vision // Proc. of Imaging understanding workshop. 1981. P. 121–133.

3. Making good features to track better / T. Tommasini [et al.] // Proc/ IEEE Computer Society Conference on Computer Vision Pattern Recognition. 1998. P. 145–149.

4. Rares A., Reinders M. J. T., Biemond J. Recovery of partially degraded colors in old movie // Proc. of EUSIPCO-2002. Toulouse, 2003.

5. Forsyth D. A., Ponce J. Computer vision: modern approach : transl. from English. M. : Williams, 2004. P. 928.

M. V. Damov, A. G. Zotin
Siberian State Aerospace University named after academician M. F. Reshetnev, Russia, Krasnoyarsk

SCENE IMAGE CONSTRUCTION BY WAY OF SEQUENCED IMAGES SUPERPOSITION

A concept of scene image construction by way of video sequence frames or photo shot series superposition is presented in the paper. Scene construction is performed on the basis of displacement map which reflects vectors of blocks shifts.

Keywords: video sequence reconstruction, image superposition, scene construction, displacement map.

The level of development of modern computing technologies allows to solve problems of great computational complexity including processing of video sequences. Construction of an image of a video sequence scene is a problem necessary to be solved in reconstruction of video sequence. Reconstruction of video sequences is an important area of work in connection with increasing requirements of potential customers such as experts working with video archives, experts in the field of production and restoration of films, TV broadcast experts as well as experts in the analysis of visual data received by various methods namely: air photography, satellite photography, laser location and other systems of sensors. But generally a video sequence scene is understood as a part of a film or a sequence of images shot from one foreshortening for some period of time. In paper [1] one of algorithms of video sequence division into scenes is presented.

A displacement map is a two-dimensional data array whose elements are two-dimensional vectors. Each vector sets a shift from a point on the first image to a corresponding point on the second image. This information is used for construction of a scene image on the basis of several neighboring frames. Thus, the algorithm of construction belongs to a spatial-temporal type. There are three basic approaches to the definition of parameters of global movement: the approach with the use of the instrument of feature points; the approach with the use of vectors of blocks movement; global search. In this article the approach with the use of vectors of blocks movement or neighborhoods is considered. The advantage of the offered approach is the use of an image pyramid. At first the displacement map is searched for greatly reduced copies of images. The found values are initial ones for displacement maps for more detailed copies etc. Thus, at each level of detail it is required only to update a displacement map which considerably reduces time of calculation and probability of finding false values. At the same time the algorithm assumes the presence of large enough objects on a scene i. e. a piecewise-smooth displacement map. Let's present an algorithm of search of a displacement map.

There is a pair of images for which it is required to construct a displacement map. Two pyramids of detail (for each image) are constructed. A pyramid in the area of images processing is the representation of a source image in a set of images of the smaller resolution additionally processed by one of smoothing filters. Pyramid formation occurs by image smoothing at the previous level and by a choice of points with a step more than one pixel with the help of bilinear interpolation. A half-width σ in Gaussian function is connected with the relation k ($k > 1$) of sizes of pyramid images at the neighboring levels:

$$k = \sigma\pi/2, \quad (1)$$

where a half-width of Gaussian function is a distance between two extreme values of the independent variable for which the value of the function is equal to half of its maximum value.

On the one hand such choice leaves in the smoothed image only those frequencies which a reduced image will contain, and on the other hand it does not lead to loss of details. At the same time the formation of a detail pyramid for a displacement map takes place. The process of search of a displacement map occurs gradually beginning from the top of a detail pyramid. Processing of images near the top of a pyramid and near its base is different. For images near the tops of pyramids (strongly reduced images) we should find geometrical transformation (affine, projective) of images in the first detail pyramid combining it as a whole with an image at the same level of the second detail pyramid . Transformation search represents the updating of an algorithm presented in work [2] and is described below. The found transformation between images at the set level of a detail pyramid allows to calculate a displacement map at the same level. In transition to the next level of detail resolution the image in the first pyramid will be transformed according to a displacement map for current level, thus, only the specification of a displacement map is searched at the next level.

Beginning from some level it is no longer possible to find simple geometrical transformation combining the images in a detail pyramid. Images break into squares of a predetermined size (from 8 to 16 pixels), and for each of them geometrical transformation (usually a shift) and a displacement map is searched the same way which was applied to the whole image. It is obvious that transformation search is appropriate only in the case when one square does not grasp objects at different distances from a camera, otherwise different parts of a square should undergo different shifts. For the purpose of reduction of displacement jumps on borders of squares it is necessary to perform averaging of displacements. The larger the degree of squares overlapping the more smoothed the displacement map is, but at the same time the operating time increases.

