

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.

For more information visit www.intechopen.com



Pseudogradient Estimation of Digital Images Interframe Geometrical Deformations

A.G. Tashlinskii

*Ulyanovsk State Technical University
Russia*

1. Introduction

Nowadays the systems of information extraction that include spatial apertures of signal sensors are widely used in robotics, for the remote exploration of Earth, in medicine, geology and in other fields. Such sensors generate dynamic arrays of data having the proper feature which is in their space-time correlation and due to which they can be represented in the form of multidimensional images (Gonzalez & Woods, 2002). When producing algorithmic software for the processing of such images it is necessary to take into account the dynamics of the scene to be observed, distortions caused by signal propagation environment, spatial movements of signal sensors and imperfection of their construction. The influence of the mentioned factors can be described through mathematical models of space-time deformations of multidimensional grids with the specified images.

The estimation of varying parameters of image spatial deformations is required not only in robotics applications, but also to solve a wide range of other problems, for example, for automated search of a fragment on the image, navigational tracking of mobile object course in the conditions of limited visibility, combination of multiregion images at remote explorations of Earth, in medical research. A lot of scientific calls for papers are devoted to different problems of image sequence space-time deformation parameters estimation (the bibliography is given, for example, in (Tashlinskii, 2000)). This chapter is devoted to one of the directions of solving this type of problems, where pseudogradient estimation of image interframe geometrical deformations (IIGDs) is considered.

Let us assume that the model of IIGDs is defined to an accuracy of a parameters vector $\bar{\alpha}$ and the estimation quality criterion is formulated in terms of some functional $J(\bar{\alpha})$ minimization showing expected losses. However, it is impossible to find optimal parameters $\bar{\alpha}^*$ in the mentioned sense in view of incompleteness of description of the images to be observed. In this case we can estimate the parameters $\bar{\alpha}$ on the basis of a given realization Z analysis of the image to be observed by means of some adaptation procedure which minimizes $J(\bar{\alpha}) = J(\bar{\alpha}, Z)$ for the given realization Z . However, it is reasonable to avoid this intermediate state of the research and estimate $\bar{\alpha}$ directly on values $J(\hat{\bar{\alpha}}, Z)$ (Polyak & Tsypkin, 1984):

$$\hat{\bar{\alpha}}_i = \hat{\bar{\alpha}}_{i-1} - \Lambda_i \nabla J(\hat{\bar{\alpha}}_{i-1}, Z), \quad (1)$$

where $\hat{\alpha}_t$ - next after $\hat{\alpha}_{t-1}$ approximation of the minimum point of $J(\hat{\alpha}, Z)$; Λ_t - gain matrix (positively defined matrix determining a value of the estimates change at the t -th iteration); $\nabla J(\hat{\alpha}_{t-1}, Z)$ - gradient of the functional $J(\hat{\alpha}_{t-1}, Z)$. The necessity of multiple cumbersome calculations hinders the procedure (1) application in the image processing. It is possible to significantly reduce computational expenses due to the usage of contraction $\nabla J(\hat{\alpha}_{t-1}, Z_i)$ instead of $\nabla J(\hat{\alpha}_{t-1}, Z)$ at some part Z_i of realization Z at each iteration choosing, for example, Z_i in the form of a sliding window. For relatively large-sized images, the analysis of approaches (Tashlinskii, 2000; Minkina et al., 2007) to the synthesis of algorithms of IIGDs estimation in real time showed that it is expedient to seek a decision, satisfying the requirements of simplicity, rapid convergence and efficiency in various real situations, among recurrent non-identification algorithms. The pseudogradient algorithms (PGAs) constitute the most representative class of such algorithms.

The conception of the pseudogradient was introduced in work (Polyak & Tsypkin, 1973). A unified approach to the analysis and synthesis of various procedures of the stochastic minimization has been developed on the basis of it. For the given problem to be solved the pseudogradient $\bar{\beta}_t$ may be represented as any random vector in the parameter space depending on a function of losses and estimates $\hat{\alpha}_{t-1}$ at the t -th iteration if the following condition is satisfied

$$[\nabla J(\hat{\alpha}_{t-1}, Z)]^T M\{\bar{\beta}_t\} \geq 0, \quad (2)$$

where T - sign of transposition; $M\{\cdot\}$ - symbol of the mathematical expectation. In the geometrical interpretation the vector $\bar{\beta}_t$ is the pseudogradient if it makes, on average, an acute angle with the exact value of the functional $J(\hat{\alpha}, Z)$ gradient. The class of PGAs includes algorithms of stochastic approximation, random search and many others.

The following procedure is used in PGA (Tsypkin, 1995)

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} - \Lambda_t \bar{\beta}_t, \quad (3)$$

where $\bar{\alpha}$ - vector of the parameters to be estimated; $t = \overline{1, T}$ - iteration number; $\hat{\alpha}_0$ - initial approximation of the parameters vector; T - number of iterations. The algorithm is considered to be pseudogradient if $\bar{\beta}_t$ is the pseudogradient at each its iteration. In this case the iterations are, on average, performed in the direction of reduction of $J(\bar{\alpha})$ and sequence $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_t, \dots$ converges to the optimal parameters when satisfying relatively weak conditions (Polyak, 1976).

If realizations α_t , $t = 1, 2, \dots$, of the parameter α to be estimated are observed as, for example, in the problems of image brightness prediction, then we can choose $\beta_t = \hat{\alpha}_t - \alpha_t$ as the pseudogradient, where the estimate $\hat{\alpha}_t$ is found on realization Z or on a part of realization Z_i . In problems of image processing the quality functional $J(\bar{\alpha}, Z)$ is often expressed through the mathematical expectation of some function $f(\bar{\alpha}, Z)$:

$$J(\bar{\alpha}, Z) = M\{f(\bar{\alpha}, Z)\}. \quad (4)$$

In particular, it can be mean square of error of some value χ :

$$f(\bar{\alpha}, Z) = (\hat{\chi}(\bar{\alpha}, Z) - \chi)^2 = \Delta^2(\bar{\alpha}, Z),$$

where χ - the exact value and $\hat{\chi}(\bar{\alpha}, Z)$ - its estimate. In this case the condition of the pseudogradientness is met if differentiation under the symbol of the mathematical expectation in (4) is possible.

We should also mention that the procedure (3) does not require compulsory finding $J(\hat{\alpha}_{t-1}, Z_t)$ or $\nabla J(\hat{\alpha}_{t-1}, Z_t)$, i.e. $J(\bar{\alpha})$ can be non-observable. It is necessary to meet only the condition of the pseudogradientness. At the non-observable realization an auxiliary observable quality functional $Q(\bar{\alpha})$ can be introduced and a noisy value $Q(\bar{\alpha})$ can be chosen as $\bar{\beta}_t$, whose point of extremum (not necessary the point of minimum) is obtained at the same optimal parameters. Later on, this chosen functional characterizing the estimation quality will be called the goal function. For example, when estimating the mathematical expectation of random value X the following value can be selected as the goal function

$$Q(\alpha, X) = M\{(X - \alpha)^2\},$$

then, in the simplest case $\beta_t = x_t - \hat{\alpha}_{t-1}$, where x_t - value X at the t -th iteration. When estimating the correlation coefficient between the centered values X and Y the goal function can be represented as

$$Q(\alpha, X, Y) = M\{(\alpha X - Y)^2\},$$

then, $\beta_t = (\hat{\alpha}_{t-1}x_t - y_t)$, where x_t and y_t - realizations of X and Y at the t -th iteration.

The problem of IIGDs estimation considered in this chapter is related to the second type of problems, where it is necessary to use the auxiliary quality functional.

Let us note one more important property of the pseudogradient procedures that consists in that, $\bar{\beta}_t$ assumes dependence on estimation values $\hat{\alpha}_p$, $p < t$ in the preceding samples and rows of the image that enables to run image processing in the order of some sweep. The last property is very important at practical realization of the algorithms.

Thus, to synthesize fast PGAs of parameters estimation $\bar{\alpha}$, it is necessary to find a relatively easily calculated pseudogradient of the given goal function of the estimation quality. In the next part, several possible ways of computational expense reduction when finding the goal function pseudogradient are considered.

2. Choice of pseudogradient

When synthesizing PGA the important stages are in the choice of a goal function and a rule of finding its pseudogradient. Let us consider some approaches to solve these problems.

Let the studied frames $\mathbf{Z}^{(1)} = \{z_{\bar{j}}^{(1)} : \bar{j} \in \Omega_{\bar{j}}\}$ and $\mathbf{Z}^{(2)} = \{z_{\bar{j}}^{(2)} : \bar{j} \in \Omega_{\bar{j}}\}$ of images specified on a regular samples grid $\Omega_{\bar{j}} = \{\bar{j} = (j_1, \dots, j_n) : j_k = \overline{1, N_k}\}$ represent additive mixture of the informational pattern $\mathbf{X} = \{x_{\bar{j}}\}$ and a pattern $\Theta = \{\theta_{\bar{j}}\}$ of an independent noise:

$$\mathbf{Z}^{(1)} = \mathbf{X}^{(1)} + \Theta^{(1)}, \quad \mathbf{Z}^{(2)} = \mathbf{X}^{(2)} + \Theta^{(2)}, \quad (5)$$

where $\bar{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m)^T$ - vector of unknown geometrical transformation parameters of the image $\mathbf{X}^{(1)} = \mathbf{X}$ into the image $\mathbf{X}^{(2)} = \mathbf{X}(\bar{j}, \bar{\alpha})$, for example, rotation, shift in some direction, scale change etc. In doing so, $x_{\bar{j}}$, $\theta_{\bar{j}}^{(1)}$ and $\theta_{\bar{j}}^{(2)}$ are homogeneous and have Gaussian distribution with zero mean and known covariance functions $R_{x_{\bar{j}}x_{\bar{l}}} = M\{x_{\bar{j}}x_{\bar{l}}\}$;

$$R_{\theta_{\bar{j}}\theta_{\bar{l}}} = \sigma_{\theta}^2 \delta_{\bar{j}, \bar{l}}, \quad \text{where } \delta_{\bar{j}, \bar{l}} = \begin{cases} 1, & \text{if } \bar{j} = \bar{l}; \\ 0, & \text{if } \bar{j} \neq \bar{l}. \end{cases} \quad \text{- Kronecker symbol; } \bar{j}, \bar{l} \in \Omega.$$

