MULTICOMPUTER SYSTEM DEDAL-2 FOR LOCAL LANDSCAPE MONITORING

I.V.Gribkov, P.P.Koltsov, N.V.Kotovich, A.A.Kravchenko, A.S.Kutsaev, V.K.Nikolaev, A.V.Zakharov
Scientific-Research Institute for System Studies, Russian Academy of Sciences (NIISI RAN),
Nakhimovskii pr.36-1, Moscow, 117218, Russia
e-mail: koltsov@niisi.msk.ru, alex@genebee.msk.su

ABSTRACT

The paper contains the description of a multicomputer system for terrestrial surface automatized monitoring. The system is able to detect moving objects on the terrain considered, to locate static objects which appear/disappear on terrain as well as to recognize such objects. The system uses photosnapshot the terrain and a set of reference images for recognition.

Keywords: Image processing, Pattern recognition, Motion analysis, Contour extracting

Image processing, recognition and understanding can be treated as important and complicated problem in computer science.This circumstance is connected in first with essentially intellectual character of this problem. Secondly, with urgent need in image processing automation in a lot of practical applications. Modern computer facilities have become a powerful tool for creation of effective technical systems. Simultaneously, it gave a pulse to researches along entire image processing chain - from input to preprocessing, actually intellectual processing and postprocessing of the videographic information.

In the paper, we consider the system DEDAL-2 for the automated processing of videosequences, obtained by a videocamera (VC). The basic application of the system is the visual monitoring of a terrestrial surface, thus the VC is onboard the aircraft. The system detects (recognizes) terrestrial surface changes, respective to the images of the same sites, received earlier.

DEDAL-2 includes elements of preprocessing, i.e. refinement and normalization of conditions of videoshooting. It also includes elements of intellectual image processing: recognition of changes and motion detecting. Finally, the system implements postprocessing: it displays the results of the work. Hence it is possible to say that the system executes the analysis of the changes detected and, by results of the analysis, visualizes its own "understanding" of the changes.

The monitoring problem splits into several tasks. It is required to determine qualitative character of changes, coordinates of appearing/disappearing objects on terrain, to estimate their sizes, to determine their shape, and also to identify changes. Appearance or disappearance on terrain of motionless objects we call static changes. Movement of objects during videoshooting, that is repositioning of object from the frame to frame, we call dynamic changes. It is required to determine both static and dynamic changes and to identify (to recognize) the detected objects.

DEDAL-2 uses orthophotoplan of a traced segment of terrain. The typical size of such segment is 100 x 100 kilometers. Orthophotoplan is a photosnapshot of a site of a terrestrial surface, provided that the camera lens is directed strictly perpendicularly to ground. Thus orthophotoplan contains orthogonal 2D-projection of a terrestrial surface.

Orthophotoplan beforehand is loaded into the system. While processing frames of a videosequence, the system carries out their binding to orthophotoplan. This allows to precisely determine coordinates of objects on terrain. The system allocates and identifies static changes on terrain, as distinctions between the current videoframes, and appropriate to them, segments of opthophotoplan. Such an allocation and identification does not depend on conditions of videoshooting. The system allocates dynamic changes, finding mobile objects in a videosequence. Using database of the reference images, it identifies (recognizes) the objects.

The total amount of calculations is significant. Thus it is necessary to use multicomputer system with interacting computing units. Each unit processes some data, transfers outcomes to the following unit and receives the data from the previous one.

DEDAL-2 is a distributed computing environment, consisting of six personal computers, integrated in five computing units.

At black-and-white graphic pallete, the volume of a 256 x 256 pixels videoframe makes 64 Kilobytes. In real-time mode it is required to transmit about 25 such videoframes per second. So the total network traffic on some computing units theoretically can exceed 3 Mbytes / sec. For maintenance of such a traffic the network includes a 100 Mbit switchboard. The network interaction on units is endured in separate flows. All computing units operate asynchronously. They get the necessary information on their demand.

The computing system includes the following units (Fig.1):