A pair of images $I_1(\vec{x})$, $I_2(\vec{x})$ and a displacement map $\vec{d}(\vec{x},\vec{\omega})$ are considered as input data. Vector $\vec{\omega}$ is a vector of parameters setting a model of a displacement map. Two models of a displacement map – a shift and similarity transformation – are considered below. Shift

$$\vec{\omega} = (t_x, t_y), \quad (2)$$

$$\vec{d}(\vec{x},\vec{\omega}) = l\vec{t}, \quad (3)$$

where $\vec{x}$ is some known coordinates of a pixel in an image; t_x, t_y are pixel displacement between images on coordinates.

Similarity transformation

$$\vec{\omega} = (t_x, t_y, \lambda, \varphi), \quad (4)$$

$$\vec{d}(\vec{x},\vec{\omega}) = (1+\lambda)R_\varphi \vec{x} + l\vec{t} - \vec{x}, \quad (5)$$

$$R_\varphi = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\phi & \cos\varphi \end{pmatrix}, \quad (6)$$

where λ is a coefficient of image scaling; φ is an angle of image turning .

The parameter l is introduced for nondimensionalization of components $\vec{\omega}$ and reduction of their values to one order of magnitude so that each component does not exceed several tenths. Its numerical value is necessarily equal to the size of a diagonal of an image:

$$l = \sqrt{w^2 + h^2}, \quad (7)$$

where w and h are width and height of an image accordingly. In both cases:

$$\vec{d}(\vec{x}, 0) = 0. \quad (8)$$

The search of a displacement map is reduced to the search of $\vec{\omega}$. Let $\varepsilon(\vec{\omega})$ denote an integral from a square of difference of the second and the transformed first images:

$$\varepsilon(\vec{\omega}) = \int_\Omega \left(I_1\left(\vec{x} + \vec{d}(\vec{x},\vec{\omega})\right) - I_2\left(\vec{x}\right) \right)^2 d\vec{x},$$

$$\varepsilon(\vec{\omega}) = \int_\Omega \left(I_1\left(\vec{x} + \vec{d}(\vec{x},\vec{\omega})\right) - I_2\left(\vec{x}\right) \right)^2 d\vec{x}, \quad (9)$$

where Ω is image sizes.

The following designation is used below:

$$\vec{y} = \vec{x} + \vec{d}(\vec{x},\vec{\omega}). \quad (10)$$

It is required to find $\vec{\omega}$, minimizing $\varepsilon(\vec{\omega})$:

$$\frac{\partial\varepsilon}{\partial\vec{\omega}} = 2\int_\Omega \left(I_1(\vec{y}) - I_2(\vec{x}) \right) \frac{\partial I_1(\vec{y})}{\partial\vec{y}} \frac{\partial\vec{y}}{\partial\vec{\omega}} d\vec{x} = 0. \quad (11)$$

The following designations are used below:

$$\vec{g}^T \equiv \frac{\partial I_1(\vec{y})}{\partial\vec{y}}, \quad (12)$$

$$Y \equiv \frac{\partial\vec{y}}{\partial\vec{\omega}} = \frac{\partial\vec{d}(\vec{x},\vec{\omega})}{\partial\vec{\omega}}. \quad (13)$$

In the assumption of smoothness $I_1(\vec{x})$, small difference of $\vec{x}$ from $\vec{y}$ and taking into account (8) correct in the first order we get:

$$\vec{g}(\vec{y}) = \vec{g}(\vec{x}), \quad (14)$$

$$I_1(\vec{y}) = I_1(\vec{x}) + (\vec{g}, \vec{y} - \vec{x}), \quad (15)$$

$$\vec{y} - \vec{x} = Y\vec{\omega}. \quad (16)$$

Using (15, 16), the equation (11) is rewritten as follows:

$$\int_\Omega \left(I_1(\vec{x}) - I_2(\vec{x}) + \vec{g}^T Y\vec{\omega} \right) \vec{g}^T Y d\vec{x} = 0. \quad (17)$$

From (17) we get a linear system relative to $\vec{\omega}$:

$$A\vec{\omega} = \vec{b}, \quad (18)$$

where

$$A = \int_\Omega Y^T \vec{g}\vec{g}^T Y^T d\vec{x}, \quad (19)$$

$$\vec{b} = \int_\Omega \left(I_2(\vec{x}) - I_1(\vec{x}) \right) Y^T \vec{g} d\vec{x}. \quad (20)$$