Under the assumed constraints the goal function for the gradient estimation of the parameters vector $\bar{\alpha}$ can be written using the conditions of the optimal estimation obtained by means of the maximum likelihood method in work (Vasiliev & Tashlinskii, 1998). In particular, it is shown that if the image $\mathbf{Z}^{(1)}$ is noisy, then the maximization of the likelihood function is almost the same as minimization of the quadratic form. Then, for the gradient of the goal function we obtain

$$\nabla J(\bar{\alpha}, Z) = \nabla \left[\left(\dot{\mathbf{Z}}_{\bar{j}}^{(2)} - \dot{\hat{\mathbf{X}}}(\bar{j}, \bar{\alpha}) \right)^T \dot{\mathbf{V}}_z^{-1} \left(\dot{\mathbf{Z}}_{\bar{l}}^{(2)} - \dot{\hat{\mathbf{X}}}(\bar{l}, \bar{\alpha}) \right) \right], \quad (6)$$

where \mathbf{V}_z - covariance matrix of the conditional distribution $w\left\{z_{\bar{j}}^{(2)}\right\} | \mathbf{Z}^{(1)}, \bar{\alpha}$, $\dot{\hat{\mathbf{X}}}(\bar{j}, \bar{\alpha})$ - prediction found on the basis of observations $\mathbf{Z}^{(1)}$, which is the best estimate in the sense of estimation error variance minimum of a deformed image. The point above the matrices denotes their lexicographic representation. In the same work it is shown that in many cases the product $\dot{\hat{\mathbf{X}}}^T(\bar{\alpha}) \dot{\mathbf{V}}_z^{-1} \dot{\hat{\mathbf{X}}}(\bar{\alpha})$ can be considered to be independent of the deformation parameters $\bar{\alpha}$. Then the gradient of the goal function is determined by the relation

$$\nabla J(\bar{\alpha}, Z) = -\nabla \left[\dot{\hat{\mathbf{X}}}(\bar{\alpha}) \dot{\mathbf{V}}_z^{-1} \dot{\mathbf{Z}}^{(2)} \right]. \quad (7)$$

We should note that in the last case to find the optimal estimates of the parameters $\bar{\alpha}^*$ the maximization of the goal function is carried out. It requires performing of recurrent algorithm iterations not in the direction of the antigradient, but in the direction of the gradient which corresponds to minus in (7).

It is obvious, the expressions (6) and (7) can not be realized in systems of continuous image processing, because it requires great computational expenses. However, their simplification enables to obtain various realizable pseudogradients of the goal function. Let us consider some possible ways of such simplification. If we assume that the image insignificantly varies

from frame to frame (i.e. $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ are noisy realizations of the images \mathbf{X} and $\mathbf{X}(\bar{j}, \bar{\alpha})$), then there is no need to calculate the unwieldy covariance matrix \mathbf{V}_z of the conditional distribution $w\left\{\left\{z_j^{(2)}\right\}\middle|\mathbf{Z}^{(1)}, \bar{\alpha}\right\}$, because in this case $\mathbf{V}_z \approx \sigma_\theta^2 \delta_{j,\bar{j}}$, where σ_θ^2 - variance of additive noise according to the model of observations (5). However, in this case calculation of the optimal prediction $\hat{x}(\bar{j}, \bar{\alpha})$ requires matrix operations, which lead to very large computational expenses for large-sized images. We can obtain their reduction by substituting the optimal prediction of values of deformed frame for prediction at limited local region of image. We can achieve even more calculations reduction using various interpolations for the prediction. When performing the interpolations at the current iteration of the algorithm the estimates $\hat{\alpha}$ obtained at the preceding iteration are employed (Minkina et al., 2007). Then, to find the pseudogradient at the t -th iteration of the algorithm it is enough to use a local sample $Z_t = \left\{z_{j,t}^{(2)}, \tilde{z}_{j,t}^{(1)}\right\}$ of samples, where $z_{j,t}^{(2)}$ - samples of the deformed image $\mathbf{Z}^{(2)}$ contained in a local sample at the t -th iteration and $\tilde{z}_{j,t}^{(1)} = \tilde{z}^{(1)}(\bar{j}_t, \hat{\alpha}_{t-1})$ - brightness values from continuous image $\tilde{\mathbf{Z}}^{(1)}$ obtained from $\mathbf{Z}^{(1)}$ through the chosen interpolation; $\bar{j}_t \in \Omega_{j,t} \in \Omega_j$ - sample coordinates $z_{j,t}^{(2)}$ ($\Omega_{j,t}$ - plan of a local sample). The number of samples $\left\{z_{j,t}^{(2)}\right\}$ in Z_t will be called the local sample size and denoted with μ . Under these assumptions the pseudogradients obtained on the basis of relations (6) and (7) will become

$$\bar{\beta}_t = \sum_{\bar{j}_t \in \Omega_{j,t}} \frac{\partial \tilde{z}_{j,t}^{(1)}}{\partial \bar{\alpha}} \Delta_{\bar{j},t} \bigg|_{\bar{\alpha} = \hat{\alpha}_{t-1}}, \quad (8)$$

$$\bar{\beta}_t = - \sum_{\bar{j}_t \in \Omega_{j,t}} \frac{\partial \tilde{z}_{j,t}^{(1)}}{\partial \bar{\alpha}} z_{j,t}^{(2)} \bigg|_{\bar{\alpha} = \hat{\alpha}_{t-1}}, \quad (9)$$

where $\Delta_{\bar{j},t} = \tilde{z}_{j,t}^{(1)} - z_{j,t}^{(2)}$.

We should note that the pseudogradient (8) is used for solving the problem of interframe difference mean square minimization. In this case it will be the goal function

$$J(\bar{\alpha}, \mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}) = \frac{1}{M} \sum_{j \in \Omega} \left(z_j^{(2)} - \tilde{z}_j^{(1)} \right)^2,$$

where M - number of samples in the frame $\mathbf{Z}^{(2)}$.

The pseudogradient (9) corresponds to the problem of interframe correlation sample coefficient maximization:

$$J(\bar{\alpha}, \mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}) = \frac{\sum_{j \in \Omega} \left(z_j^{(2)} - z_m^{(2)} \right) \left(\tilde{z}_j^{(1)} - \tilde{x}_m^{(1)} \right)}{M \hat{\sigma}_{z_1} \hat{\sigma}_{z_2}},$$

where: $\hat{\sigma}_{z1}^2 = \frac{1}{M-1} \sum_{j \in \Omega} (z_j^{(2)} - z_m^{(2)})^2$ and $\hat{\sigma}_{z2}^2 = \frac{1}{M-1} \sum_{j \in \Omega} (\tilde{z}_j^{(1)} - \tilde{x}_m^{(1)})^2$ - variance estimates of the images $\mathbf{Z}^{(2)}$ and $\tilde{\mathbf{Z}}^{(1)}$; $z_{jm}^{(2)} = \frac{1}{M} \sum_{j \in \Omega} z_j^{(2)}$; $\tilde{x}_m^{(1)} = \frac{1}{M} \sum_{j \in \Omega} \tilde{z}_j^{(1)}$; $j \in \Omega_j$.

Thus, in practical problems of IIGDs estimation the basic goal functions can be the interframe difference mean square and the interframe correlation sample coefficient. We should note that the pseudogradient (9) in contrast to (8) is invariant to the total variability of samples brightness of the image $\mathbf{Z}^{(2)}$. The choice of interframe difference mean square as the goal function is expedient in absence of multiplicative distortions and noncentered interference in the observable image models.

The vector $\bar{\beta}_t = \bar{\phi}(\nabla Q(\bar{\alpha}_{t-1}, Z_t))$ can be chosen as a pseudogradient, where $\bar{\phi}(\cdot)$ - vector function of the same dimensionality as ∇Q . For example, the function $\bar{\phi}(\cdot)$ can be linear transformation with the positively determined matrix. At that, if errors with respect to the true gradient have symmetric distributions with reference to zero, then the condition of the pseudogradientness (2) holds for any odd function $\bar{\phi}(\cdot)$. In particular, very simple and at the same time well converging algorithms of the parameters estimation are obtained when choosing the following sign function as $\bar{\phi}(\cdot)$ (Korn & Korn, 1968)

$$\bar{\beta}_t = \text{sgn}(\nabla Q(\hat{\alpha}_{t-1}, Z_t)), \quad (10)$$

where $\text{sgn}(\nabla Q) = (\text{sgn}(\nabla Q_1), \dots, \text{sgn}(\nabla Q_m))^T$. When using the pseudogradient (10) and the diagonal gain matrix in PGA (3) the i -th component of the vector $\hat{\alpha}_t$ is different from the corresponding component of the vector $\hat{\alpha}_{t-1}$ by $\pm \lambda_{i,t}$, where $\pm \lambda_{i,t}$ - corresponding diagonal element of the gain matrix Λ_t ; $i = \overline{1, m}$. At that, PGA iterations are carried out at finite and a priori known number of directions of the space of the parameters to be estimated. If each component of the error (10) in relation to the true gradient takes positive and negative values with equal probabilities, then the pseudogradientness condition (2) is met. Let us note the algorithms that use the pseudogradients of type (10) have wide application in various problems requiring IIGDs estimation in the conditions of complex noise assemblage.

