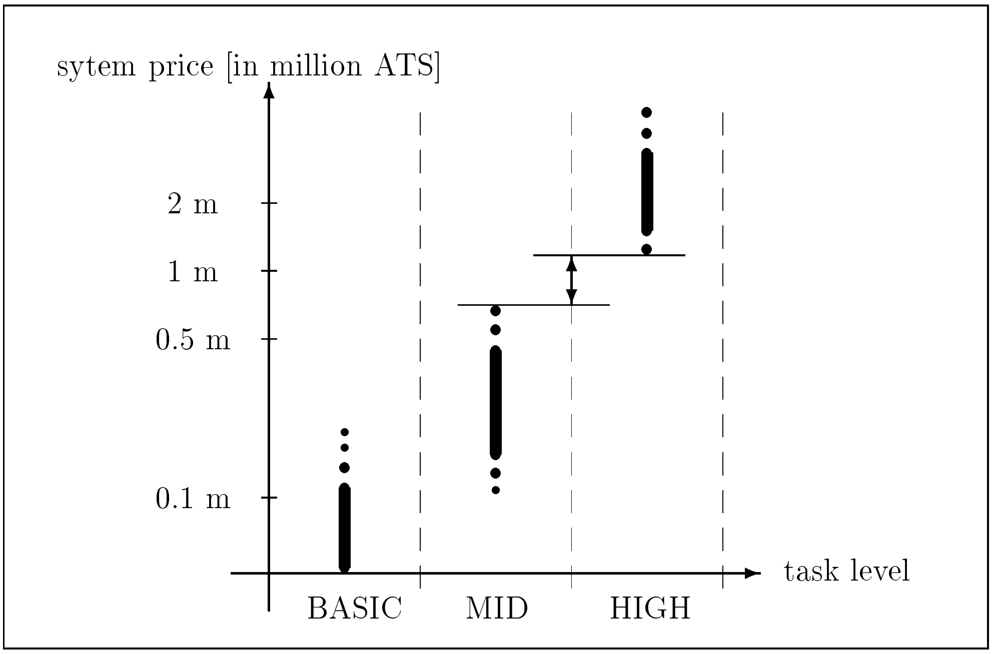

Figure 2: Estimated system prices

sufficient (full bar). Advanced educational tasks like a master's or PhD thesis, are done on a mid- or high-level. Fundamental theoretical research might be carried out without any image processing system (the fat dots indicate that a high-level is more likely than a mid-level - fine dots). Applied research can take place at any task level, while consulting will be done at mid- and high-levels.

3 Types of Image Processing Systems

Having talked about possible ways of viewing image analysis tasks, we are trying out a similar approach for the types of image processing systems and divide them into the following three categories:

- General: This is either a PC with image processing software or a general purpose workstation with "simple" software (there are several packages which aim primarily at visualization, but also contain an image processing module).
- Hardware oriented: At the low end there are PCs with add-on boards (frame grabber, image memory, processors, etc.), then there are many VME-bus-based systems of rather different architecture. VME-add-on boards are still available, but most of the more recent workstations do not have a VME-bus any more. There are complete VME-systems with a single board computer (e.g. VME-PC-386, Sparcengine, etc.), but the most flexible solution seems to be a detached special purpose hardware unit, which is coupled to a host computer (typically a workstation).

