Results of comparative study on the problem of concrete strength characteristics prediction

Approach	Estimation of an error expectation, %	Estimated error variance, %
GASEN	4.119	0.040
GA based 1	4.113	0.047
GA based 2	4.012	0.036
Proposed approach	3.521	0.028

The results allow us to draw a conclusion that the proposed approach can effectively automatically generate neural networks ensembles and calculate the final solution of the ensemble. Thus, this integrated approach should be used to improve the effectiveness of the solution of complex practical problems which were traditionally solved by systems using a single artificial neural network, for example, complex problems of approximation and prediction.

References

1. Hansen L. K., Salamon P. Neural network ensembles // IEEE Trans. Pattern Analysis and Machine Intelligence. 1990. № 12 (10). P. 993–1001.

2. Cherkauer K. J. Human expert level performance on a scientific image analysis task by a system using combined artificial neural networks // Proc. AAAI-96 Workshop on Integrating Multiple Learned Models for Improving and Scaling Machine Learning Algorithms : P. Chan, S. Stolfo, D. Wolpert (eds.). Portland, OR : AAAI Press ; Menlo Park, CA. 1996. P. 15–21.

3. Hampshire J., Waibel A. A novel objective function for improved phoneme recognition using timedelay neural networks // IEEE Transactions on Neural Networks. 1990. № 1 (2). P. 216–228.

4. Goldberg D. E. Genetic algorithms in search, optimization and machine learning. Reading. MA: Addison-Wesley, 1989.

5. Semenkin E. S., Sopov E. A. Probabilities-based evolutionary algorithms of complex systems optimization. In Intelligent Systems (AIS'05) and Intelligent CAD (CAD-2005). M. : FIZMATLIT, 2005. Vol. 1. P. 77–79.

6. Perrone M. P., Cooper L. N. When networks disagree: ensemble method for neural networks // Artificial Neural Networks for Speech and Vision ; R. J. Mammone (ed.). N. Y. : Chapman & Hall, 1993. P. 126–142.

7. Jimenez D. Dynamically weighted ensemble neural networks for classification // Proc. IJCNN-98. Vol. 1. Anchorage: AK. IEEE Computer Society Press; Los Alamitos, CA. 1998. P. 753–756.

8. Zhou Z. H., Wu J., Tang W. Ensembling neural networks: Many could be better than all // Artif. Intell. 2002. Vol. 137, № 1–2. P. 239–263.

9. Koza J. R. The Genetic Programming Paradigm: Genetically Breeding Populations of Computer Programs to Solve Problems. Cambridge, MA : MIT Press, 1992.

10. Yeh I.-C. Modeling slump flow of concrete using second-order regressions and artificial neural networks // Cement and Concrete Composites. 2007. Vol. 29. № 6. P. 474–480.

11. URL: http://www.archive.ics.uci.edu/ml/datasets/ Concrete+Slump+Test

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

BACKGROUND RESTORATION IN FRAME AREAS WITH SMALL-SIZE OBJECTS IN VIDEO SEQUENCES

A general concept of removal of artificially overlaid images, natural damages of video images and other small-size objects is presented. The classification of artificially overlaid images is developed. The algorithms of feature points detection and feature points tracking used in video sequence restoration are considered.

Keywords: movement tracking, video sequence, feature points, texture, texture filling.

The task of restoration of video sequences is getting more and more actual as a result of computer engineering development. The restoration of the original images under artificially overlaid object (TV channel logotypes, subtitles etc.) and other small-size objects such as a man, a tree, a stone etc. on a certain background as well as the images distorted as a result of damage of information carrier (scratches on the film etc.) is of primary importance. The solution of this problem in general will result in reduction of costs of video reutilization such as old film remastering, original video forwarding by different TV channels with removal of earlier overlaid, but irrelevant now, computer graphics, and accidentally shot objects, for example, advertising structures.