3. Pseudogradient algorithms for interframe geometrical deformations parameters estimation

3.1 Algorithms at given set of geometrical deformations model parameters

If a parameters set $\bar{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m)^T$ of possible IIGDs is known, then at chosen goal function the problem amounts to estimation of their values that are constant on the images $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$. For example, if for (3) the interframe difference mean square is chosen as a goal function and its pseudogradient is given by relations (8) and (10), then to estimate $\bar{\alpha}$ we accordingly obtain the following algorithms:

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} - \Lambda_t \left(\sum_{\bar{j} \in \Omega_{\bar{j},t}} \frac{\partial \tilde{z}_{\bar{j},t}^{(1)}}{\partial \alpha} \Delta_{\bar{j},t} \right)_{\bar{\alpha} = \hat{\alpha}_{t-1}} ; \quad (11)$$

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} - \Lambda_t \operatorname{sgn} \left(\sum_{\bar{j} \in \Omega_{\bar{j},t}} \frac{\partial \tilde{z}_{\bar{j},t}^{(1)}}{\partial \alpha} \Delta_{\bar{j},t} \right)_{\bar{\alpha} = \hat{\alpha}_{t-1}} . \quad (12)$$

Experimental study show it is expedient to extend the algorithms set (11)-(12) by adding two more ones

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} - \Lambda_t \left(\sum_{\bar{j} \in \Omega_{\bar{j},t}} \frac{\partial \tilde{z}_{\bar{j},t}^{(1)}}{\partial \alpha} \operatorname{sgn} \Delta_{\bar{j},t} \right)_{\bar{\alpha} = \hat{\alpha}_{t-1}} ; \quad (13)$$

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} - \Lambda_t \left(\sum_{\bar{j} \in \Omega_{\bar{j},t}} \Delta_{\bar{j},t} \operatorname{sgn} \frac{\partial \tilde{z}_{\bar{j},t}^{(1)}}{\partial \alpha} \right)_{\bar{\alpha} = \hat{\alpha}_{t-1}} . \quad (14)$$

In the algorithm (11) all the components of estimation increment vector depend on interframe differences $\Delta_{\bar{j},t}$, $\bar{j} \in \Omega_{\bar{j},t} \in \Omega_{\bar{j}}$. It determines higher estimation convergence speed compared to other given algorithms. However, at finite number of iterations the precision of (11) is often lower because we can not always attain little $\Delta_{\bar{j},t}$ and then, the estimates variation steps can be too large. In the algorithm (13) only signs of $\Delta_{\bar{j},t}$ are used. It is preferable to apply it when we want to avoid excessive influence of modulo large values $\Delta_{\bar{j},t}$, for example, in presence of infrequent but intensive impulse interference on image. The algorithm (12) is even more immune to interference, but it may not operate well in the neighborhood of zero values $\Delta_{\bar{j},t}$. If high accuracy of estimation is attained at some iteration, then the next step can be taken in the direction backward from optimal values of the parameters. To eliminate this disadvantage we can employ a sign function with expanded zero:

$$\operatorname{sgn}_\varepsilon(x) = \begin{cases} -1, & x < -\varepsilon, \\ 0, & |x| \leq \varepsilon, \\ 1, & x > \varepsilon, \end{cases} \quad \varepsilon > 0.$$

In the algorithm (14) the increase of convergence speed at large values $\Delta_{\bar{j},t}$ combines with immunity to derivatives estimation errors. As a result this algorithm is more resistant to interference than the algorithm (11).

When choosing the interframe correlation sample coefficient as a goal function the properties of the corresponding algorithms are close to the properties of the algorithms (11)-(14). The advantage in this case is in high immunity to additive noise and close to linear brightness distortions. Among disadvantages are larger computational expenses (determined by large size of the local sample) and also high sensibility to local extremums of the goal function.

It is necessary to note that the convergence speed of the estimates, formed by PGAs, is higher if the sequence of local samples, which is the basis for parameters estimation, is not correlated. To reduce correlation of the observations sequence it is expedient to choose random coordinates of samples of the local sample.

The algorithms (11)–(14) demonstrated high efficiency when estimating interframe deformations of simulated and real images. In particular, for two-dimensional images of size 64×64 pixels formed by means of the wave model (Krashennikov, 2003) shifted by several steps of sample grid and turned at any angle, the shift was estimated with error variance of about $2 \cdot 10^{-4}$ steps of the sample grid and rotation – $5 \cdot 10^{-5}$ radians. Let us give the results of analytical calculation of the probability $P(\Delta)$ of parallel shift estimate \hat{h}_1 error spillover of the given interval $\Delta = [-a, a]$. The calculation was carried out on the basis of the accuracy analysis method of PGAs estimates at finite number of iterations (Tashlinskii & Tikhonov, 2001) for the PGA (12) and the following parameters: the images with a Gaussian brightness distribution and the autocorrelation function with correlation radius equal to 5; the signal variance-to-noise variance ratio $g = 100$; local sample size $\mu = 10$; initial error of the shift is $\bar{h}_0 = (5, 4)^T$; $a = 1.0, 0.3$ и 0.1 (here, by correlation radius we imply the distance in steps of the sample grid when the autocorrelation function of the image is equal to 0.5). The value of estimate shift increment in one case was chosen to be constant $\lambda_{1,t} = \lambda_{2,t} = \lambda = \text{const} = 0.1$ and in the other – falling off according to the rule $\lambda_{1,t} = \lambda_{2,t} = \text{var} = \frac{1}{(10 + 0.01t)}$. The plots of the probability $P(\Delta)$ are shown in Fig. 1. It

follows from the analysis of the plots that at constant λ the balance between tendency of the estimate to the true value and error, caused by λ , comes at a certain iteration. The further increase of iterations number does not lead to estimates improvement. It enables to find the minimum number of iterations that is necessary to achieve the highest possible accuracy of parameter estimation.

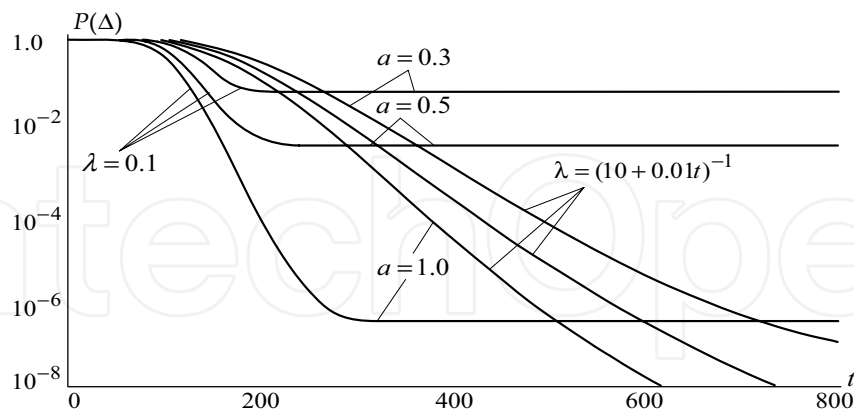


Fig. 1. Probability of estimate error spillover of the confidence interval

For the above-mentioned characteristics of images and PGA in Fig. 2 the shift h_1 estimate error probability distribution change at both constant (Fig. 2,a) and varying (Fig. 2,b) value

of the shift estimate increment is shown. For clearness the distributions are presented only for 12 iterations from the range of 10 to 100. The interframe difference mean square at local sample size $\mu=4$ was used as a goal function. The analysis of the plots shows that at constant shift increment step the process of probability distributions forming stabilizes after about 500th iteration. At varying shift increment step the process of probability distribution forming does not have an equilibrium state and the estimate variance theoretically permanently decreases.

Let us note that at known set of IIGD parameters the algorithms (11)-(14) have shown a good performance at automated search of local fragments on images.

3.2 Algorithms at unknown set of geometrical deformations model parameters

If the form of IIGD is not given then we can specify a certain sample grid deformations model

$$\bar{\alpha}(\bar{j}) = (\bar{j} + \bar{h}_j) = (j_1+h_1, j_2+h_2, \dots, j_n+h_n), \tag{15}$$

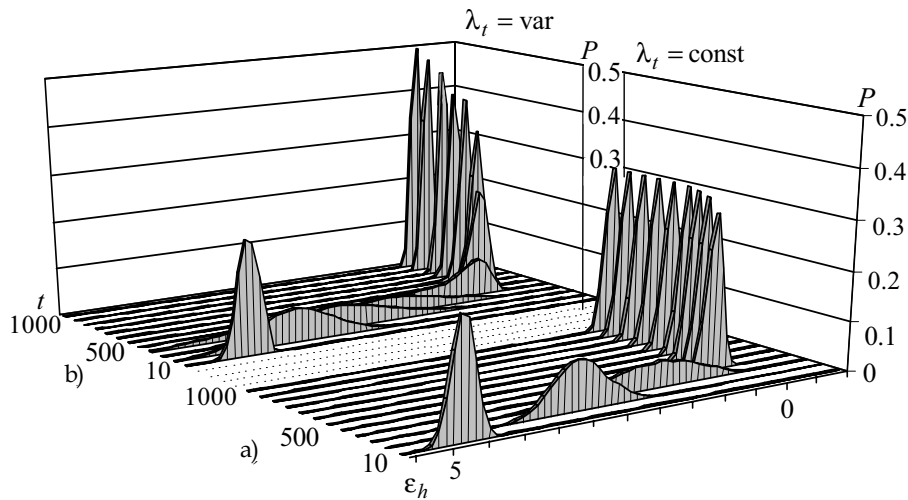


Fig. 2. Probability distributions of the shift estimate error at constant and varying estimate increment steps

considering its parameters to be varying, where $j_k \in \bar{j} \in \Omega_j, j_k = \overline{1, N_k}, k = \overline{1, n}; \Omega_j - n$ -dimensional rectangular grid. Then, the algorithm (3) can be written in the form

$$\hat{h}_t = \hat{h}_{t-1} - \Lambda_t \bar{\beta}_t(Z_t, \bar{j}_t, \hat{h}_{t-1}) \tag{16}$$

In this case to ensure variability of the estimates the components of the gain matrix Λ_t in (16) have to be bounded below. Then, assuming $\Lambda_t = \Lambda$ and choosing the interframe difference mean square as a goal function, we come to the algorithm

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} - \Lambda_t \left(\sum_{\bar{j}_t \in \Omega_{\bar{j},t}} \frac{\partial \tilde{z}^{(1)}(\bar{j}_t, \bar{h})}{\partial \bar{h}} \left(\tilde{z}^{(1)}(\bar{j}_t, \bar{h}) - z_{\bar{j},t}^{(2)} \right) \right)_{\bar{h} = \hat{h}_{t-1}}. \quad (17)$$

In work (Tashlinskii, 2000) it is shown that if point-to-point correlation is within the limits $0.8 \div 0.99$ and signal variance-to-noise variance ratio is more than 50, then for many problems of IIGDs estimation it is enough to choose $\mu = 1$ in (17). In spite of the simplicity the algorithm (17) is rather effective at high correlation of the shifts \bar{h}_j in the direction of image scanning and relatively minor interframe brightness distortions. However, only one-dimensional filtering of deformation parameters that does not take into account interrow and interframe correlation is carried out in it. The simplest way to take into account this correlation can be in refinement of the estimates, obtained at one pass through the images. For that we can perform repeated passes at lesser values λ in the backward and other directions (along columns, diagonals) during which the obtained estimates are corrected. The algorithms of the type (17) have shown a good performance at automatic combination of image fragments that have reciprocal spatial and amplitude deformations (the problem of image «pasting»). This problem often occurs when forming a unified image from a sequence of frames, obtained from a mobile object, that have small common regions on the adjacent frames. When solving the mentioned problem it is required as a rule to ensure continuity of spatial and brightness characteristics on the resulting image. The considered algorithms enable to do it. To illustrate it in Fig.3 an example of «pasting» of images is presented, where a) and b) - are the images of size 100×160 elements to be connected, having parallel shift $\bar{h} = (-0.5, -0.2)^T$, rotation angle $\varphi = 0.5^\circ$ and scale coefficient $k = 1.2$; c) - the result of «pasting» before spatial correction using the obtained estimates; d) - the result of «pasting» after correction. The estimates accounted for $\hat{\bar{h}} = (-0.676, -0.13)$, $\hat{\varphi} = 0.509^\circ$, $\hat{k} = 1.18$.