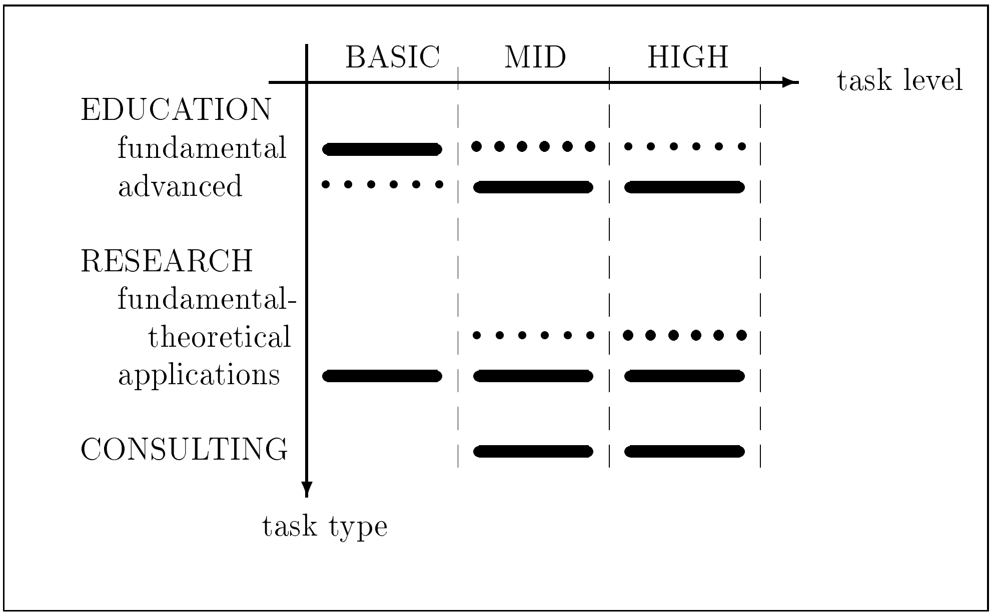

Figure 3: Task type versus task level

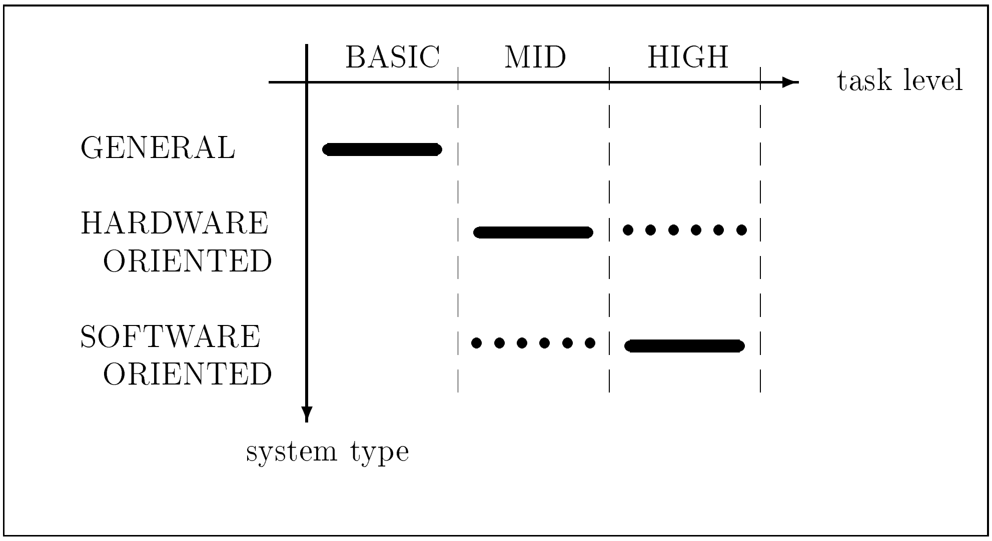

Figure 4: System type versus task level

- Software oriented: In addition to the “simple” image processing software mentioned above, there are several sophisticated software packages for image processing available. They usually run under UNIX and make use of the capabilities of “graphic” workstations (we will discuss the differences between graphics and image processing and the resulting problems later). This type of system has major advantages: it is open and expandable, standardization is affordable, high-level interfaces are available (knowledge-based image processing).

Figure 4 shows our view of a relationship between task levels and system types. While there is an obvious relation between the general system type and the basic-level, one might argue about the rest of the diagram. Today, most of the systems which provide high-level capabilities are hardware oriented (fat dots), but the emerging software oriented systems certainly aim at this high-level (full bar), so that Figure 4 may be showing a situation in the near future.

4 Results

The diagrams in Figure 3 and 4 illustrate that "the appropriate system" does not exist. As a consequence, we propose a *network consisting of several different image analysis components* and view the complete network as the system (Figure 5). In such an environment we find application- or task-specific components, while at the same time there is the possibility of sharing rare resources (e.g. one camera and frame grabber for ten basic educational systems).

Besides the obvious advantages, there are several problems with this heterogeneous solution. There will be a heavy traffic of image data on the network. Neither a standard ethernet nor the use of fibre optical cable (FDDI) can provide sufficient throughput, so that very effective *image data compression* is required. When the components are used remotely (i.e. controlled by another workstation in the network), problems are bound to arise, especially with hardware oriented systems and in cases where interactive facilities are required (e.g. a high-resolution display connected to special purpose hardware or interactive coordinate entry). Finally, system compatibility, handling and maintenance, will be difficult because the system is heterogeneous. This could be overcome by a standardization of image analysis systems, a current research effort which will certainly continue [4].

While the previous sections of this article stayed at a very high level of abstraction, we want to close with a discussion of a few points in more detail. These "hot points of argument" might provide helpful clues to others involved in a similar evaluation process.