Overlaid computer graphic images occurred in video can be divided into several groups. There are TV channels logotypes that can be defined as small-size images arranged in one or several frame corners or frame borders; the second group is titles, that is, text areas with information about film makers arranged in any place of a frame; the third group is subtitles, which can be defined as text areas near the top or bottom frame borders with

periodically changed static text; and finally a creeping line that can be defined as a text area at the top or bottom frame borders with the text moving according to generally accepted rules of reading and writing.

All variety of overlaid computer graphic images can be classified according to different criteria. Let's consider the most frequently used classifications. They are the following: according to their size: small (less than 5 % of the frame), middle (less than 20 % of the frame), huge (less than 35 % of the frame); according to their position: corner; stretched along the horizontal border of the frame; stretched along the vertical border of the frame, in accordance with Substation Alpha standard; or some other kind; according to their action: static (image is always constant), moderately changing (image size is constant), fully dynamic (image changes its size and some other video sequences may be overlaid within this size); according to their duration: always constant in video sequence, periodically absent in video sequence; according to their color: monochrome; black and white; gradient; with the limited number of colours; full-colour; according to their transparency: transparent and opaque; according to the presence of a border line: bordered, borderless; according to the presence of their own background: with background, without background [1].

Usually images distortions as a result of damage of information carrier have stretched geometric structure and can appear in any place of the frame with different angles of inclination. Their presence in several consecutive frames unassociated with the change of scene foreshortening is a unique feature of such structures. Several similar structures can be present in video sequence and each of them is characterized by its own behavior and can overlay others. Indeterminacy and unpredictability of damages appearance make the problem difficult to solve automatically. Only the hypothesis about the stretch of geometric structure, its small size relative to full frame, homogeneous brightness and stability of existence in frame sequence allows to devise a method of allocation and restoration of similar images.

Accidentally shot and unnecessary objects in the frame must be characterized by small size (less than 10 % of the image) relative to the size of frame, also they can be characterized by static position on dynamic background; by dynamic position on static background; by dynamic position on dynamic background.

The right classification of objects to be removed allows to choose the right complex of algorithms for restoration of original video sequence. The general order of restoration of original video sequence is presented below with the detailed description of every step:

Step 1. Determination of characteristics of a video sequence (feature points of a frame, movement vectors in a frame, object and texture movement in a frame).

One of the technologies used to extract structured and intelligent information from video sequence is feature points tracking in image sequence. A feature point is defined as a point in the scene situated on the plane area of the scene surface. Meanwhile, the depiction of environment of this point could be distinguished from the depictions of environment of other points by some particular environment of this point.

Harris detector that computes the value of special corner response function for every pixel is often used for the detection of feature points of an image or a frame. This function evaluates the degree of similarity between the depiction of point environment and the corner. To evaluate this degree matrix is calculated the following way:

$$\mathbf{M} = \begin{bmatrix} \left(\dfrac{\partial I}{\partial x}\right)^2 & \left(\dfrac{\partial I}{\partial x}\right)\left(\dfrac{\partial I}{\partial y}\right) \\ \left(\dfrac{\partial I}{\partial x}\right)\left(\dfrac{\partial I}{\partial y}\right) & \left(\dfrac{\partial I}{\partial y}\right)^2 \end{bmatrix},$$

where $I(x, y)$ is image brightness in the point with the coordinates (x, y).

If both eigenvalues are large then even small displacement of the point (x, y) provokes considerable changes in brightness that corresponds to the feature point of the image and corner response function is written as follows:

$$R = \det \mathbf{M} - k(\text{trace}\,(\mathbf{M}))^2,$$

where $k = 0.04$ (coefficient suggested by Harris); trace (**M**) is the sum function calculation of matrix elements on the principal diagonal.

The points of image corresponding to local maximum of corner response function are considered to be the feature points.

Let us consider a simple scheme of feature points tracking [2]:

1. Detection and evaluation.

1.1. Detection of a set of feature points {F} in terms of characteristics of feature points such as the degree of corner response function extremity, the position of feature point (in the image center, at the image borders, at the image corners), the position of feature point relative to other feature points or compactness of feature points in some area of the image.