Fig. 3. An example of automated «pasting» of images

Basically, the model (15) enables to define any IIGDs. However, if at the chosen order of the images pass the shifts \bar{h}_j change rapidly then their estimation by means of PGA (17) is difficult. It is due to the fact that when variation speed of the shifts increases in the direction of the image pass it is necessary to increase steps $(\Lambda_t, \bar{\beta}_t)$ of PGA (17). The last, in its turn, leads to estimation error increase. In this situation we can not improve the estimates even by repeated passes. The mentioned contradiction can be solved due to the usage at each following pass information about estimates, obtained at the preceding passes. Let us

consider the algorithm that forms the deformation matrix \mathbf{H}_l of size $N_1 \times N_2 \times \dots \times N_n$ (where n - image dimensionality) as an example of such an approach. This matrix contains shifts estimates $\hat{h}_j^{(l)}$ of all image pixels, corresponding to sample grid nodes of the frame $\mathbf{Z}^{(1)}$ after the l -th pass. The method of estimates forming can be various, for example, it can be determined by available conceptions regarding physical nature of geometrical deformations. Assuming, that all elements of the matrix \mathbf{H}_{l-1} at the $(l-1)$ -th pass have been determined we can write

$$\mathbf{H}_l = \left\| \hat{h}_j^{(l)} \right\| = f \left(\hat{h}_j^{(l-1)}, \{ \hat{\alpha}_t^{(l)} \} \right),$$

$$\hat{\alpha}_t^{(l)} = \hat{\alpha}_{t-1}^{(l)} - \mathbf{\Lambda}_{l,t} \bar{\beta} \left(z_{\bar{j},t}^{(2)}, \tilde{z}^{(1)}(\bar{j}_t + h_{\bar{j},t}^{(l-1)}), \hat{\alpha}_{t-1}^{(l)} \right),$$

where $\hat{\alpha}_t^{(l)} = (\hat{\alpha}_{1,t}^{(l)}, \hat{\alpha}_{2,t}^{(l)}, \dots, \hat{\alpha}_{m,t}^{(l)})^T$; $\bar{j}_t \in \Omega_t^{(l)} \in \Omega$; $\mathbf{\Lambda}_{l,t} = \left\| \begin{matrix} \lambda_{1,t}^{(l)} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \lambda_{n,t}^{(l)} \end{matrix} \right\|$; $z_{\bar{j},t}^{(2)}$,

$\tilde{z}^{(1)}(\bar{j}_t + h_{\bar{j},t}^{(l-1)}) \in Z_t^{(l)}$; $Z_t^{(l)}$ - local sample of the goal function for the pseudogradient estimation at the t -th iteration of the l -th pass of the algorithm; $l = \overline{1, L}$ - pass number; L - given number of passes. Choosing various goal functions and pseudogradients we can obtain different algorithms. For example, if we choose the pseudogradient (10) for the interframe difference mean square and assume $\lambda_{ii}^{(l)} = \lambda_i$; $\hat{h}_j^{(l)} = \hat{h}_j^{(l-1)} + \hat{\alpha}_j^{(l)}$, where $\hat{\alpha}_j^{(l)}$ - current estimate $\hat{\alpha}^{(l)}$ at the point \bar{j} then we obtain

$$\mathbf{H}_l = \left\| \hat{h}_j^{(l)} \right\|, \hat{h}_j^{(l)} = \hat{h}_j^{(l-1)} + \hat{\alpha}_j^{(l)}, l = \overline{1, L},$$

$$\hat{\alpha}_t^{(l)} = \hat{\alpha}_{t-1}^{(l)} - \lambda_l \operatorname{sgn} \left(\sum_{\bar{j}_t \in \Omega_{j,t}^{(l)}} \frac{\tilde{z}^{(1)}(\bar{j}_t + \hat{h}_{\bar{j},t}^{(l-1)} + \hat{\alpha}_{t-1}^{(l)})}{\partial \hat{\alpha}} \left(\tilde{z}^{(1)}(\bar{j}_t + \hat{h}_{\bar{j},t}^{(l-1)} + \hat{\alpha}_{t-1}^{(l)}) - z_{\bar{j},t}^{(2)} \right) \right)_{\hat{\alpha} = \hat{\alpha}_{t-1}^{(l)}}. \quad (18)$$

If IIGDs with known parameters set (for example, the common for all image shift, rotation etc.) are present along with deformations of unknown type, then values of these parameters can be estimated, specified and taken into account when forming elements of matrix \mathbf{H}_l at each algorithm pass.

Another advantage of the algorithms of the type (18) is that they enable to estimate IIGDs that do not satisfy the continuity requirement.

An example of such estimation is shown in Fig. 4, where a) and b) - images of size 256×256 elements having reciprocal shifts ($h_1 = 1.5$, $h_2 = 3.5$), besides in the lower image the continuity of geometrical deformations is violated (5 rows are missing) and the fragment is

absent; c) – result of parameter h_1 estimation at $L=40$ and $\mu=1$. The sudden change corresponding to the break of the parameter h_1 is well visible. Besides in the region, corresponding to the absent fragment, the estimates have significantly differing statistical properties and due to which these regions can be easily identified.

One of the disadvantages of PGA at IIGD parameters estimation is in a relatively minor definition domain of parameters, where effective convergence of estimates is ensured (not large operating range). The size of this region is determined by sample correlation that can appear in the local sample Z_t . The situation is also complicated by the fact that in real images samples of reference and deformed images taken rather far from each other are almost non-correlated. At operating with real images another serious disadvantage of PGAs is in the possibility of the estimates to converge to points of false extremums of the goal function in the parameter space.

3.3 Algorithms with adaptive forming of local sample size

In view of random character of images and noise, the estimate of the goal function is not unimodal and besides the global extremum it also contains false local extremums. The local extremums appear because of correlatedness of separate extensive objects on the image and are exposed if a portion of samples of the local sample appears into these regions, i.e. they are caused by limited size of the local sample. As local sample increases or changes the probability of this effect appearance sharply decreases. As a result it is reasonable to verify on the goal function local extremums attributes at each iteration of estimation and if any, to increase sample size or change it. Here, the sample size μ becomes an adaptive value.

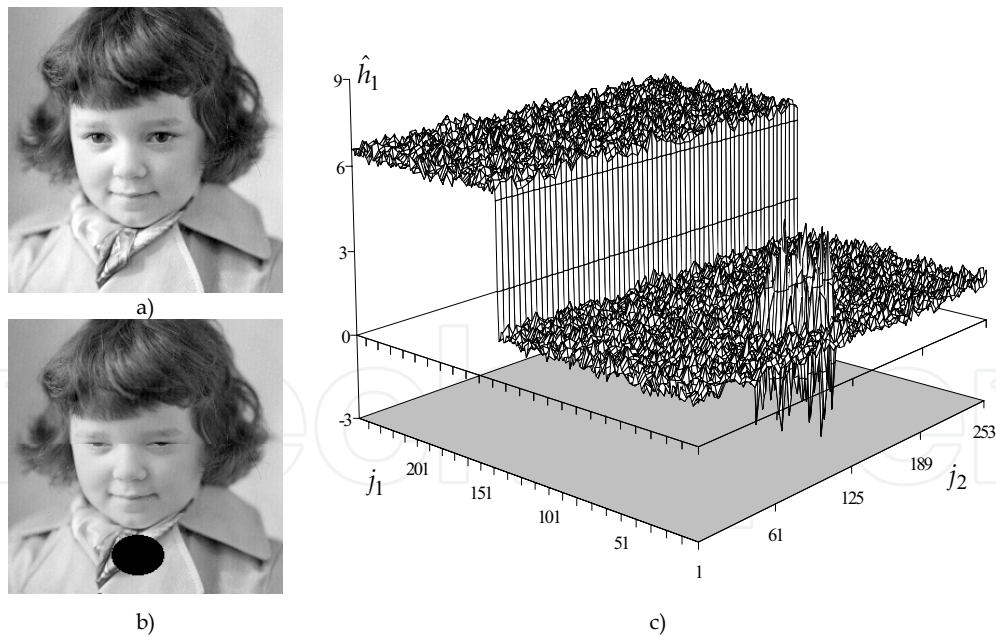


Fig. 4. An example of geometrical deformations estimation that does not satisfy the continuity requirement

Let us consider only one example of construction of IIGD parameter estimation PGA, where μ is adjusted automatically during the procedure performing at each iteration. A current iteration of parameter estimation is carried out when a certain condition is met. If at minimal μ for the current iteration the condition is not met, then μ increases step-by-step until the condition is fulfilled. So for the local sample, formed at the given iteration, its size is to be minimum to meet the condition of the iteration realization. For definiteness, we assume that the PGA (3) with the pseudogradient (10) and the diagonal gain matrix Λ_t is used. Then, the estimates for the i -th parameter are formed according to the following rule:

$$\hat{\alpha}_{i,t+1} = \hat{\alpha}_{i,t} \mp \lambda_{t+1} \operatorname{sgn}(\beta_{i,t+1}(\hat{\alpha}_t, Z_t)), \quad t = \overline{1, T}, \quad i = \overline{1, m},$$

where the signs «-» and «+» correspond to the problems of minimization and maximization of the goal function.

