SOME RESEACHES OF PYRAMID STRUCTURE IN ROBOT VISION SYSTEM

*Dapeng Zhang, **Zhaoqi Bian

* Institute of Automation,Academic Sinica,Beijing,PRC
** Dept. of Automation,Tsinghua University,Beijing,PRC

ABSTRACT

In this paper, we investigate some essential characteristics of a robot vision system (RVS), including hierarchy processing, multi-knowledge, real-time rate and mixed control strategy, and explore the technical approaches of integrated design which should be adopted. Then an expanded pyramid structure suitable for the RVS is discussed, mainly composed of hierarchic array, distributed knowledge base, multiple controller and MIMD network. Finally, the implementation mechnism of the system is described.

INTRODUCTION

A robot vision system (RVS) is defined as a system whose scene informations are obtained, processed, classified and described, and scene knowledges are represented, stored and utilized effectively.(1) Generally speaking, the system is mainly composed of four processing modules, including iconic processing(IP), transforming interface(TI), symble processing(SP) and knowledge base(KB). They are organized as shown in Fig.1. It is obvious that the RVS should be able to combine "symbolic,semantic" and "iconic,pictorial" information, using each where appropriate, and to process image fast enough to handle scenes input in real-time from the the real world. This means that we have to develop special computing system. In other words, an appropriate structure must be designed in an extremely parallel fashion.

One of the attractive structures is pyramid. A number of researchers have been exploring this structure due to its many good qualities, such as handling global communications and global processes with great efficiency, reducing message-passing, computing complex function hierarchically, and achieving massive speedups.(2) Especially, it is suitable for image processing, pattern recognition and some computer vision where information is simultaneously transformed and anverged (or, moving in the other direction, broadcost and diverged). However, to implement robot vision processing, some special requirements must be considerated and some integrated technologies must be adopted in the design of this structure.

In this paper, we attempt to explore some relative problems about the RVS and present an expanded pyramid structure suitable for the system. For simplicity, we assume that the raw image is single digitized optical image and the goal of the system is to do object recognition.

SOME DESIGN CONSIDERATIONS

To design an apprapriate RVS structure, it is necessary to investigate the essential characteristics of the system and to explore the technical approaches that should be adopted.(1) Some main characteristics of the RVS are considerated as follows:

.Hierarchy Processing

The RVS main processing modules mentioned previously form a distinct hierarchy, i.e. their IP, TI and SP corresponding to low level, intermediate level and high level processing respectively. According to this feature, the overall system organization designed should be a hierarchic global structure.

.Multi-Knowledge

The system incorporates a great quantity of knowledge about the class of scenes to be analyzed and types of objects that occur in them,as well as the rules governing how this knowledge is used in analizing the image. The knowledge is generally divided into three different types, involving the physical knowledge, the perceptual knowledge and the semantic knowledge. As a result the knowledge base designed should be broken down into relatively local sub-knowledge bases.

.Real-Time Rate

This system can recognize, describe and track complex objects in real-time in the scene. It will be put into practical use only when the system must be designed to perceive fast enough to meet real world real-time constraints. This kind of real-time demands make it absolutely necessary to use a predominantly parallel multi-computer approach.

.Mixed Control Stratery

The RVS should be a flexible,robust and extendable system, where a mixed top-down, bottom-up/feedback control strategy must be adopted. The knowledge base is explored in a top-down mode. While the analyzing procedure itself is applied in a bottom-up mode and a feedback mode. Obviously, the structure designed should meet this strategy with multiple controller.

STRUCTURE DESCRIPTION

Based on the considerations above, the RVS structure should be integratively designed. Here an expanded pyramid structure for the RVS is developed. It is mainly composed of hierarchical array, distributed KB, multiple controller and MIMD network. The block diagram of the structure is shown in Fig.2.

.Hierarchical Array

It can be divided into three different array hierarchies of processing level, i.e. IP array (IPA), TI array (TIA), and SP array (SPA), which carries out low level, intermediate level and high level processing, respectively. In fact, each processing level array also consists of many layer arrays, while every layer array can perform a fixed operation that belongs to the corresponding processing level.