1.2. Detection of quality of all feature points – $Q\{F\}$. The feature points with large degree of corner response function extremity, sufficiently remote from the frame border, with low compactness of feature points in the area of the frame of interest are considered to be the most qualitative feature points. Methods of multi-attributive decision making, for example, the method of ordered preference by way of similarity with the ideal decision, can be used for evaluating of the quality of feature points.

1.3. Choice of feature points with the quality higher than some earlier or dynamically defined threshold, and generation of the set of feature points {G}.

2. Tracking and evaluation.

For each next frame:

2.1. Finding of a new position of all feature points from set {G} – tracking in the current frame;

2.2. Estimation of current quality of all the elements of set {G};

2.3. Choice of such feature points only whose quality satisfies a certain criterion. As a rule, this is an integral

criterion or a degree of corner response function extremity.

2.4. If the number of tracking feature points becomes less than the required number then a detector of feature points is applied to the current image and a number of new points is added to set {G}.

Modification of Lucas-Kanade algorithms is applied for tracking of feature points coordinates changes [3]. The last modification of Lucas-Kanade algorithms is Favaro-Soatto algorithm that takes into consideration the displacement of feature points, affine distortions of feature points, affine changes of brightness of feature points. The problem of feature points tracking is reduced to the problem of definition of movement parameters and distortion of feature point space. Then the difference is minimized:

$$c = \iint_W \left(J(Ax+d) - I(x) \right)^2 w(x)dx,$$

where W is feature point space, and $w(x)$ is weight function which can be equal to one in the whole space, $J(x)$ и $I(x)$ are two images; $Ax + d$ is the displacement of points.

The expression is differentiated relative to the movement parameters, and derivative function is equal to zero. Then the system is linearized by means of image function expansion in Taylor's series:

$$J(Ax+d) = J(x) + g^T(u).$$

It gives us a linear system of six equations with six unknown quantities:

$$\mathbf{Tz} = \mathbf{a},$$

where all required parameters are merged in vector $\mathbf{z}$:

$$\mathbf{z^T} = [d_{xx} \quad d_{yx} \quad d_{xy} \quad d_{yy} \quad d_x \quad d_y].$$

Error vector $\mathbf{a}$ is written as:

$$\mathbf{a} = \iint_W (I(x) - J(x)) \begin{bmatrix} xg_x \\ xg_y \\ yg_x \\ yg_y \\ g_x \\ g_y \end{bmatrix} wdx,$$

and matrix $\mathbf{T}$ can be presented as:

$$\mathbf{T} = \iint_W \begin{bmatrix} \mathbf{U} & \mathbf{V} \\ \mathbf{V^T} & \mathbf{Z} \end{bmatrix} wdx,$$

$$\mathbf{U} = \begin{bmatrix} x^2 g_x^2 & x^2 g_x g_y & xyg_x^2 & xyg_x g_y \\ x^2 g_x g_y & x^2 g_y^2 & xyg_x g_y & xyg_y^2 \\ xyg_x^2 & xyg_x g_y & y^2 g_x^2 & y^2 g_x g_y \\ xyg_x g_y & xyg_y^2 & y^2 g_x g_y & y^2 g_y^2 \end{bmatrix},$$

$$\mathbf{V^T} = \begin{bmatrix} xg_x^2 & xg_x g_y & yg_x^2 & y \\ xg_x g_y & xg_y^2 & yg_x g_y & yg_y^2 \end{bmatrix},$$

$$\mathbf{Z} = \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix}.$$

The obtained system of equations is resolved also iteratively with the help of Newton–Raphson method.

If movement is considered to be not affine, and just a displacement, then first four elements of the required vector $\mathbf{z}$ become zero and only the last two remain meaningful. The algorithm is turned into Tomashi-Kanade algorithm.

Let us supplement the algorithm for variable brightness case.