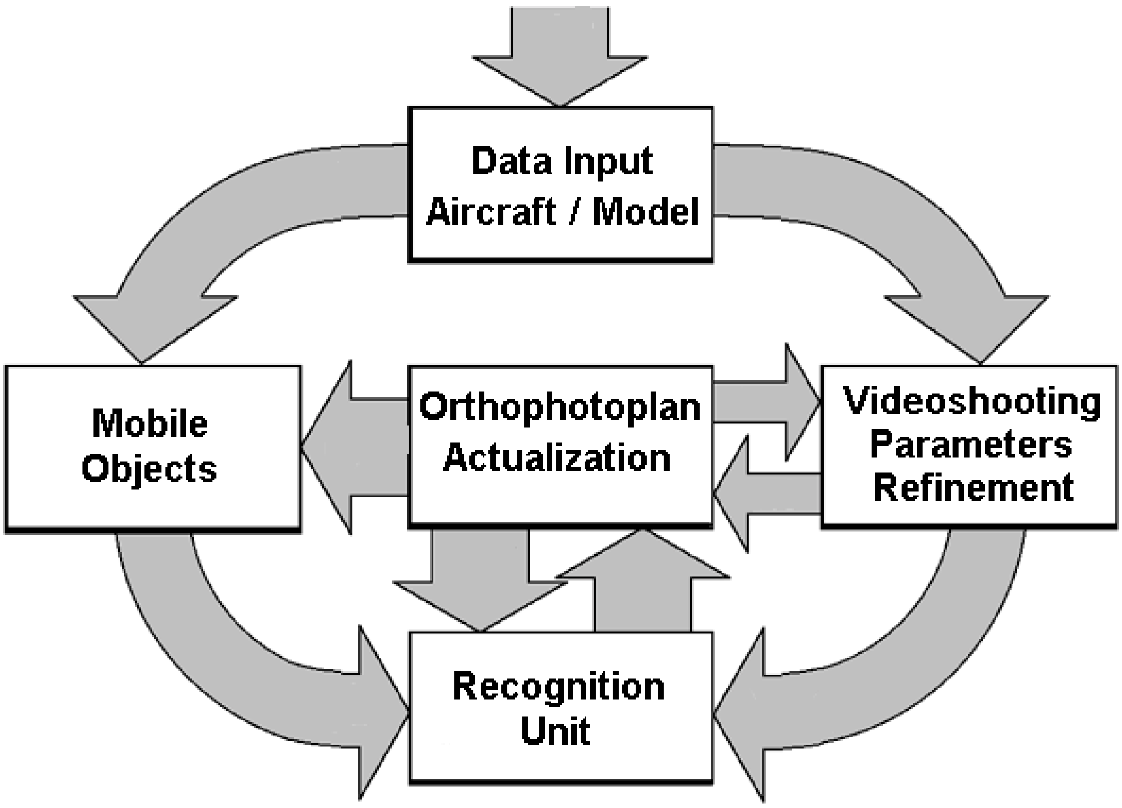

Figure 1. The structure of DEDAL-2

1.The videoshooting parameters refinement unit. An aircraft transmits data containing videosequences and parameters of videoshooting, such as coordinates, altitude and others. It is supposed that parameters of videoshooting contain an appreciable error. It makes impossible their direct usage for binding the images to orthophotoplan. The given unit, using available orthophotoplan, makes parameters of videoshooting more exact (refinement procedure). For this purpose, at the beginning of processing of each videoframe from the input videosequence, the system binds the frame to orthophotoplan. The exact binding is executed by comparizon of a videoframe and a corresponding segment on the ortophotoplan. Such an approach permits to achieve almost absolute accuracy of binding.

Essential feature of the approach is the preprocessing of both orthophotoplan and image [Grib97]. Both on orthophotoplan and on the image the borders are allocated and all further processing is carried out not with images, but with preparations of borders. This essentially reduces the volume of calculations. Thus the problem of finding of such videoshooting parameters is solved, at which one the greatest possible conformity of these boundaries is reached.

For allocation of boundaries the Canny method was used [Canny86], which one is obtained outgoing from following, rather natural, efficiency criteria:

- Good relation of a signal to a noise. Pithyly it means, that the method should select all true boundaries and thus to not select false.
- Good localization - the points marked as of boundary, should place as it is possible closer to their true position.
- Uniqueness of response on boundary.

As a result of binding the frame is supplied with precise values of coordinates, parameters of orientation and scaling respective to the corresponding fragment of orthophotoplan. After that there is possible exact localization of objects present on videoframe. Besides the binding to orthophotoplan allows to allocate both dynamic changes on terrain, and static. The outcomes of such allocation are subjected to further analysis for their identification.