- General purpose computer versus special purpose image processing hardware: In affordable cases we would prefer the general purpose computer. By "affordable" we mean: sufficient performance for the required task, not too expensive, sufficient resolution, interactivity, and the like. However, in certain cases special purpose hardware will remain the best solution (e.g. real time, research in parallel processing). When talking about parallelism, it should be *massive* [14], we therefore do not favor transputer systems with a maximum number of several tens of transputers.

- The difference between "Computer Graphics" and "Image Processing": Today's graphic workstations do not reflect the needs of typical image processing environments. To mention a few obvious points: We need at least 24 bits/pixel in *independent* channels (e.g. 3x8 bits), *plus* independent overlays when working with true color (e.g. RGB) images. Independent zoom and pan of the channels should be possible. We need several independent look-up-tables (one LUT/window), better coordinate input facilities

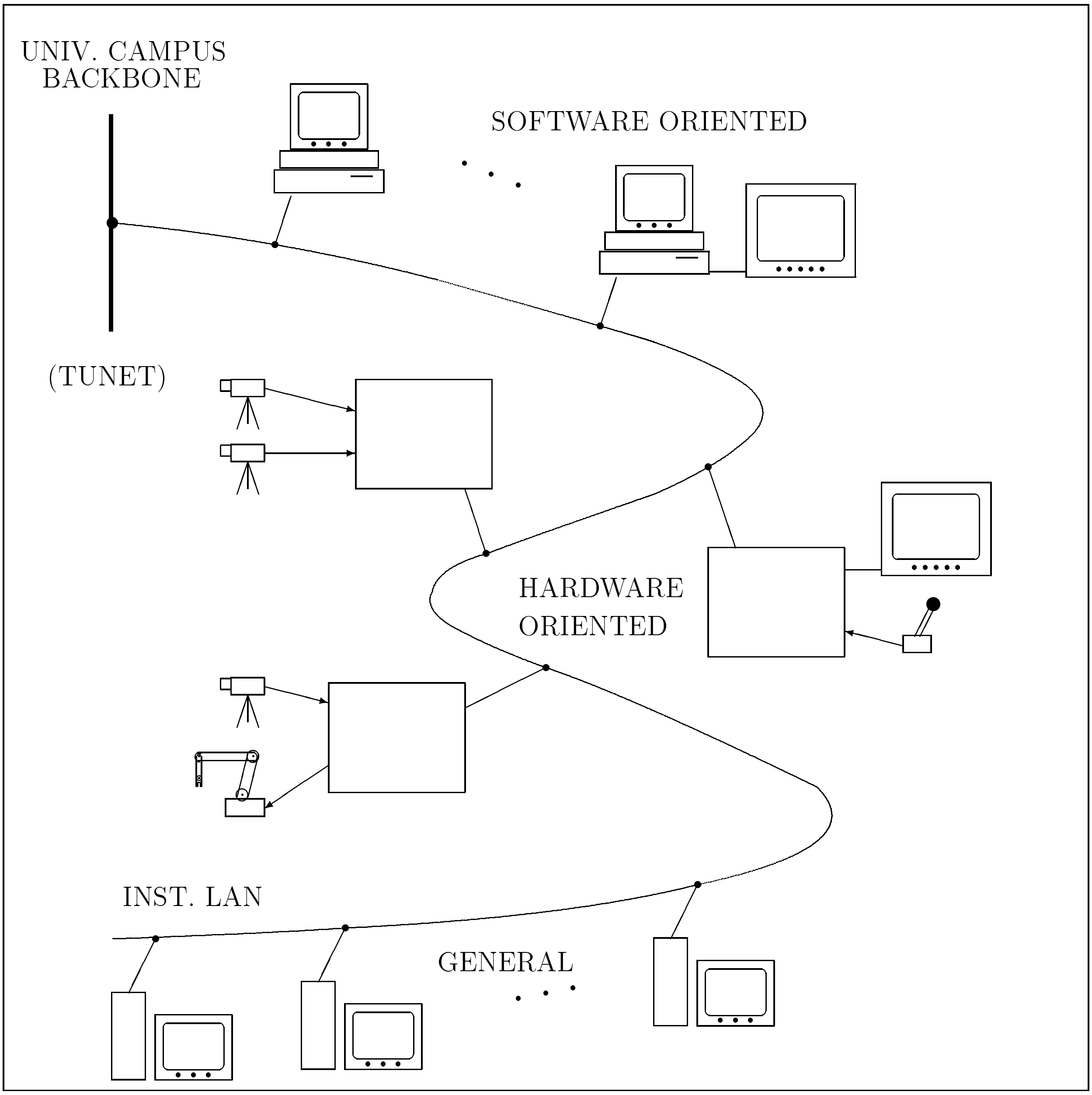

Figure 5: The network is the system

(a mouse is frequently not sufficient), and a lot of RAM. We *do not need* a very large number of vectors/second, a feature which makes these machines extremely expensive.