Let the scene surface for which the scene feature point is detected will be Lambert one. Then the intensity of point brightness $x = P(X)$, where X is scene point, P is projection operator, x is image point, can be described as:

$$I(x,0) = \upsilon_E E(X) + \xi_E \ \forall_X \in W_U,$$

where $E(X)$ is albedo (reflection power) of scene point; U is environment of scene point; υ and E are permanent parameters which present the changes of contrast and brightness respectively. While a camera is moving the parameters are changing, that is, they depend on time. Changes of brightness in time can be written as:

$$I(x,0) = \upsilon(t) I(x,t) + \xi(t) \ \forall_X \in W_U,$$

where

$$\upsilon(t) = \frac{\upsilon_E(t)}{\upsilon_E(t)}, \ \xi(t) = \xi_E(t) - \frac{\upsilon_E(t)}{\upsilon_E(t)} \xi_E(0),$$

if $t > 0$. Having merged affine movement of environment of feature point with changes of brightness we get the following expression:

$$I(x,0) = \upsilon(t) I(Ax+d,t) + \xi(t) \ \forall_X \in W.$$

This equation will never be realized in reality because of noise in the image, approximate modeling of movement and changes of brightness. That's why the task of feature points tracking is reduced to minimization of difference between the environments of current and new feature point position:

$$c = \int_W \left(I(x,0) - \upsilon I(Ax+d,t) + \xi \right)^2 w(x)dx,$$

where $w(x)$ is weight function. With the help of expansion in Taylor's series in the environment $d = 0$, $\upsilon = 1$, $\xi = 0$ we get:

$$\upsilon I(y,t) + \xi \approx \upsilon I(x,t) + \xi + \nabla I \frac{\partial y}{\partial u}(u - u_0),$$

$$y = Ax + d \ A = \{d_{ij}\},$$

$$A = \{d_{ij}\},$$

$$\mathbf{d} = [d_1 \quad d_2]^T,$$

$$\mathbf{u} = [d_{11} \quad d_{12} \quad d_{21} \quad d_{22} \quad d_1 \quad d_2]^T,$$

$$\mathbf{u_0} = [1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0].$$

Having rewritten the obtained equality in matrix form we get the following equation:

$$I(x,0) = F(x,t)^T z,$$

where

$$\mathbf{F}(x,t) = \begin{bmatrix} xI_x & yI_x & yI_y & I_x & I_y & I & 1 \end{bmatrix},$$

$$\mathbf{z} = \begin{bmatrix} d_{11} & d_{12} & d_{21} & d_{22} & d_1 & d_2 & \upsilon & \xi \end{bmatrix}^T.$$

Having multiplied by $F(x,t)^T$ and having integrated over all space of feature point W with weight function w, we get a system of eight equations with eight unknown quantities:

$$\mathbf{Sz} = a,$$

$$a = \int_W F(x,t)^T I(x,0) w(x) dx,$$

$$S = \int_W F(x,t)^T F(x,t) w(x) dx.$$

Having changed integration into the sum of all pixels in space W we get the following system of linear equations:

$$\mathbf{Sz} = a,$$

$$\mathbf{S} = \sum_{x \in W} \begin{bmatrix} \mathbf{T} & \mathbf{U} \\ \mathbf{U^T} & \mathbf{V} \end{bmatrix},$$

$$\mathbf{T} = \begin{bmatrix} x^2 I_x^2 & xy I_x^2 & x^2 I_x I_y & xy I_x I_y & x I_x^2 & x I_x I_y \\ xy I_x^2 & y^2 I_x^2 & xy I_x I_y & y^2 I_x I_y & y I_x^2 & y I_x I_y \\ x^2 I_x I_y & xy I_x I_y & x^2 I_y^2 & xy I_y^2 & x I_x I_y & x I_y^2 \\ xy I_x I_y & y^2 I_x I_y & xy I_y^2 & y^2 I_y^2 & y I_x I_y & y I_y^2 \\ x I_x^2 & y I_x I_y & x I_x I_y & y I_x I_y & I_x^2 & I_x I_y \\ x I_x I_y & y I_y^2 & x I_y^2 & y I_y^2 & I_x I_y & I_y^2 \end{bmatrix},$$

$$\mathbf{U^T} = \begin{bmatrix} x I_x I & y I_x I & x I_y I & y I_y I & I_x I & I_y I \\ x I_x & y I_x & x I_y & y I_y & I_x & I_y \end{bmatrix},$$

$$\mathbf{V} = \begin{bmatrix} I^2 & I \\ I & 1 \end{bmatrix}.$$

If matrix **S** is invertible then the solution of the system of linear equations can be written as follows:

$$\mathbf{z} = \mathbf{S}^{-1} a.$$

As in all tracking algorithms the system is solved iteratively with the help of Newton–Raphson method. Iterations are carried out until the solution changes become negligibly small.