Given a certain initial sample size $\mu_{t\min}$, whose minimum value at interframe correlation sample coefficient maximization must be not less than 2 and at the interframe difference mean square minimization – is equal to 1. To find numerical values in the conditions of the iteration realization we use the goal function estimates obtained at the corresponding μ . Let us denote the goal function estimate with $q_t(\mu_k)$ which is calculated at the t -th iteration on the samples $z_{j,t}^{(2)}$ and $\tilde{z}_{j,t}^{(1)}$ of the local sample of size μ_k and as $q_t^{\pm}(\mu_k)$ – goal function estimate at the t -th iteration at the same μ_k calculated on samples $z_{j,t}^{(2)}$ and $\tilde{z}_{\pm}^{(1)}(\bar{j}_t, \hat{\alpha}_{1,t-1}, \dots, \hat{\alpha}_{i,t-1} \pm \Delta_{\alpha_i}, \dots, \hat{\alpha}_{m,t-1})$, i.e. when a certain increment $\Delta_{\alpha_i} > 0$ is specified $\bar{j}_t \in \Omega_t \in \Omega$ for the parameter α_i estimate.

Let us fulfill the following condition of the iteration realization: the iteration of finding the current estimate $\hat{\alpha}_{i,t+1}$ is not performed and $\mu_{t\min}$ increases by 1 (a new pair of samples $z_{j,t}^{(2)}$ and $\tilde{z}_{j,t}^{(1)}$ is added to the local sample) in two cases:

- if at the current t -th iteration the estimate $q_t(\mu_{t\min})$ for the local sample size $\mu_{t\min}$ is «better» than both the values $q_t^+(\mu_{t\min})$ and $q_t^-(\mu_{t\min})$;
- if at the current t -th iteration the estimate $q_t(\mu_{t\min})$ for the local sample size $\mu_{t\min}$ is «worse» than the values $q_t^+(\mu_{t\min})$ and $q_t^-(\mu_{t\min})$ but at that $q_t^+(\mu_{t\min}) = q_t^-(\mu_{t\min})$.

After that the sample size increases by one ($\mu_{t\min} + 1$) and the above-mentioned conditions are verified again. If one of them is fulfilled, then μ increases by one again and so on right up to a certain value μ_{\max} . If μ_{\max} is attained, the following iteration of the parameter α_i estimation is performed. If at the current μ the conditions are not met, then the next $(t+1)$ -th iteration of the estimate $\hat{\alpha}_{i,t+1}$ forming for the parameter α_i is carried out.

In particular, when maximizing the goal function we can write the procedure of parameter α_i estimation in the following form

$$\begin{aligned}
 \hat{\alpha}_{i,t+1} &= \hat{\alpha}_{i,t} - \lambda_{t+1} \beta_{i,t+1}(\mu_t); \\
 \mu_t &= \begin{cases} \mu_k + 1, & \text{if } (q_t(\mu_k) < q_t^+(\mu_k) \wedge q_t^-(\mu_k)) \vee (q_t(\mu_k) > q_t^+(\mu_k) = q_t^-(\mu_k)), \\ \mu_{\max}, & \text{if } \mu_k = \mu_{\max}, \\ \mu_k, & \text{in the other case;} \end{cases} \\
 \beta_{i,t+1} &= \begin{cases} 1, & \text{if } q_t^+(\mu_t) < q_t^-(\mu_t) \wedge q_t(\mu_t) > \min(q_t^+(\mu_t), q_t^-(\mu_t)); \\ 0, & \text{if } q_t^+(\mu_t) = q_t^-(\mu_t); \\ -1, & \text{if } q_t^+(\mu_t) > q_t^-(\mu_t) \wedge q_t(\mu_t) > \min(q_t^+(\mu_t), q_t^-(\mu_t)). \end{cases}
 \end{aligned} \tag{19}$$

Let us note that as t increases the value $\mu_{t\min}$ varies according to a certain prescribed rule that is defined by the problem to be solved, in particular, in the simplest case $\mu_{t\min} = \text{const}$.

In Fig. 5 experimental results obtained for the algorithm (19) realization are presented. In the experiment a real image of optical range with correlation radius equal to 5 steps on the sample grid was used. A parameter to be estimated was the parallel shift $\bar{h} = (10, 0.5)^T$. The shifted image was additionally noised by an independent centered Gaussian noise. The dependencies of μ_t as a function of the number of iterations, averaged on 50 realizations, are shown in Fig. 5,a. Here, the dependence 1 corresponds to signal variance-to-noise variance ratio $g=100$ and the dependence 2 - $g=50$. It is seen that for great errors of the estimate the sample size is small (for $g=100$ at $t=10$ it is equal to about 2 and at $t=500$ - to about 2.3) and it increases monotonously on average as the number of iterations increases (attaining about 6 at $g=100$ and $t=2000$). In Fig. 5,b plots of the estimation error ε_n versus the number of iterations are presented, where curve 1 corresponds to the results, obtained for the adaptive forming of μ_t , and curve 2 - at constant $\mu = \mu_m$, where

$$\mu_m = \sum_{k=1}^{2000} \mu_k - \text{average sample size for } t \text{ varying from 1 to 2000. The results are averaged on}$$

250 realizations. It is obvious at small number of iterations there is a loss in estimation accuracy (at the 100th iteration it is about 5 per cent). It can be explained by high speed of the algorithm convergence with constant μ at the initial stage of estimation (due to a greater average of μ). However, at equal computational expenses (to the 2000th iteration) a gain in accuracy of about 2.4 times as large is guaranteed.

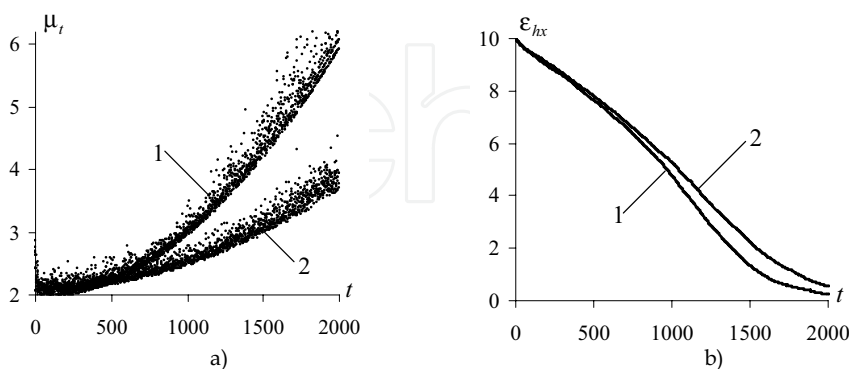


Fig. 5. An example of dependence of local sample size and estimate error versus the number of iterations

Thus, due to the fact that the proposed PGA with adaptive size of local sample facilitates the estimates vector recovery from local extremums of the goal function, it enables to significantly increase parameter estimate convergence speed in comparison with PGA with constant sample size at equal computational expenses. One more example of construction of PGA with adjustable local sample size is presented in work (Tashlinskii, 2003).

Let us note that if the problem of IIGD parameter estimation is a part of the problem of identification with a decision rule based on goal function values, then in order to achieve the required confidence probability of identification it may require large sample size μ_1 that is not justified in the process of the estimates $\hat{\alpha}$ convergence. In this case it is expedient to use adaptive adjustment of sample size and we can choose $\mu_{\max} = \mu_1$ as its maximum. Then the attainment of μ_{\max} will simultaneously mean the identification problem solution. An example of such a problem can be search of a fragment location on the reference image where a criterion of correspondence is in excess of a certain correlation sample coefficient value between the fragment and the reference image.

4. Structural optimization of pseudogradient algorithms

In practical problems of IIGDs estimation by means of pseudogradient procedures the required accuracy of parameter $\bar{\alpha}$ estimates is not obtained in the whole domain $\Omega_{\bar{\alpha}}$ of possible values $\bar{\alpha}$, but only in a certain subdomain bounded by an operating range of the procedures. This leads to the necessity of decomposing $\Omega_{\bar{\alpha}}$ into N subdomains $\Omega_{\bar{\alpha}}^{(i)}(\hat{\alpha}_{0,i})$, $i = \overline{1, N}$ corresponding to the operating range of the employed procedures where $\hat{\alpha}_{0,i}$ - an initial approximation of the parameters for the procedure operating in the i -th subdomain. Let the procedure, operating in the subdomain which contains the sought vector $\bar{\alpha}_v$ of parameters, be called a V-procedure (from veritas - true) and the corresponding subdomain - V-subdomain. Subdomains that do not include $\bar{\alpha}_v$ are called P-subdomains (from pseudo - false) and the corresponding procedures - P-procedures. As the result of all these procedures operation N vectors $\hat{\alpha}_i$ of IIGD parameter estimates are formed and the problem of determination of a V-subdomain among these estimates with required accuracy where the goal function attains its extremum arises.

4.1 The principle of pseudogradient procedure set control

In the problems of IIGDs estimation the number of subdomains can run up to dozens of thousands. Thus, bringing all the procedures operating in subdomains to the number of iterations that ensures the necessary accuracy of estimation requires great computational expenses. At such an approach the choice of the V-subdomain requires additional calculations. To reduce computational expenses the following principle of structural optimization can be used (pseudogradient procedures set control). At each step of the algorithm the priority of the current iteration realization is given to the procedure, having the least value of a certain penalty function ψ characterizing the level of the priority (Tashlinskii, 2006). Here, by «algorithm step» we imply a set of operations that includes performing of the current iteration by the procedure with the least penalty function, finding

a new value of the penalty function and obtaining a procedure with the least penalty function.