- Image analysis benchmarks: would help very much in the selection of the appropriate system. Unfortunately, they are not yet performed for many current systems and differentiate only between very rough characteristics of the systems. Any definite decision must be based on the evaluation of the specific needs of the intended range of applications. However, the results published may help in finding those companies which have the necessary knowhow in this area.

- Product lifetime: Compared to computer graphics, the market for image processing is still rather small. Therefore a graphic workstation will become obsolete much faster than is the case for image processing components, so that the technology gap is growing. As has been mentioned above, that most of the image processing hardware still uses the VME-bus, which is not favored by workstation manufacturers any more.
- Standardization: is urgently required, but will be extremely difficult.
- Input/output and communication of images: It has already been stated, that very efficient compression methods are required. This might involve “fuzzy” compression, where the original and the restored image may differ “slightly”.

Acknowledgement

The authors are very grateful to Michèle Oliva for the invaluable help in the preparation of the English text.

References

[1] J. Dongarra, J. L. Martin, and J. Worlton. Computer benchmarking: Paths and pitfalls. *IEEE Spectrum*, 24(7):38–43, July 1987.

[2] M. Duff. How not to benchmark image processors. In L. Uhr, S. Levialdi, K. Preston, Jr., and M. Duff, editors, *Evaluation of Multicomputers for Image Processing*, pages 3–12. Academic Press Inc., 1986.

[3] A. R. Hanson and E. M. Riseman. *Computer Vision Systems.* Academic Press, 1978.

[4] ISO/IEC JTC1/SC24/WG1. *Imaging: Image Processing and Interchange (ITI)*, 07 1990.

[5] B. Jähne. *Digitale Bildverarbeitung.* Springer, 1989.

[6] G. L. Marchant, M. B. Stephenson, and T. Crowfoot. A set of benchmarks for evaluating engineering workstations. *IEEE Computer Graphics and Applications*, 9(3):29–33, May 1989.

[7] T. Matsuyama. Tc5 activities in 1989-1990. *IAPR Newsletter*, 13(4):6, Feb. 1991.

[8] A. Pinz. Digitale Bildverarbeitung am Inst. f. Vermessungswesen u. Fernerkundung: das System DBIVF. In F. Pichler and A. Pinz, editors, *Statistik und Mustererkennung*, OCG Schriftenreihe 42, pages 111–124. Oldenbourg, 1988.

[9] K. Preston, Jr. Benchmark results - the Abingdon Cross. In *Evaluation of Multicomputers for Image Processing*, pages 23–54. Academic Press, 1986.

[10] K. Preston, Jr. The Abingdon Cross benchmark survey. *IEEE Computer*, 22(7):9–18, July 1989.

[11] W. J. Price. A benchmark tutorial. *IEEE Micro*, 9(5):28–43, Oct. 1988.

[12] H. Ranzinger. DIBAG - ein portables Bildverarbeitungssystem. In F. Leberl and H. Ranzinger, editors, *Mustererkennung und Bildverarbeitung in Österreich, ÖAGM-Tagung Ramsau*, pages 95–112. Oldenbourg, 1981.

[13] W. B. Teeuw and R. P. W. Duin. An algorithm for benchmarking an SIMD pyramid with the Abingdon Cross. *Pattern Recognition Letters*, 11(7):501–506, July 1990.

[14] J. K. Tsotsos. Analyzing vision at the complexity level. *Behavioral and brain sciences*, 13:423–469, 1990.

[15] C. Weems, E. Riseman, A. Hanson, and A. Rosenfeld. The DARPA image understanding benchmark for parallel computers. *Journal of Parallel and Distributed Computing*, 11:1–24, 1991.

[16] R. P. Weicker. An overview of common benchmarks. *IEEE Computer*, 23(12):65–75, Dec. 1988.

[17] J. Wiesel. Didak, a digital image processing system. In *Proc. Intern. Symp. on Image Processing, Graz*, 1977.

[18] G. Winkler. Zur Beurteilung der Eignung von Bildauswertesystemen für die Lösung von Sichtaufgaben. In H. Ernst, editor, *Maschinelles Sehen*. Europa Fachpresse Verlag, München, 1990.