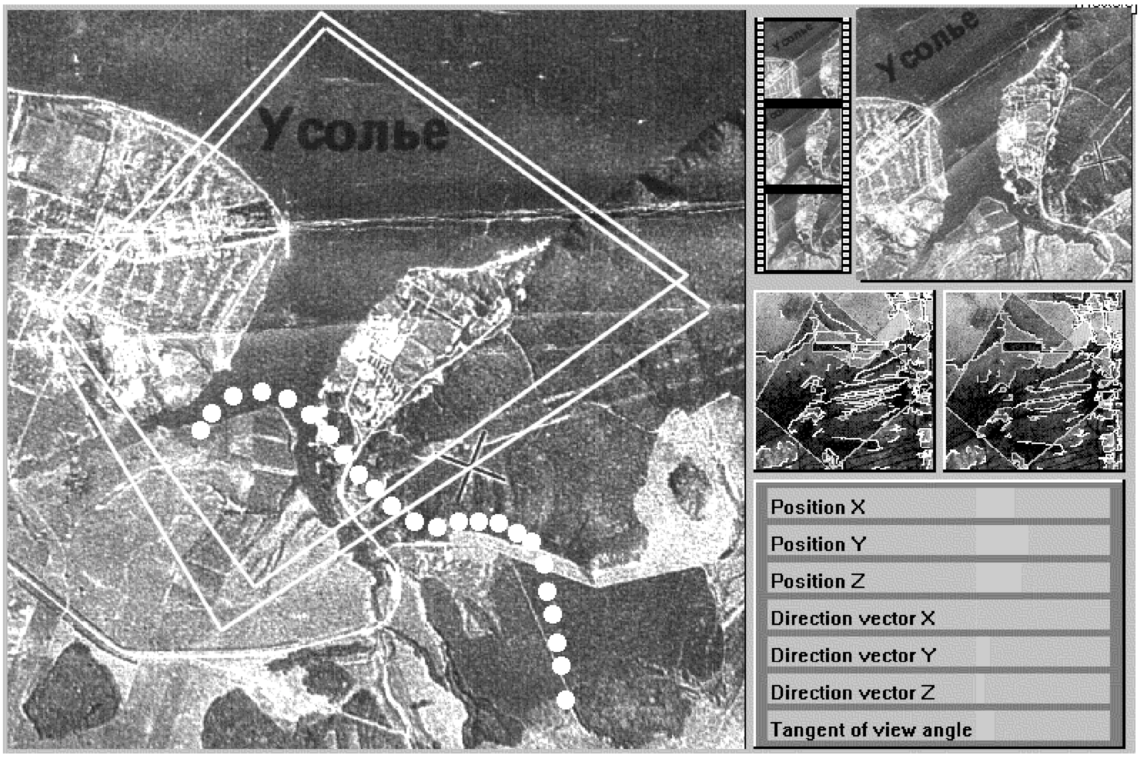

Figure 2. Binging a frame to orthophotoplan using border extraction

Binding of videoframe to orthophotoplan and videoshooting parameters refinement are illustrated in Fig.2. At the left: orthophotoplan, where the white circles show position of the aircraft at the moment of videoshooting. The two trapezoids demonstrate projections of boundaries of current frame before and after videoshooting parameters refinement. On the right: last three received frames, the current frame (which is larger); the last videoframe positioned on the orthophotoplan (before and after refinement); the boundaries are figured. Below is the chart showing the difference between initial and refined videoshooting parameters.

2.The motion detecting unit. The unit deals with a videosequence. It is required to allocate and to accompany with mobile objects.

In many real videoes, the scene is static and its contents varies only due to motion of the VC (dominant motion). For allocation of objects, driving in the scene, first it is necessary to establish parameters of the dominant motion. After compensation of the dominant motion, the next step is allocation and tracking actually mobile objects. Thus for each tracked object, the parameters of its motion are determined [Agiv85], [Ayer95], [Bober94], [Wang94]. It is supposed, that mobile objects take only small part of the images.

Figure 3. A mobile object is detected

At the beginning of mobile objects allocation process, the preprocessing of the images is made. For example, the images are resulted in one level of light exposure. Further the program determines the dominant motion conditioned by moving of the VC. Then the two consecuent images are transformed to common coordinate system. Finally, on the common part of two obtained images, the mobile objects are allocated (Fig.3). The history of motion of previously allocated objects is also taken into account.

After allocation and tracking of mobile objects, they can be transmitted to the recognition unit.

3.The recognition unit - detection of changes on terrain, identification of the detected changes and mobile objects. The detected objects are identified by means of statistical and spectral methods. Thus the reference images database is used.

The unit of comparison and recognition is intended for the analysis of character of static and dynamic changes. It consists of two subunits. First subunit detects distinctions between two images. Second one implements recognition of objects by comparison them with reference objects.

Detection of distinctions between two images (Fig.4) is done as follows. Let's name one of the images reference, another - tested. Due to videoshooting parameters refinement the two images have the same format, and the conterminous parts have identical scale, orientation, and are not shifted from each other. However, these two images can have different histograms of brightness; the tested image can be noicy, and also there can be various little objects, which to find out it is not required.