A characteristic property of such an approach is in the necessity to compare the «penalties» of the procedures which have performed different number of iterations. Studies have shown that when minimizing the goal function the following penalty functions satisfies such requirements

$$\psi_t^{(i)} = \sum_{k=1}^t (q_k^{(i)} - q_{\text{inf}}), \quad i = \overline{1, N},$$

where $q_k^{(i)}$ - goal function estimate at the k -th iteration; $q_{\text{inf}} \leq \inf\{q_k^{(i)}\}$ - value which is less than the lower bound of the possible estimates set of the goal function. If the goal function is to be maximized then

$$\psi_t^{(i)} = \sum_{k=1}^t (q_{\text{sup}} - q_k^{(i)}), \quad i = \overline{1, N},$$

where $q_{\text{sup}} \geq \sup\{q_k^{(i)}\}$.

In the process of parameters estimates convergence to the optimal values the probabilistic properties of the goal function estimates are changing, which leads to the change of the probabilistic properties of the penalty function ψ . So when studying properties of ψ it is necessary to know its probability distribution density $w_t(\psi)$ at each iteration of estimation. At that, $w_t(\psi)$ depends on the local sample Z_t i.e. a rate of the correspondence (similarity) of the sets $\{z_{j,t}^{(1)}\}$ and $\{z_{j,t}^{(2)}\}$ involved in the local sample. It is expedient to use correlation sample coefficient ρ as a value characterizing this correspondence. For isotropic images ρ is a one-dimensional characteristic for any number of parameters to be estimated which simplifies calculations. Then for probability distributions of the penalty function increment $\Delta\psi$ at the t -th iteration we can write

$$w_t(\Delta\psi) = \int_{-1}^1 w_t(\Delta\psi|\rho)w(\rho)d\rho, \quad (20)$$

where $w_t(\Delta\psi|\rho)$ - conditional density of increment; $w(\rho)$ - probability distribution density of the correlation coefficient. Let us note that for V-procedures $w_t(\Delta\psi|\rho)$ depends on the iteration number because ρ increases as the estimates vector $\hat{\alpha}$ converges to the optimal values.

Without loss of generality, we can assume that ψ takes only positive values. Then, for calculation of the distribution density of ψ at the t -th iteration we can obtain the recurrent expression

$$w_t(\psi) = \int_{0-1}^{\infty} \int_{-1}^1 w_{t-1}(\psi - \Delta\psi_t)w_t(\Delta\psi|\rho)w(\rho) d\Delta\psi d\rho. \quad (21)$$

For the P-procedures $w_t(\Delta\psi|\rho)$ does not depend on the iteration number. Then,

$$w_t(\psi) = \int_0^{\infty} w_{t-1}(\psi - \Delta\psi_t)w(\Delta\psi)d\Delta\psi, \quad (22)$$

where $w(\Delta\psi) = w_t(\Delta\psi|\rho=0)$.

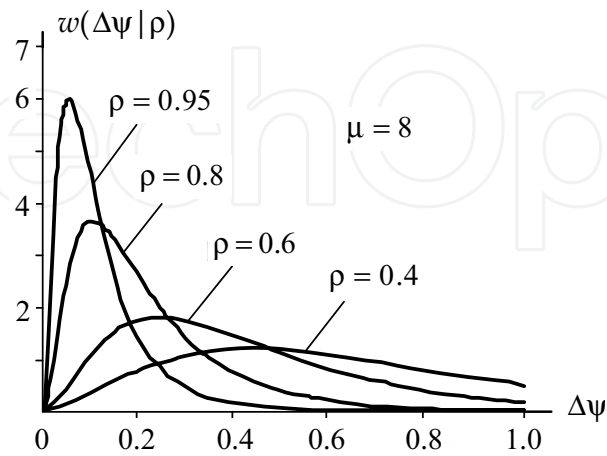
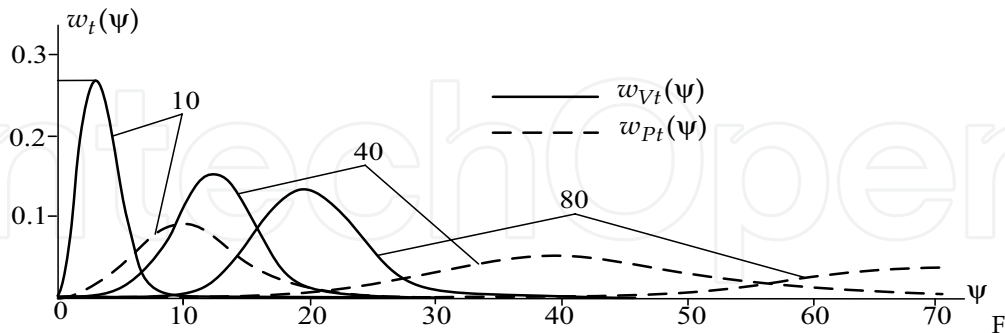


Fig. 6. Conditional distribution density of the penalty function increment

The expressions (20)–(22) enable to easily find the penalty function distribution density for the goal functions obtained in the second part. As an example, in Fig. 6 curves $w(\Delta\psi|\rho)$ of the increment $\Delta\psi$ of the interframe difference mean square at $\rho = 0.4, 0.6, 0.8, 0.95$ are presented. Other parameters of calculation were the following: the images $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ are Gaussian with correlation radius equal to 5 steps of the sample grid; the signal/noise ratio $g = 100$; the local sample size $\mu = 4$. In Fig. 7 the probability distribution densities of the interframe correlation sample coefficient for V-procedure ($w_{V_t}(\psi)$) and P-procedure ($w_{P_t}(\psi)$) at the number of iterations 10, 40 and 80 and $\mu = 7$ are presented. From the plot it is seen that the area of intersection between $w_{V_t}(\psi)$ and $w_{P_t}(\psi)$ decreases sharply as the number of iterations increases, which facilitates reliable separation of the procedures of P- and V-type.



ig. 7. Penalty function distribution density at the interframe correlation sample coefficient

4.2 Testing of the hypothesis about goal function extremum absence in the parameter definition domain

If the existence of V-subdomains in the IIGD definitional domain to be studied is not known a priori, then the problem of testing of the hypothesis that there is no V-subdomain among N studied subdomains appears. Let us consider the possibility of construction of a simple criterion for testing of such a hypothesis. In doing so, we use the circumstance that when using the principle of the structural optimization, the basic feature of the V-subdomain is in the number of iterations performed by the procedure that operates in it. If among the procedures to be studied there is a procedure corresponding to the V-subdomain, then the number of iterations performed by it is, as a rule, more than the number of iterations performed by the procedures, which operate in P-subdomains. So the V-procedure attains the given number of iterations on average faster than P-procedure. In absence of a V-subdomain in the domain to be studied all the procedures carry out, on average, equal number of iterations and the leading procedure attains the threshold number of iterations for a larger number of algorithm steps. Thus, the total number M of iterations performed by all procedures, the number of steps of the algorithm before the leading procedure attains a certain threshold value T_M can be chosen as a numerical value of the criterion of acceptance of the hypothesis about absence of the V-subdomains. In doing so, the statistical criterion of the hypothesis testing becomes very simple: if during $M = M_c$ steps of the algorithm none of procedures has reached the T_M -th iteration then there is no subdomain of V-type among the analyzed subdomains. The choice of values T_M and M_c enables to obtain the necessary error probabilities of the first and the second kind.

The value T_M in a number of cases can be defined by the problem of the investigation following the hypothesis testing, for example, in the problem of fragment search on a large image – by the accuracy of its location parameters determining. In this case it is expedient to find the number of the algorithm steps M_{c1} or M_{c2} ensuring the required error probability of the first or the second kind basing on the preassigned value T_M . Let us note that a certain value of the error probability of the second kind corresponds to each value M_{c1} assuring a given error probability of the first kind and vice versa. We can specify only one probability whereas the second will be determined through the given probability value and the algorithm parameters. Simultaneous meeting of both the conditions (Here,, one of them will be limited) can be obtained by a choice of a value T_M . If T_M and the algorithms parameters are given, then we can find the required number of steps M_{c1} for the error probability of the first kind and M_{c2} – for the error probability of the second kind and use the biggest of them in the algorithm as the threshold value. In the process of the algorithm realization, depending on the number of iterations t_L performed by the leading procedure and the total number M of algorithm steps, several various situations are possible that characterize veracity of the criterion of testing of the hypothesis about absence of a V-subdomain in the domain to be studied. Possible variants are presented in Tab.

t_L	M	Conclusion
$t_L < T_M$	$M < M_{c1} \wedge M < M_{c2}$	The error probabilities of the first and the second kind exceed the given values $P^{(1)}$ and $P^{(2)}$.
$t_L < T_M$	$M = \min \{M_{c1}, M_{c2}\}$	If $M_{c1} < M_{c2}$ then the error probability of the first does not exceed the value $P^{(1)}$; if $M_{c1} > M_{c2}$, then the error probability of the second does not exceed the value $P^{(2)}$.
$t_L < T_M$	$M = \max \{M_{c1}, M_{c2}\}$	The error probabilities of the first and the second kind do not exceed the given values $P^{(1)}$ and $P^{(2)}$.
$t_L \geq T_M$	$M < M_{c1} \wedge M < M_{c2}$	The error probabilities of the first kind exceeds the given values $P^{(1)}$ and $P^{(2)}$.

Table. Veracity of the criterion of testing of the hypothesis about absence of the V-subdomains in the parameters domain to be studied

Let us consider the veracity of the proposed criterion. First let us find the number of algorithm steps M_{c2} guaranteeing the hypothesis testing with the specified error probability of the second kind $P^{(2)}$. This error may appear when two conditions are simultaneously satisfied:

- the V-procedure has not reached the T_M -th iteration (we will consider the probability of this condition meeting to be $P_V^{(2)}$);
- none of P-procedures has reached the T_M -th iteration iteration (we will consider the probability of this condition meeting to be $P_p^{(2)}$).

