Architecture of Computer Vision Software System

A.Ivkin, A.Doudkin, R.Sadykhov

Institute of Engineering Cybernetics of the National Academy of Sciences of Belarus
6 Surganov Str., Minsk, 220012, Belarus
sadyhov@newman.basnet.minsk.by, doudkin@newman.bas-net.by

Abstract

The software of the computer vision systems is complicated. The authors discuss the principles of computer vision systems organization and propose the architecture for such software system.

Introduction

This article describes the architecture for high performance software system that supports distributed computing and parallel processing. Vision is one of the most difficult abilities to be implemented artificially. Developments in this area often require knowledge from a vast range of areas, including artificial intelligence, mathematics, physics, computer architectures and neurosciences.

Computer vision rose unique challenges in computer software and hardware technologies. The challenges include making computer vision systems less expensive, more compact and more robust in various conditions. These systems often have critical missions and require real-time processing of data to meet millisecond-level control cycle requirements. A comprehensive approach to developing, operating and maintaining a computer vision system should cover all these critical issues[1]:

- a scalable, open architecture;
- a comprehensive, computer-aided rapid development and deployment environment;
- real-time control and operator-interaction capability.

The ultimate goal is to propose a unified architecture solution for the computer vision software. Scraping existing systems and building new architecture from scratch is not practical. So, computer vision technologies must evolute incrementally. First, let's look at fundamental computer vision technologies.

1. Function chart of the computer vision system

The common function chart of the computer vision system is given in fig. 1. First, the image of an object is transferred to the light – signal converter through the optical device, then the electrical signal is amplified and is stored in the pre-processing device. The image analysis (secondary processing) device carries out location and recognition of the object, determination of its coordinates and orientation. The information on the object is displayed by the device of the visual control in case of need. On the basis of the received information the communication controller chooses control signals, that bring in action executive mechanisms, carrying out influence on the object. Besides the system of computer vision can carry out record of the results of the analysis of the image. E.g., it can generate the description of the electronic circuit on the basis of the processed image of the VLSI chip.

The purpose of the primary processing device is to reduce the common time of the image processing. In the time of pre-processing representation of the image is done as a set of characteristic points (the centre of the object, its important characteristics: corners, distances from centre up to edge, etc.) Thus, at the first stage direct processing and representation of the information are done in form that is convenient for further standard transformations which usually are carried out on the universal computer.

The secondary processing device carries out recognition of the image, i.e. it uses its characteristic attributes for determination, to which class the object belongs. On the input of the device there is a code of the image, obtained at the stage of the primary processing. On the output there are results of the processing of the image as the description in a language of the subject area. In the case, when the system is intended for automated control, on the output we have control orders, translated by the controller of communication in signals, determining movings and actions of the manipulator.

The important part of system is the control device. Its functions are control of the processing devices parameters and synchronization of the processes, carried out in the system.

2. Further development of the computer vision system architecture

For further specification of the classes structure of the computer vision system we list some basic methods of image processing and recognition used in similar systems [2] (fig. 2).

It is necessary to mean, that the process of image input and generation and processing is affected by noise. A goal of pre-processing is elimination of noise distortions, brought in at different stages of formation, i.e. restoration of true values of the brightness function [3].

This purpose is pursued by image filtering. We shall consider the basic filtering methods:

1. threshold;
2. anisotropic and reccurent;
3. discrete Fourier transform, and also Hadamard and Walsh transforms that are less exact, but faster.

At the following stage of processing the image segmentation is carried out [4]. The segmentation is a separation of the image into components: objects, their fragments and characteristics. Two approaches to the decision of a problem of segmentation are known.

The first one is based on contour location on the images, that is assumed as a set of borders between object and background, between various objects or between adjacent areas of the same object. Three basic methods of contour location of the image are known:

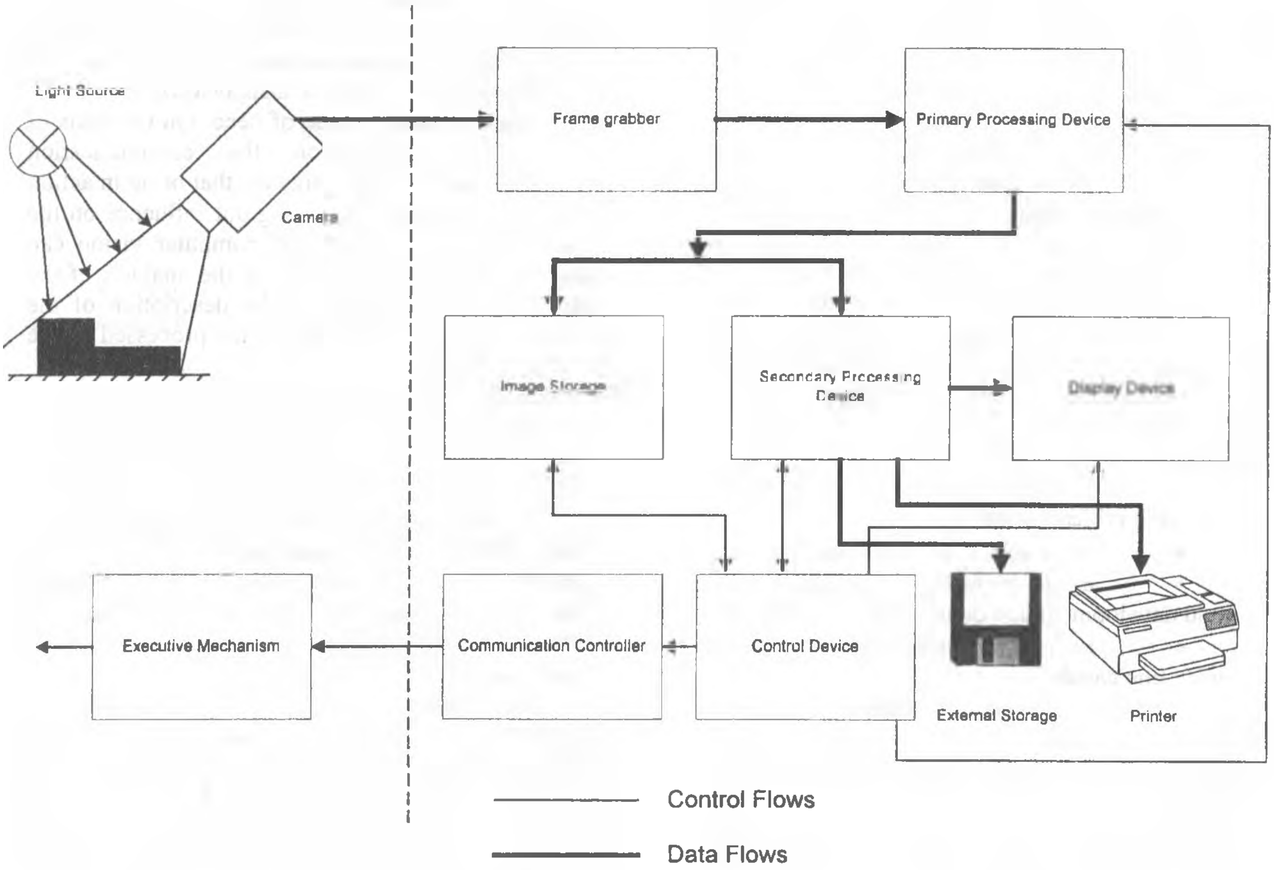

Fig.1. Function Chart of the Computer Vision System

1. method of spatial differentiation;
2. method of functional aproximation;
3. method of high-frequency filtering.

The specified methods of contour location of the image are realized in the computer differently. The realization of a method of spatial filtering brings