Step 2. Division of a video sequence into scenes.

To increase the quality of restoration of video sequence it is necessary to divide the video sequence into scenes. The division occurs according to the following algorithm:

1. Calculation of distance from every feature point to the central point of frame:

$$R_{ij} = \sqrt{(x_{G_i} - x_c)^2 + (y_{G_i} - y_c)^2},$$

where x_{G_i}, y_{G_i} are coordinates of the i-th feature point; x_c, y_c are coordinates of the central point of the frame.

2. Calculation of point displacement:

$$\left| R_{ij} - R_{ij-1} \right| < e,$$

where e is the threshold of point displacement in the frame.

3. Calculation of a number of strongly displaced points in j-th frame:

$$f(R, e, j) = \text{count}(e > e_n),$$

where e_n is the general threshold of displacement. If function f obtains local maximum in the current frame j then current and next frames of video sequence are the borders of the scene. The quality of detection of the scene border can be described by the following parameters:

– accuracy, that is the probability that the detected scene border is not wrong:

$$P = \frac{C}{C+F};$$

– threshold signal that is the probability that the expected scene border will be found:

$$V = \frac{C}{C+M};$$

– synthetic measure of accuracy based on accuracy and threshold signal:

$$F1 = \frac{2PV}{P+V},$$

where C is the number of right actuations; M is the number of missed scenes; F is the number of wrong actuations.

Step 3. Detection of a scene type (with movement, without movement).

If feature points movement vectors were detected successfully at the stage of tracking of feature points of video sequences then let us consider that this video sequence and scene as its part have movement. Hence at the stage of restoration of the original video sequence let us use the spatial-temporal algorithm of obtaining information from the neighboring frame. When movement vectors can not be detected successfully at the stage of tracking of feature points of video sequences then let us consider that the scene does not have movement and use spatial algorithm of restoration of video sequences based on obtaining the information from neighboring area in the current frame.

Step 4. Detection of borders of artificially overlaid computer graphics areas in case of restoration of video sequence of such type.

The localization of text areas with artificially overlaid computer graphics is based on modification of Rares–Reinders–Biemond spatial algorithm [4]. This algorithm is constructed on the principle of detection of extreme brightness areas on the basis of soft and hard dynamic thresholds. To detect the areas of extreme brightness we must define some thresholds for searching and localization of bright and dim pixels. However, the use of fixed thresholds is not desirable as brightness changes from frame to frame. A hard threshold is a good solution

for detection of such areas. A soft threshold results in a large quantity of wrong actuations. To avoid these problems the dynamic threshold is used to detect extreme brightness areas. The algorithm works in our case rather efficiently. The main idea of choosing the dynamic threshold is that firstly a hard threshold is chosen. Then only the areas with brightness values higher than the initial threshold are detected. The areas obtained at this stage are expanded by the neighboring areas satisfying the soft threshold.

Step 5. Detection of characteristics of areas to be restored, choice of the complex of algorithms for restoration of video sequence.

For video sequence with movement signs in the frame the structure of several previous frames of video sequence and the changes of the obtained structure of the previous frames in comparison with the current frame are analyzed. The decision about modification of the current frame with the use of information obtained from the previous frames taking into consideration the changes of the structure of the previous frame is made on the basis of the obtained data. For video sequence without movement signs the texture of the neighboring area is analyzed. Then the structure and the probability of its changes are defined. Texture area analysis with the help of dynamic sized space and comparison of the image elements on space borders can be a good alternative. It can be suggested that when the main elements of the image coincide on the space border then the image inside the space is a desired texton and on the basis of this image it can be acceptable to generate the texture for filling of the object to be removed. The filling of restoration area is carried out taking into consideration the data obtained.