Then assuming $P_V^{(2)}$ and $P_p^{(2)}$ to be independent we obtain

$$P^{(2)} = P_V^{(2)} P_p^{(2)}.$$

To find the probabilities $P_V^{(2)}$ and $P_p^{(2)}$ it is necessary to know the discrete distributions of the number of iterations for V- and P-procedures at the given number M of algorithm steps. Let us denote the discrete distribution of the number of iterations for the V-procedure as P_{Vt} , $t = \overline{1, T_M}$, where P_{Vt} - probability of the event that the V-procedure has performed t iterations at M algorithm steps. Then, the probabilities of the event that the V-procedure will not reach the T_M -th iteration are

$$P_V^{(2)} = \sum_{t=1}^{T_M-1} P_{Vt}.$$

Accordingly, let us denote the discrete distribution of the number of iterations for the P-procedure as P_{P_t} , $t = \overline{1, T_M}$, where P_{P_t} - probability of the event that the P-procedure has performed t iterations at M algorithm steps. Here, the probability of the event that none of the P-procedures reaches the T_M -th iteration is

$$P_P^{(2)} = \left(\sum_{t=1}^{T_M-1} P_{P_t} \right)^{N-1}.$$

Let us note that P_{V_t} and P_{P_t} depend on the total number M of algorithm steps accordingly various M correspond to different distribution density $w_{T_M}(\psi)$ of the penalty function

If the number of algorithm steps M and the probability $P^{(2)}$ are given, then we can find the corresponding number of iterations T_M from the following considerations. On average, the

V-procedure fulfills $\sum_{t=1}^{T_M-1} tP_{V_t}$ iterations at M algorithm steps and each of P-procedures - $\sum_{t=1}^{T_M-1} tP_{P_t}$ iterations. Then, we can obtain T_M from the condition

$$\sum_{t=1}^{T_M-1} tP_{P_t} + (N-1) \sum_{t=1}^{T_M-1} tP_{V_t} = M. \quad (23)$$

The condition (23) can be simplified if we take into account the fact that the V-procedure in comparison with the P-procedure guarantees significantly high convergence speed and will fulfill the number of iterations close to T_M at M steps. Then,

$$T_M + (N-1) \sum_{t=1}^{T_M-1} tP_{V_t} \approx M.$$

Having determined T_M it is easy to find probabilities P_{V_t} and P_{P_t} :

$$P_{P_t} = \int_0^{\infty} w_{T_M}(\psi)(1 - F_{P_t}(\psi))d\psi - \sum_{i=1}^{t-1} P_{P_i},$$

$$P_{V_t} = \int_0^{\infty} w_{T_M}(\psi)(1 - F_{V_t}(\psi))d\psi - \sum_{i=1}^{t-1} P_{V_i},$$

where $F_{P_t}(\psi) = \int_0^{\psi} w_{P_t}(x)dx$ и $F_{V_t}(\psi) = \int_0^{\psi} w_{V_t}(x)dx$ - distribution functions of ψ at the t -th iteration for P- and V-procedures; $w_{T_M}(\psi)$ - penalty function distribution density at the T_M -th iteration defined by the procedures of both V-type and P-type. At one V-procedure

$$w_{T_M}(\psi) = w_{V_{T_M}}(\psi)(1 - F_{P_{T_M}}(\psi))^{N-1} + w_{P_{T_M}}^{(N-1)}(\psi)(1 - F_{V_{T_M}}(\psi)), \quad (24)$$

where $w_{P_{T_M}}^{(N-1)}(\psi)$ - distribution density of the minimum of $(N-1)$ penalty functions values of P-procedures at the T_M -th iteration.

Thus,

$$P^{(2)} = \sum_{t=1}^{T_M-1} P_{Vt} \left(\sum_{i=1}^{T_M-1} P_{Pi} \right)^{N-1} = \left(\int_0^{\infty} w_{T_M}(\psi) (1 - F_{V(T_M-1)}(\psi)) d\psi \right) \left(\int_0^{\infty} w_{T_M}(\psi) (1 - F_{P(T_M-1)}(\psi)) d\psi \right)^{N-1}. \quad (25)$$

Similarly we can find the error probability of the second type $P^{(1)}$ (the probability of the event that at least one of N P-procedures performs not less than T_M iterations at M algorithm steps):

$$P^{(1)} = 1 - \left(\sum_{i=1}^{T_M-1} P_{Pi} \right)^N.$$

In this case V-procedures are absent so $w_{T_M}(\psi) = w_{PT_M}^{(N)}(\psi)$ - probability distribution of minimum of N penalty function values of P-procedures. Then

$$P^{(1)} = 1 - \left(\sum_{t=1}^{T_M-1} P_{Pt}^{(1)} \right)^N = 1 - \int_0^{\infty} w_{PT_M}^{(N)}(\psi) (1 - F_{P(T_M-1)}(\psi)) d\psi. \quad (26)$$

4.3 The probability of goal function extremum subdomain erroneous choice

One of the most important characteristics of the considered structural optimizations of PGAs is the probability P_{er} of the V-subdomain erroneous choice. First, let us consider the simplest case, when the domain $\Omega_{\bar{\alpha}}$ is decomposed into only two subdomains and one of them is of P-type and the other is of V-type. A subdomain of P-type will be chosen if the procedure operating in the P-subdomain is the first to reach the final T -th iteration, i.e. when the condition $\psi_{PT} < \psi_V$ is met, where ψ_{PT} - penalty function value for the P-procedure at the T -th iteration, ψ_V - penalty function value for the V-procedure that can theoretically perform from 1 to $(T-1)$ iterations. Let us assume we know a priori the value ψ_s of the penalty function which is considered to be a criterion of the V-subdomain choice. Then, to have a error of the V-subdomain choice it is necessary the simultaneous performing of two events: the penalty function of the V-procedure has exceeded the value ψ_s and of the P-procedure has not exceeded the value ψ_s . The probabilities of these events are determined by expressions

$$\int_{\psi_s}^{\infty} w_{VT}(\psi) d\psi = 1 - F_{VT}(\psi) \quad \text{and} \quad \int_0^{\psi_s} w_{PT}(\psi) d\psi = F_{PT}(\psi),$$

where $F_{VT}(\psi) = \int_0^{\psi} w_{VT}(x) dx$, $F_{PT}(\psi) = \int_0^{\psi} w_{PT}(x) dx$ - distribution function of ψ at the T -th iteration for procedures of V- and P-type correspondingly. Assuming the independence of the mentioned events the probability of the erroneous choice amounts to:

$$P_{er} = (1 - F_{VT}(\psi_s)) F_{PT}(\psi_s).$$

However, the value ψ_s is unknown a priori. Then in presence of the V-subdomain in the parameter definition domain the conditional probability $P_{er|V}$ of the choice of the P-procedure instead of the V-procedure is the probability of the event that $\psi_{VT} > \psi_{PT}$ at the T -th iteration:

$$P_{er|V} = \int_0^{\infty} w_{VT}(\psi) \int_0^{\psi} w_{PT}(x) dx d\psi = \int_0^{\infty} w_{PT}(\psi) (1 - F_{VT}(\psi)) d\psi.$$

Let us note that $w_{VT}(\psi)$ and $w_{PT}(\psi)$ depend on the distribution density $w(\hat{\alpha}_0)$ of the initial approximation $\hat{\alpha}_0$ of vector of the parameters to be estimated and in this sense they are considered to be estimated the conditional probabilities too. Below for definiteness, we assume that the initial approximation $\hat{\alpha}_0$ for the V-procedure gives the worst estimate in the operating range and the sought probability $P_{er|V}$ corresponds to the upper bound of the error probability when choosing a V-subdomain.

It is easy to show the probability $P_{er|V}$ for the case of presence of one V-subdomain and $(N-1)$ P-subdomain in the parameter definitional domain is equal to

$$P_{er|V} = \int_0^{\infty} w_{VT}(\psi) \left(1 - (1 - F_{PT}(\psi))^{N-1}\right) d\psi. \quad (27)$$

Here, the penalty function values of the procedures operating in P-subdomains are considered to be independent. Let us note that assumed constraint on independence of local samples from different P-procedures is not rigid, because in real images the samples in the domains corresponding to P-procedures are almost noncorrelated.

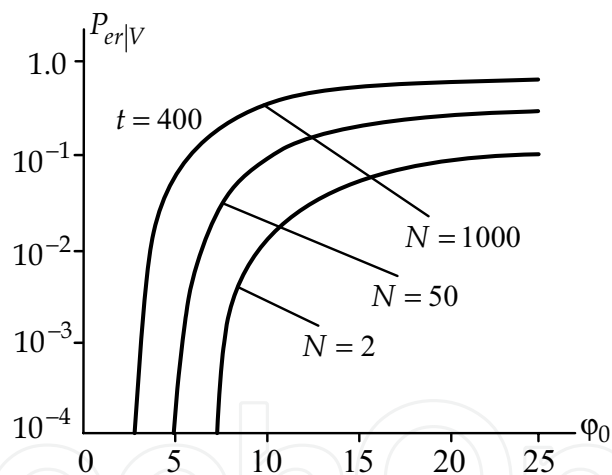


Fig. 8. The probability of omission of the sought fragment on the studied image

As an example, in Fig 8. graphs of the probability $P_{er|V}$ of omission of the sought fragment on the image to be studied are given. The IIGD parameters include parallel shift $\bar{h} = (h_1, h_2)^T$ of the sought and standard fragments and rotation angle ϕ between them. The calculation was carried out for the case of usage of the pseudogradient (10) and diagonal gain matrix with diagonal elements $\lambda_{ht} = 0.1/(1+0.1t)$ and $\lambda_{\phi t} = 0.1/(1+0.01t)$ for \bar{h} and ϕ , respectively. The situations when image size required the partitioning of the parameter definitional domain $\Omega_{\bar{\alpha}}$ into 2, 50 and 1000 subdomains have been studied. The presented

plots correspond to the 400th iteration of the estimation. Here, the initial approximation φ_0 of the rotation angle varied from 25 to 0 degrees and the initial approximation h_0 of the parallel shift was fixed and equal to 5 sample grid steps. The results have been obtained for the image parameters corresponding to the previous examples. On the basis of the graphs it is possible to find, for example, the maximum rotation angle at which the required probability $P_{er|V}$ is guaranteed. As φ_0 decreases the probability $P_{er|V}$ increases too, for example, if $\varphi_0 = 10^\circ$ then at $N = 50$ the probability $P_{er|V} \leq 10^{-2}$ and at $N = 1000$ - $P_{er|V} \leq 10^{-3}$.