The search of the solution of system (18) demands accuracy as an image (or its fragment) may not possess the texture open enough for reliable determination of $\vec{\omega}$. The idea consists in estimating an error of matrix **A** according to a priori value of noise n on an image and smoothing radius σ of original images. Further system (18) is solved with the help of *SVD*-decomposition, and singular values of the smaller threshold are nulled [2]. To estimate a threshold singular value of matrix **A** we will consider a random image with dispersion of pixels intensity $\langle n^2 \rangle$ and zero average value: $\langle n \rangle = 0$. In case of absence of borders in a local area for $\sigma \to 1$ the estimation for dispersion of the considered image derivatives [2] is carried out:

$$\mu^2 \equiv \frac{\langle |\vec{g}|^2 \rangle}{2} = \frac{\langle n^2 \rangle}{8\pi\sigma^4}. \quad (21)$$

Estimation (21) can be easily received by transition into a spectral area. A considered image can be presented in the form of a spectrum of a homogeneous and stochastic function whose dispersion can be calculated from the theorem of integral equality from a square of a signal module to an integral from a square of its spectrum. Image differentiation is reduced to multiplication of its spectrum by a corresponding derivative of Gaussian function which allows to estimate μ^2. Knowing the model of required transformation and estimation (21) it is possible to estimate a minimum singular value of matrix **A**:

$$S_{\min} = \mu^2 l^2 wh . \qquad (22)$$

Singular values of matrix **A** smaller than $S_{\min}$, should be nulled in calculation of a pseudo inverse matrix. In case of estimation of a minimum singular value during the work with a fragment variables ω, h, l the sizes of a fragment are designated in expression (22).

Let's present concrete models of displacement maps and an algorithm of superposition of the images received on the basis of analytical calculations.

Shift:

$$Y = \begin{pmatrix} l & 0 \\ 0 & l \end{pmatrix}. \qquad (23)$$

Similarity transformation:

$$Y = \begin{pmatrix} l & 0 & x & -y \\ 0 & l & y & -x \end{pmatrix}. \qquad (24)$$

An algorithm of images superposition:

Input: the first image I_1, the second image I_2.

Output: the transformation converting the first image into the second image.

Step 1. To calculate derivatives of the first image by convolution with Gaussian function. A half-width of Gaussian function should be of an order of expected values in a displacement map.

Step 2. To make matrix **A** of a systems according to expression (19).

Step 3. To make right part $\vec{b}$ according to expression (20) and to solve system of equations (18).

Step 4. To transform the second image according to the reversed transformation constructed from $\vec{\omega}$. It excludes the necessity to recalculate **A** on each iteration.

Step 5. If the module of vector $\vec{\omega}$ does not exceed the set value then pass to step 7.

Step 6. Pass to step 3.

Step 7. The algorithm ends.

As one can see from tests (sce figure) for piecewise-smooth surfaces the algorithm works well. Incorrect values are received near borders of objects. The algorithm operating time is proportional to a number of pixels on images.

The considered algorithm has two drawbacks: firstly, there are wrong values of displacement on borders of objects; secondly, the offered approach is in principle inapplicable in case of presence of small objects located close to a camera. It is connected with the fact that at low resolution the object disappears, and at high resolution the object shift turns out to be too big for the offered algorithm to be able to fix it. However the analysis of a difference frame can reveal similar problem objects and for them the application of other methods for post-processing is possible.

It is necessary to note also that as a result of algorithm work one-to-one correspondences between pairs of pixels of the first and second image which are true for a significant amount of points of an image (except occlusions) are received. It creates the possibility to use stable methods to find a fundamental matrix in epipolar geometry. Thus, the second pass of the algorithm which is carried out at imposed epipolar restrictions is possible.

It is necessary to note that data received by means of an instrument of feature points or an instrument of global search can be also used as initial data for the construction of a displacement map. A shortcoming of algorithms using special points as a base is low productivity, which is caused by high computing complexity of search operations (Feature Selection) and of feature points tracking (Feature Tracking) [3]. The advantage of similar algorithms consists in high reliability of definition of local movement. Algorithms of global search define optimum transformation without use of information on local movement [4]. Calculating complexity of these algorithms is so high that it practically excludes the possibility of their application for the solution of real problems.

a *b*

The image of a test video sequence scene and vectors of movement after the algorithm work (*a*); one of frames of a test video sequence with vectors of movement before the algorithm work (*b*)

Thus, at this stage there are two partially overlapping images I_1 and I_2; a displacement map representing an array of displacement vectors (Δx, Δy); the transformation converting the first image into the second (similarity or shift transformation).