For a video sequence scene with movement signs let us suggest that the position presented in rectangular coordinates relative to the top left corner of the frame $(x1_a, y1_a, x2_a, y2_a)$ and linear sizes $(d_x = x2_a - x1_a; d_y = y2_a - y1_a)$ of the restoration area are well-known, the frame moves in one direction in general, its motion being uniform and linear. After finishing of Lucas–Kanade algorithm's work we know the position of the set of feature points G_i for each frame, the previous position of the set of feature points G_{i-1} in one of previous frames and vector (x_v, y_v) or direction and movement value of each feature point between a couple of related frames. Having this information we get the possibility to calculate the frame number relative to the current frame from which we shall obtain the information for restoration. The description of work of this algorithm is given below.

In general the restoration area can be rectangular in shape that is why the frame number n is computed as minimal so that the restoration point is beyond the restoration area $n = \min(d_x/x_v; d_y/y_v)$ meanwhile the displacement of the replacement point with respect to the i-th frame will be $x_{i-n} = n \cdot x_v; y_{i-n} = n \cdot y_v$, and the coordinates will be the following:

$$\begin{bmatrix} x'_i \\ y'_i \end{bmatrix} = \begin{bmatrix} x_{i-n} \\ y_{i-n} \end{bmatrix} = \begin{bmatrix} x_i - n \cdot x_v \\ y_i - n \cdot y_v \end{bmatrix},$$

there i is the current frame; n is the frame displacement; $i-n$ is the previous frame containing information for restoration; $[x_i, y_i]$ is the restoration point; $[x'_i, y'_i]$ is the point to be reconstructed, coordinates value contains colour; $[x_{i-n}, y_{i-n}]$ is the point on the previous frame used for restoration.

Restoration process is repeated for each point of the restoration area for each frame of the restoration scene of video sequence in order to restore the whole scene of video sequence. We can use the reconstructed points for the restoration of other points of the same frame or the reconstructed frames for the restoration of other frames of the scene to be reconstructed.

This algorithm is applicable for the restoration of areas of any foreground objects mentioned above, but constraints are imposed on the size and position of the restoration area. In case if the restoration area is not situated near the frame border it can occupy up to 90 % of the frame on its larger side and 10 % on its smaller side. In case if the restoration area is situated near the frame border then the side length of the restoration area cannot exceed 10 % of the frame border length.

For the scene of video sequence without movement signs for small-sized restoration areas we can use a modified algorithm introduced by Ponce and Forsyth [5]:

1. Choice of a texture fragment in a desired located area in terms of the hypothesis about the continuation of similar texture in the restoration area.

2. Insertion of a texture fragment into the restoration area of image (until the restoration area is filled) in the cycle.

3. Selection of the environment of the position after the example of image, ignoring the estimation of the similarity of position with undefined values in process of calculation; choice of the value for this position from the set of relevant values of the chosen environments in a random and equiprobable way (until all the points values on the borders of the restoration area are selected in the cycle).

4. End of the cycle, step 3 and step 2.

A disadvantage of the presented algorithm is its computational complexity and high dependence on random values. Advantages are the solution of the problem of pouring the structure of areas with the uncertain form, and joining of the generated and the initial image. The results of the algorithm can also be improved by using a median filter. The desirable video sequence should be the result of algorithms work, but the main drawback is that the results can be estimated only subjectively or by a number of experts.

References

1. Damov M. V. Spatial method of localization of logotypes images in video sequences // Science. Technology. Innovations. NTI-2008 : Materials of the All-Russian scientific Conference of young scientists. P. 1. Novosibirsk, 2008. P. 191–193.

2. Lucas B. D., Kanade T. An iterative image registration technique with an application to stereo

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:

$$k = \sigma\pi/2, \tag{1}$$

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