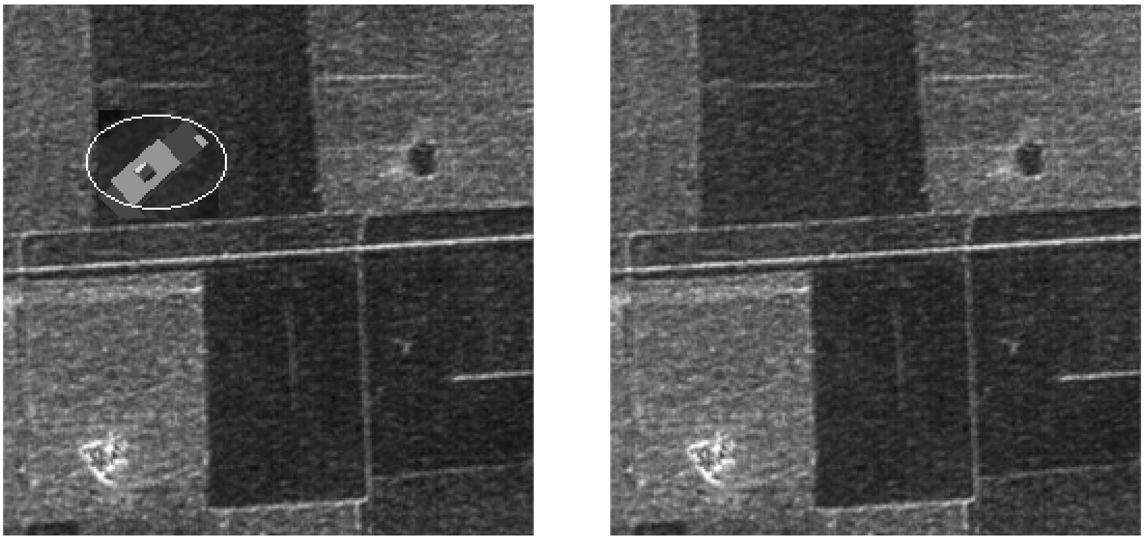

Figure 4. A static change is detected

To detect the distinctions between two images, each of them is cut up on small squares. The size of a square is a setting of the system. It depends on the size of small objects, which are not required to be found out. Usually side of such square is 16x16 pixels provided that the format of the whole image is 256x256 pixels. Program compares the corresponding squares on reference and tested images. For this purpose, it carries out the Fourier transforms of the images which have got into the specified squares and calculates correlation of their spectra. The correlation picture can strongly depend on histograms of brightness and noice. To reduce this dependence, high-frequency kill of spectra is previously made. For acceleration of the computation, the actually used version of fast Fourier transform is arranged so, that only those low-frequency factors of a spectrum are calculated which are necessary in the further work. Then the factors of spectra are set norms on unit as complex numbers and further cross correlation of spectra is evaluated. In outcome, the value of a correlation peak and its coordinates inside a square turn out for each pair of squares. This data is subject to cluster analysis. Pairs of the applicable each other squares, which contain only a little bit distinguishing images, give approximately identical values of correlation peaks. Thus their coordinates generate a cluster "of typical values". If in two squares the images differ, the values of peaks and their coordinates do not belong to this cluster (but usually also group in clusters). So it is possible to allocate squares containing distinctions and hence to find distinctions on two images.

Recognition of objects by matching them with reference means the following. The tested object should fall into to the class determined by some reference object, irrespective of tesetd object's position on the image, its scale, orientation, partial occlusion by other objects, luminosity histogram, structure of a background and - up to a certain degree - noice. Besides it is necessary to estimate a position of the identified object on the image.

The reference object is a 2D-image with large enough size and with a homogeneous background. The set of the reference objects is loaded into system.

4.*The unit for the orthophotoplan actualization.* This unit implements most evident operative performance of the areas shooted. It realizes the following opportunities:

- Display of the initial and updated orthophotoplans;
- Transformation of the next frame and its attachment to the resulting image;
- Allocation of areas with changes;
- Labeling special marks on the recognized objects.

The activity of the unit is as follows. At the beginning, the unit receives an orthophotoplan from the network. It contains the selected terrains. Then the reception of videoframes, refined videoshooting parameters, recognized objects and detected distinctions starts. The refined videoshooting parameters allow to map each videoframe onto orthophotoplan. This mapping results in a non-rectangular image - an orthofragment - whose scale and orientation meet to parameters of the entire orthophotoplan. The data processing is, basically, transformation of frame to orthofragment. Thus the factor of a stretching from the frame to orthophotoplan is taken into account. Those pixels of frame will be transformed only, for which the stretching is not too great. Through the network, the orthofragment is sent to the appropriate units for detection of distinctions between orthophotoplan and the videoshooting data. The distinctions and recognized objects are labeled on the orthophotoplan. Simultaneously, orthofragment is added to initial orthophotoplan, that gives actualized orthophotoplan - a picture of terrain on last moment (Fig.5).