The array of vectors found by an algorithm of movement definition has some characteristic features, namely, a movement vector is highly probably found incorrectly, if:

– the error of definition of a movement vector is great (it means that there is little similarity between compared blocks);

– the movement vector differs from the neighboring vectors considerably (it follows from smoothness of a field of movement vectors); the neighborhood dispersion is small (as a rule, in this case there is little difference between one block and neighboring ones). Hence it is necessary to introduce the quality or trust estimation taking into account the features of a field of vectors.

In paper [5] for the definition of movement vectors reliability it is offered to introduce the function of trust of the following kind:

$$fb_{x,y} = \left(a \cdot e_{x,y} + \frac{b}{d_{x,y}^{\;2}} + c \cdot \sigma_{x,y} \right)^{-1}, \qquad (25)$$

where a, b, c are parameters; $e_{x,y}$ is an error of a movement vector definition (a sum of absolute differences of pixels brightness in the neighboring area of a vector in a positions (x, y) from a current frame and a compared neighboring area from the previous frame; $d_{x,y}$ is a neighboring area dispersion in a position (x, y) from a current frame; $\sigma_{x,y}$ is a standard deviation of a movement vector in a position (x, y) from the neighboring vectors

$$\sigma_{x,y} = \frac{1}{4} \sum_{i=-1}^{1} \sum_{j=-1}^{1} \left(\left(p_{x,y}^{X} - p_{x+i,y+j}^{X} \right)^2 + \left(p_{x,y}^{Y} - p_{x+i,y+j}^{Y} \right)^2 \right), \qquad (26)$$

where $p_{x,y}^{X}$, $p_{x,y}^{Y}$ is a projection of a movement vector in a position (x, y) on the abscissa axis and ordinate axis accordingly. In paper [5] the following values of parameters of trust function received by an experimental way are given: $a = 0{,}25$; $b = 32$; $c = 1$. Movement vectors with small value of trust function are rejected; the other vectors form a set of reliable vectors.

In the process of formation of the scene I_s image it is necessary to take into consideration features of a displacement map. A displacement map displays only vectors for similar points located on overlapping images I_1 and I_2. In this connection for the points belonging to objects which appeared or disappeared again vectors of displacement will be absent. On a displacement map the display of such points will look like a vector of a zero size.

For generation of a scene on the basis of a displacement map one can use two approaches. The essence of the first approach is that a completion of an initial frame takes place. During this completion there occurs an expansion of borders and filling of new blocks with information according to the data from the next frame which underwent certain transformations. The second approach is that for each new frame the change of the size occurs by addition of a so-called loop received as a result of transformation of the boundary information from the previous frames and comparison.

As the area on frame borders is of greatest interest for us during the construction of the full image of a scene so for speed increase it is expedient to use boundary data for an analyzed frame. Taking into account the possibility of non-uniform movement on frame borders it is necessary to carry out calculation of displacement parameters for analyzed regions, however inside the region there can also be a non-uniformity. On the basis of information given above it was decided to apply the information about vectors of local movement for segments and data about similarity and shift transformations as a basis for formation of a scene image .

The general algorithm for a scene image formation on the basis of a displacement map will consist of the following steps:

Step 1. Formation of preliminary boundary segments. On this step boundary segments with width and height $W_\omega \times W_h$ accordingly are formed depending on the size of an analyzed area. We calculate them from dimension of a shot dividing it by 8. Then the generation of a card of segments M_{segm} taking into consideration their spatial position in conformity with initial division takes place.

Step 2. Calculation of vectors of local movement from frame I_1 to frame I_2. During this step calculation of average, maximum and minimum vectors of movement for each segment occurs. To calculate an average vector of movement one can choose the following ways to receive value VSs_i as an arithmetical mean or median value from corresponding elements of a displacement array .

Step 3. The analysis of local movement vectors and formation of displacement vectors . On this step local vectors of movement from frame I_1 to frame I_2 for each segment are analyzed and distances from a comparable segment for frame I_2 with borders of shot I_2 are calculated.

Step 4. The analysis of the received vectors of displacement for all segments and formation of an expansion map M_{scale}. The given map is formed as a rectangular area taking into consideration the greatest possible displacement for each segment. Unused areas, i. e. those areas in which movement is not defined unequivocally are marked on completion of an expansion map generation.