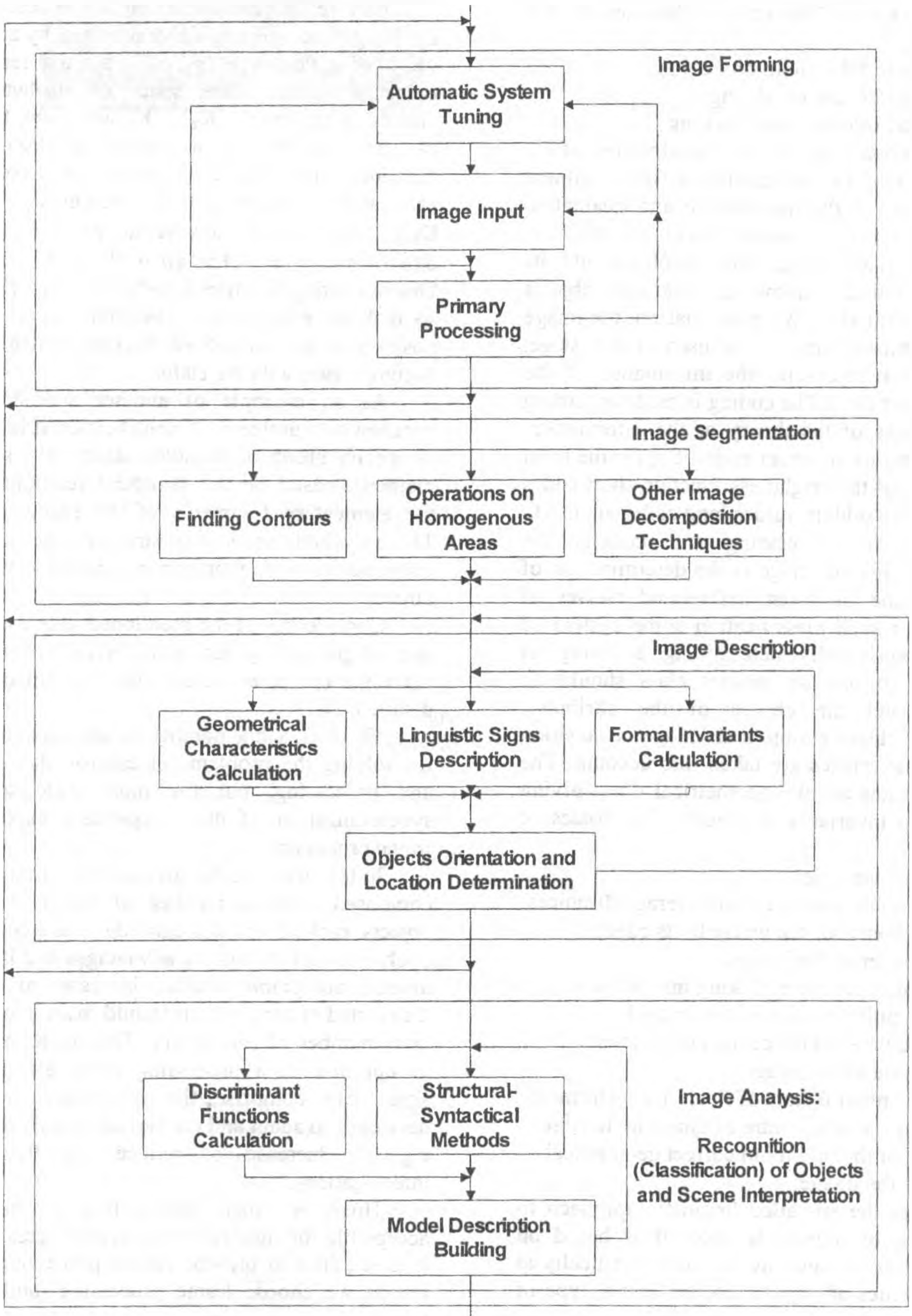

Fig.2. Structural Scheme of Image Processing in the Computer Vision System

in large volume of calculations. The method of spatial differentiation is most frequently applied in practice.

The second approach to the decision of a problem of segmentation is based on the concept of uniformity of the image points, laying in some area, limited by a contour. The basic techniques of this approach are:

1. method of threshold processing;
2. method of area escalating;
3. method of relaxation marking.

The next processing stage is construction of the image description, i.e. representation of brightness function as a set of the quantitative and qualitative characteristics, forming a set of attributes, which is hereinafter used for recognition of object and its classification, and also allows to determine objects location and orientation. We note, that not the image itself, i.e. brightness function, is used at this stage, but its code, that represents the information of the image in a proper kind. The coding is made according to the basic rules of the theory of the information, related to finding of optimum codes. So, for the most frequent values of the brightness function short codes are given, and for seldom values long codes are used.

After primary processing and coding, the analysis of the filtered image is the determination of its fitting to one of given beforehand classes of objects [5]. For such classification some system of attributes, unequivocally determining a fitting of visible object to one or another class should be allocated. During the choice of the attributes, algorithms of classification and image acquisition hardware characteristics are taken into account. The simpliest attributes are the geometrical ones, giving some universal invariants of objects. The basics of them are:

1. area of the objects;
2. maximum, minimum and average distances from centre of inertia of the image to its edge;
3. perimeter of the image;
4. adjusted perimeter (taking into account weight of each point on edge of the image);
5. coordinates of the centres of various moment of inertia of the image;
6. some initial factors of Fourier transform of the distance from inertia centre of image up to edge;
7. sizes of the minimum correct geometrical shape, limiting the image.

Sometimes the so-called linguistic approach to the description of objects is used. It is based on attributes, that are located by human heuristically at study of properties of objects (corner, arrow, type of crossing and other).