.Distributed KB

It is configured by some individual KB, e.g. KB for low level processing (LKB), for intermediate level (IKB) and for high level (HKB), as well as their coordination mechanism. Each KB possesses its own knowledge processor and searching device. If the knowledge is required in the processing, a corresponding KB is enabled via the relative searching device and the knowledge is derived from the knowledge processor. Obviously, various kind of knowledge are interacted on each other by the KB coordinator.

.Multiple Controller

In reality, a tree structure of controllers is formed in the RVS, in turn including host computer, level controller, layer controller, region controller,or even processor controller. The host is used to control and coordinate all processing modules in the system. The level controller governs the information processing in the level and data communication between the levels, while the layer controller manages its array to perform an assigned operation. In some conditions, a region controller in the layer is required in order to carry out a special region processing. In addition, each processor of layer array can have its own controller. This kind of tree form control strategy makes the system even more flexible.

.MIMD Network

It plays the role of information transmission. Information can be frequently passed between different layers and different levels of the system via this network. At higher layers, the network may most appropriately be effected. It is also used for feedback to, and interaction with,low level processing at lower layers.

IMPLEMENTATION MECHANISM

In order to further comprehend the structure of the system, a simple implementation outline is given out. Here we assume that the raw image has been transduced and input to the lowest layer array of the system,called retina.

IPA is mainly used to perform the operations on pixels and local neibourhoods of pixels,such as smoothing, sharping,averaging,detecting gradients, segmenting and extracting features. These operations are successively implemented by their corresponding layer arrays with SIMD mode and bit plane processing. The sizes of the layer arrays are defined in terms of different processing requirements. In addition, the knowledge required is derived from LKB for refining the processing result in the case when feedback from high level to low level occurs.

TIA provides an interface between pixel-based representation and symbolic elements representing visual knowledge stored in IKB. It will involve grouping, splitting and labeling processing. TIA is constructed by multi-SIMD layer arrays to perform associative processing. Each cell has a multi-bit (e.g. 16 bit) parallel ALU. The results of TIA processing are representations of image entities such as regions, line segments and vertices.

SPA performs symbolic interpretation using the relative object descriptions stored in HKB. It will do the operations involved in recognition of objects, classification and handle top-down processing, e.g. telling IPA what kinds of things to look for. This array is driven by MIMD operation mode. Its processor has more powerful, e.g. LISP processor, and execute full word (64 bit,or more) operation.

In this way, the RVS can successively transform the raw image's information into more abstract representations, and do the corresponding processing until the object is recognized. Since the information is passed through the structure in pipeline/ parallel, the system can successively process real world images in realtime.

CONCLUSION

In this paper, an integrated hierarchical structure--expanded pyramid for RVS is developed. It has an interesting mixture of advantages, including good integrated technical coordination, good message passing capabilities and good global/local control strategies. It is designed to use for the RVS, but it is also a powerful system for a number of other areas in AI.

REFERENCES

1. A.Rosenfeld, (ed.), "Expert" Vision Systems: Some Issues.Computer Vision, Graphics, and Image Processing 34, 99-117, 1986.
2. L.Uhr, Multi-Computer Architectures for Artificial Intelligence, John Wiley & Sons, 1987.
3. S.Levialdi(ed.), Integrated Technology for Parallel Image Processing, Academic Press,Inc.1985.

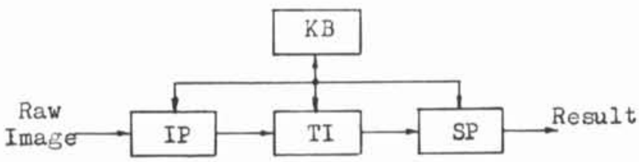

Fig.1. Main Processing Modules of a RVS

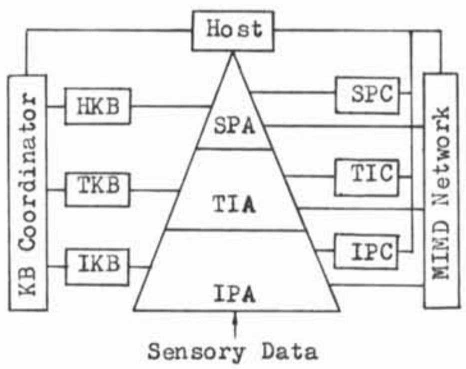

Fig.2. Block Diagram of the RVS Structure