Figure 5. Actualized orthophotoplan

At mapping the pieces of the actualized orthophotoplan upon its initial version, the recent data always replaces previous. Common boundary of the pieces mapped is also constructed.

The input and output packages have the following structure.

Orthophotoplan of terrain (received). The package contains the rectangular raster images selected from the database and appropriate to region of videoshooting which is determined by the flight mission. The format of the graphic file is BMP (file header, parameters and graphic pallet, massif of bits). The map is usually monochrome, though it is not necessary.

Refined videoframe from a videosequence (received). The package contains the image, number of the frame, refined videoshooting parameters. The videoshooting parameters consist of 12 floating-point numbers including coordinates of a videoshooting point, VC orientation angle, focal distance. The image is the videoshooting data in BMP graphic format (monochrome).

The data for motion detecting unit (transmitted). The package contains rectangular raster monochrome image in BMP format, image's coordinates on orthophotoplan and number of the frame.

Distinctions detected (received). The package contains rectangular raster monochrome image, whose non-zero pixels respond the retrieved distinctions. Also it contains coordinates of the image on orthophotoplan and number of the frame. The processing of the distinctions detected is a definition of a characteristic rectangle, which one allocates distinctions on the image. This processing consists of counting mean coordinates and sizes on the marked pixels.

Recognition results (received). The package contains number of the frame, refined videoshooting parameters, frame size, text of the description, coordinate of the center of object on the frame, size in scale of orthophotoplan and number of object under the catalogue.

5.The unit of data entry, including simulations of flight and videoshooting process. In the simulation mode, the route of motion, videoshooting parameters and conditions are preset. The system constructs a videosequence using orthophotoplan of 3D-model of a terrain. The videosequence is transmitted to the other units of the system [Zakh.97a], [Zakh.99.b].

The interface is arranged so that the system operator is not sidetracked by waste details. Each region of interest always can be highligted and zoomed. The display of distinctions and recognized objects can be switched off, if necessary. If parts of frames overlap, the objects and distinctions responding last frame are left only.

The approach, considered in the paper, allows to transform each computing unit to a multicomputer cluster. This way the problem of real time monitoring can be solved. Besides, further development of the system could be aimed at construction of image-understanding environment of a flow of images. Such a tool would gain to a new quality: adaptivity of each unit for the task to be solved. Due to this, one could test a number various methods for recognition, movement detecting, border extracting, etc. to construct the methods and heuristics with quality close to optimum.

REFERENCES

[Adiv85] G. Adiv, "Determining 3D-Motion and Structure from Optical Flows Generated by Several Moving Objects," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. *7,* no. 4, pp. 384-401, 1985.

[Ayer95] S. Ayer, "Sequential and Competitive Methods for the Estimation of Multiple Motions," PhD thesis, EPFL, Lausanne, 1995.

[Bober94] M. Bober and J. Kittler, "Robust Motion Analysis," *Proc. CVPR,* pp. 947-952, Seattle, June 1994.

[Canny86] J.Canny, "A Computational Approach to Edge Detection," *IEEE Pattern Anal. Machin. Intell.*, vol PAMI-8, pp 679-698, Nov 1986.

[Grib97] I.V.Gribkov, A.V.Zakharov, P.P.Koltsov, N.V.Kotovich, V.K.Nikolaev, "Preprocessing and 2D Image Recognition ", *Voprosy Kibernetiki, Red. V.B.Betelin*, Moscow, 1997, pp. 3-32.

[Wang93] J.Y.A. Wang and E.H. Adelson, "Layered Representation for Motion Analysis," *Proc. Computer Vision & Pattern Recognition Conf.,*pp. 361-366, New York, June 1993.

[Zakh97a] A.V.Zakharov, P.P.Koltsov, N.V.Kotovich, A.A.Kravchenko, A.S.Kutsaev, V.K.Nikolaev, "Virtual Reality: Methods of Realization", *Voprosy Kibernetiki, Red. V.B.Betelin*, Moscow, 1997, pp.105-154.

[Zakh99b] A.V.Zakharov, P.P.Koltsov, N.V.Kotovich, A.A.Kravchenko, V.K.Nikolaev "3D-Modeling and Problems of Recognition and Identification", *Voprosy Kibernetiki, Red. V.B.Betelin*, Moscow, 1999, pp. 92-109.