The listed geometrical attributes in the majority of cases do not provide the complete description of visible objects, sufficient for their classification, and frequently it is necessary to use more complex formal invariants, for example, moment functions of the image, and other.

As a result of construction of the description of the image each object is characterized by an ordered vector of attributes $p = (p_1, p_2, \dots, p_n)$, that represents a point in n-dimensional space of attributes. If N classes of objects $K_1, K_2, \dots, K_N$ are given, the object can be classified by a method of discriminating functions, that in a general case consists in calculation of values of N of functions $d_1(p)$, $d_2(p)$,... $D_N(p)$, such that for any vector $p^* \in K_i$, $d_i(p^*) > d_j(p)$ while $i \neq j$ for all $j = 1, 2, \dots, N$. Discriminating function $d_i(p^*)$ $(i=1, 2, \dots, N)$ gets out so that the error of classification was as small as possible. A special case of discriminationg functions is comparison with the etalon.

As an example of another approach to the decision of a problem of identification it is necessary to specify group of so-called structural - syntactical methods, based on the structural relations between the elementary fragments of the image of object. These methods use known structures of the discretic mathematics: formal grammars, graphs, networks and other.

Considering of the mentioned above analysis in area of processing and analysis of the images, we make the computer vision system architecture more detailed.

There is a big number of architectural models for solving the problems of control, data collection and processing, but the most wide spread are synchronization of the independent executors and frame processing.

In the first case the architecture of the system is composed from a number of rather independent objects, each of which is carried out as a control flow. Such architecture has the advantages and is, perhaps, unique acceptable model in case of designing distributed system, which should make processing of large number of parameters. This model also allows to optimize data processing more effective (each agent can comprise the information how it is necessary to adapt and change the system parameters, e.g., to increase or reduce the frequency of interrogation).

However, such architecture is not always acceptable for rigid real-time systems creation, when it is required to provide robust processing. For this reason we choose frame processing model for our system.

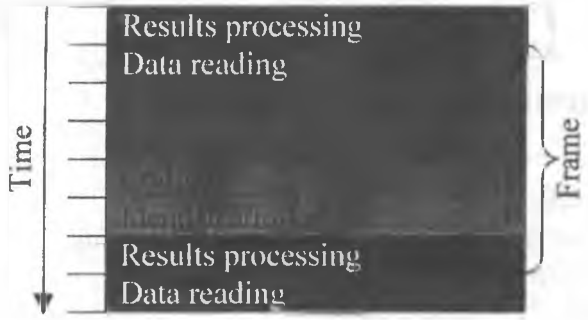

Fig.3. Frame Processing

As it is shown in fig. 3, the process of functioning comes true in this case as a sequence of reading of the image, filtering, segmentation, coding, identification and processing of results through certain intervals of time. Each element of such sequence is a frame. In one's turn, it is possible to divide the frame on a number of sub-frames, taking into account their functional behaviour. The basic advantage of such model is, that we can more rigid supervise a sequence of actions of data collection and processing system.

5. Conclusion

We have proposed architecture of the computer vision software system on the basis of analysis of most common processing methods. This architecture is scalable and flexible in the terms of possibility to use different algorithms of recognition and primary processing of images. Another feature of the offered approach is the multipurpose orientation of the developed architecture that distinct this system from similar ones [6,7].

Our system provides the facility necassary for experimenting with various choices at different abstraction levels to find the best match for the system under development.

6. References

1. G.Booch. Object-Oriented Analysis and Design. Addison-Wesley, New York, 1994.
2. R. Lewis. Practical Digital Image Processing, Ellis Horwood Ltd., Chichester, UK, 1990.
3. U. Seger, H.-G. Graf, and M.E. Landgraf. Vision Assistance in Scenes with Extreme Contrast, *IEEE Micro*, Vol. 13, No. 1, Feb. 1993
4. L.S. Davis. A Survey of Edge Detection Techniques, Computer Graphic and Image Processing, Vol. 4, 1975.
5. R. Duda and P. Hart. Pattern Classification and Scene Analysis, John Wiley & Sons, New York, 1973.
6. J.M. Sanchiz, F. Pla, and J.A. Marchant. A Framework for Feature-Base Motionrecovery in Ground Plane Vehicle Navigation, Proc. Computer Analysis of Images and Patterns '97, Springer-Verlag, New York, 1997
7. VISION Computer Vision System. http://www.llnl.gov/eng/ee/erd/siprg/vision.html.

The research was supported by the INTAS (project № 97-2028).