Step 5. Filling of a scene image on the basis of an expansion map and vectors of displacement taking into account possible transformation. For each vector of displacement (Δx, Δy) of vicinities (x, y) not equal to zero this vicinity is transferred into a scene image. For each vector of displacement (Δx, Δy) of vicinities (x, y) equal to zero the neighboring vectors are analyzed. The location of such vicinity in the scene image is calculated using data of the neighboring vectors. Then a vicinity is placed into a scene image using the calculated location.

After the work of an algorithm of scene formation there is an image of a scene of some height and width in the form of pixels array. The reliability of a scene image constructed on the basis of several neighboring shots of one scene of a video sequence can be estimated visually by means of an expert estimation. In fig. 1 the result of work of an algorithm after processing a scene of test video sequence as well as the found field of movement vectors are presented.

The result of research work is the algorithm of scene image formation on the basis of the neighboring frames of a video sequence scene. The offered algorithm is applicable for the solution of many important problems, for example, for construction of a panoramic image, video sequence stabilization and compression, video sequence reconstruction in systems of video editing and electronic video archives.

References

1. Damov M. V. Restoration of background in parts of a frame with small-size objects in videosequence // Vestnik. Scientific Journal of Siberian state aerospace university named after academician M. F. Reshetnev. Issue 1(27). 2010. P. 52–56.

2. Mowler K., Forsyte G., Malcolm M. Machine methods of mathematical calculations, transl : from English. Mosckov : World (Mir), 1980.

3. Making good features to track better / T. Tommasini [et al.] // Proc. IEEE Computer Society Conference on Computer Vision Pattern Recognition. 1998. P. 145–149.

4. Litvin A., Konrad J., Karl W. C. Probabilistic Video Stabilization Using Kalman Filtering and Mosaicking // Proc. of SPIE Conference on Electronic Imaging. 2003.

5. Soldatov S. A., Strelnikov K. N., Vatolin D. S. Quick and reliable definition of global motion in videosequences. // Materials of 16-th International conf. on computer graphics and its applications. 2006. P. 430 – 437.

I. M. Danilin
Siberian State Aerospace University named after academician M. F. Reshetnev, Russia, Krasnoyarsk

A. I. Danilin
Institute of Forest named after V. N. Sukachev, Siberian department of Russian Academy of Sciences, Russia, Krasnoyarsk

D. I. Svischev
Krasnoyarsk State Agrarian University, Russia, Krasnoyarsk

LASER LOCATION AND DIGITAL AERIAL SURVEY AS A SUBSATELLITE COMPONENT IN THE SYSTEM OF INFORMATION SUPPORT OF INVENTORY, MONITORING AND CADASTRE OF FOREST LANDS

Approaches and solutions in the area of forest remote sensing methods for the purposes of information support of inventory, monitoring and cadastre of forest lands with the use of innovation methods and high-end technologies of laser location, digital aerial survey and global satellite positioning are considered in the paper.

Keywords: laser location, digital aerial survey, satellite positioning, inventory, monitoring and cadastre of forest lands.

In the contemporary practice of forest use, monitoring and cadastre to receive reliable and efficient information about state and dynamics of forest lands is an urgent problem concerning both natural resources and environment and nature protection.

Nowadays laser location and digital aerial survey, which are major constituents of geomatics – the newest integral direction in the development of remote sensing methods of the Earth (aerial- and space survey), geoinformation technologies, digital photogrammetry and cartography, satellite geopositioning and telecommunications are more and more widely used in the solution of this problem in many countries of the world as well as in Russia.

Those advanced and highly effective methods today find wide application in many branches of economy. In fact they are an information basis of natural resources cadastres, land- and forest planning, ecological monitoring, systems of data gathering, processing and analysis. Judging by their accuracy and economical efficiency they exceed other methods of studying and measuring parameters of the Earth cover and natural systems [1–7].

Modern high-end aviation laser location systems develop quickly and today they have a scanning frequency of more than 200 thousand pulses (measurements) per second (fig. 1).

The highest density of scanning pulses is 1 pulse per 5–7 cm of the surface, the accuracy of measuring geometric parameters of the ground objects and vegetation morphostructural elements in plan and profile projections is about ± (5–10) cm. The accuracy of satellite positioning of forest line contours and boundaries, sample plots, separate trees and morphostructural elements of their stems and crowns, including under crown space, is practically unlimited and is determined by technical parameters of the GPS instruments [1; 4].