Let us note that if the operating range of procedures exceeds the subdomains $\Omega_{\alpha}^{(i)}(\hat{\alpha}_{0i})$, then several V-procedures can correspond to one extremum of the goal function which increases the error probability $P_{er|V}$ when doing V-subdomain search. It can be explained by the fact that in this situation the error condition at the V-subdomain choice is in a lesser than that of all V-procedures penalty function value of the P-procedure, having the largest number of iterations. To find the probability $P_{er|V}$ in this case we can use the expression (27). However, in doing so, the probability distribution $w_{VT}(\psi)$ at that has the sense of probability distribution of the penalty function minimum value among all V-procedures.

It is also of interest also to study the situation when the goal function has several extremums as it is, for example, in the problem of search of several similar objects locations on an image. In this case derivation of the corresponding relations enabling to find the probability of omission of one or several subdomains of extremum position does not cause theoretical difficulties.

4.4 Computational expense analysis

Probability distributions of the number of iterations, performed by the V-procedure and P-procedures when attaining the threshold T -th iteration by one of the procedures, contain information necessary for analysis of average computational expenses.

Assume it is known a priori that there is one subdomain of V-type among N subdomains. Suppose also that a penalty function value of the leading procedure at the T -th iteration to be known and equaled to ψ_T . Then, the conditional probability of the event that the value ψ_{P1} of the penalty function of the P-procedure exceeds ψ_T after first iteration is

$$P_P(1|\psi_T) = P(\psi_{P1} > \psi_T) = \int_{\psi_T}^{\infty} w_{P1}(\psi) d\psi.$$

Here, on average $(N-1)P_P(1|\psi_T)$ procedures of $(N-1)$ P-procedures perform only one iteration. The probability $P_P(t|\psi_T)$ of the event that the excess of the value ψ_T occurs directly after the t -th iteration is equal to

$$P_P(t|\psi_T) = \int_{\psi_T}^{\infty} w_{P_t}(\psi) d\psi - \int_{\psi_T}^{\infty} w_{P_{(t-1)}}(\psi) d\psi = \int_{\psi_T}^{\infty} w_{P_t}(\psi) d\psi - \sum_{i=1}^{t-1} P_P(i|\psi_T). \quad (27)$$

Here, on average $(N-1)P_P(t|\psi_T)$ procedures fulfill t iterations. The relation (27) at $t = \overline{1, T-1}$ corresponds to the conditional discrete distribution of the number of iterations

performed by the P-procedure. We can also obtain a similar conditional probability distribution $P_V(t|\psi_T)$, $t = \overline{1, T-1}$ for the V-procedure. The total average number of the procedures which have performed t iteration constitutes

$$N_m(t) = P_V(t|\psi_T) + (N-1)P_P(t|\psi_T).$$

To find unconditional probability distribution $\{P_t\}$, $t = \overline{1, T}$ it is necessary to take into account the fact that a value ψ_T is a random one. Here, the V-procedure will be the leading one with the probability $P(\min \psi_{PTi} > \psi_T)$, $i = \overline{1, N-1}$, where $\min \psi_{PTi}$ - minimum of penalty function values for P-procedures. If $\psi_{VT} > \psi_T$ then one of P-procedures will be the leading one with the probability $P(\psi_{VT} > \psi_T) = 1 - F_{VT}(\psi)$. Then, for the probability distribution $w(\psi_T)$ of the minimum penalty function value ψ_T for all procedures

$$w(\psi_T) = w_{VT}(\psi)P(\min \psi_{PTi} > \psi_T) + w_{PT}(\psi)P(\psi_{VT} > \psi_T). \quad (28)$$

On the assumption of noncorrelation of the local samples of P-procedures

$$P(\min \psi_{PTi} > \psi_T) = (1 - F_{PT}(\psi))^{N-1}.$$

Taking into account the relations (27) и (28) for the unconditional distribution $\{P_t\}$ of the number of iterations we can write

$$P_t = \int_0^\infty w_T(\psi) \int_\psi^\infty (w_{Pt}(x) - w_{P(t-1)}(x)) dx d\psi, \quad t = \overline{1, T-1}, \quad (29)$$

where

$$w_T(\psi) = w_{VT}(\psi)(1 - F_{PT}(\psi))^{N-1} + w_{PT}^{(N-1)}(\psi)(1 - F_{VT}(\psi)); \quad (30)$$

$w_{PT}^{(N-1)}(\psi)$ - probability distribution of the penalty function minimum value for P-procedures at the T -th iteration.

Let us note if we need calculation of all the components of the distribution (29) then to reduce computational expenses it is convenient to use the following recurrent relation

$$P_t = \int_0^\infty w_T(\psi)(1 - F_{Pt}) d\psi - \sum_{i=1}^{t-1} P_{Pi}, \quad (31)$$

and for calculation of the probability $P_{t-l,t}$ of finding of the number of iterations in the range from $(t-l)$ to t , $l = \overline{2, t-1}$ the expression

$$P_{t-l,t} = \int_0^\infty w_T(\psi) \int_\psi^\infty (w_{Pt}(x) - w_{P(t-l)}(x)) dx d\psi.$$

In absence of the goal function extremum in the parameter definitional domain the expression (30) is simplified $w_T(\psi) = w_{PT}^{(N)}(\psi)$.

Then for the distribution $\{P_t\}$ we obtain

$$P_t = \int_0^\infty w_{PT}^{(N)}(\psi)(1-F_{P_t})d\psi - \sum_i^{t-1} P_{P_i}, t = \overline{1, T-1}. \quad (32)$$

In Fig. 9 examples of the probability distribution of number of iterations for the problem of fragment search on the image for the situations of presence (Fig. 9,a) and absence (Fig. 9,b) of the sought fragment on the image are presented. The results have been obtained using the relations (31) and (32) for the case of attainment of the 100th iteration by the leading procedure. The fragment shift with reference to the subdomain center was equal to 5 sample grid steps, interframe correlation sample coefficient was used as the goal function, the pseudogradient (10) and diagonal gain matrix with elements $\lambda_{ht} = 0.05/(1+0.04t)$, $g = 100$ were applied in procedures. For the given example in absence of the sought fragment on the image (in absence of the goal function extremum in the parameter definitional domain) the total number of all the procedures is about 2.3 times as large.

In the diagrams the experimental results obtained at the same parameters on the simulated Gaussian images, when the image domain was decomposed into 900 subdomains averaged on 200 realizations are also shown (with circles). A good correspondence of the theoretical and the experimental results can be seen. The performed simulation showed that at $T > 200$ the distribution $\{P_t\}$ is normalized which enables to use Gaussian approximation.

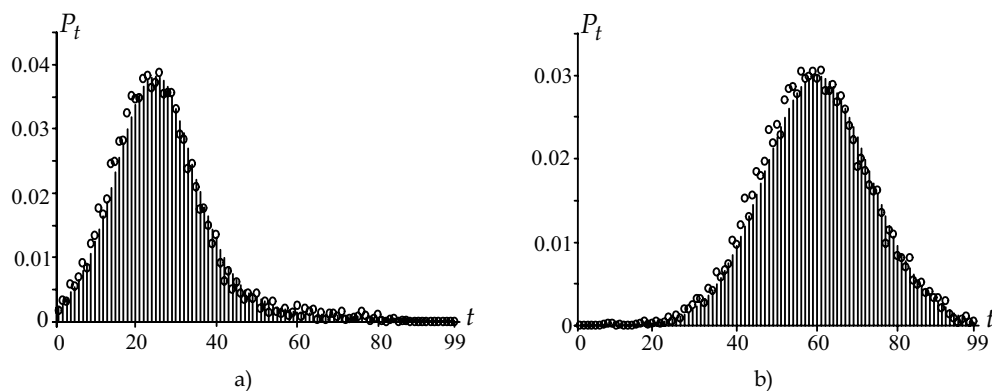


Fig. 9. Probability discrete distributions of the number of iterations

When using structural optimization the computational expenses depend on both the type of pseudogradient procedures to be used and computational resources due to which the algorithm is realized. Let us consider only approximate relations primarily characterizing the total number of iterations performed by the procedures to attain the required result. Assume the time necessary to perform one step of the PGA to be the sum of the time τ_1 , spent on performing of one iteration by the procedure, and the time τ_ψ of finding the procedure that has the least value of the penalty function (for example, by arranging the penalty function values of all the procedures in ascending order and by choosing the procedure leading in this arranged sequence).

The computational expenses $E^{(1)}$ on testing of the hypothesis about absence of the V -subdomain at the given error $P^{(1)}$ of the first type are the simplest to find. In this case the

threshold value of the statistical criterion of hypothesis acceptance is the number M_{c1} of iterations performed by all procedures. Accordingly,

$$E^{(1)} = M_{c1}(\tau_I + \tau_\Psi).$$

If the hypothesis is rejected then the number of iterations T_M is chosen as a rule basing on the required probability of V-subdomain omission. Then for the mathematical expectation of the total computational expenses E_T we can write

$$E_T = (\tau_I + \tau_\Psi) \left(T + (N-1) \sum_{t=1}^{T-1} t P_t \right),$$

where P_t - probability of the event that the procedure has performed t iterations; N - total number of iterations. If all procedures have attained T iterations then the computational expenses constitute at least

$$E_{TA} = \tau_I T N.$$

Accordingly the gain in computational expenses when using structural optimization in comparison with the case when all procedures are brought to the threshold number of iterations is determined by the relation

$$G = \frac{E_{TA}}{E_T} = \frac{\tau_I}{\tau_I + \tau_\Psi} \cdot \frac{1}{\frac{1}{N} + \left(\frac{1}{T} - \frac{1}{NT} \right) \sum_{t=1}^{T-1} t P_t}. \quad (33)$$

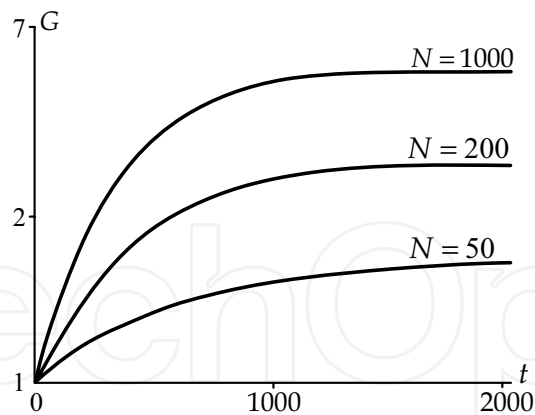


Fig. 10. Gain in computational expenses when using structural optimization

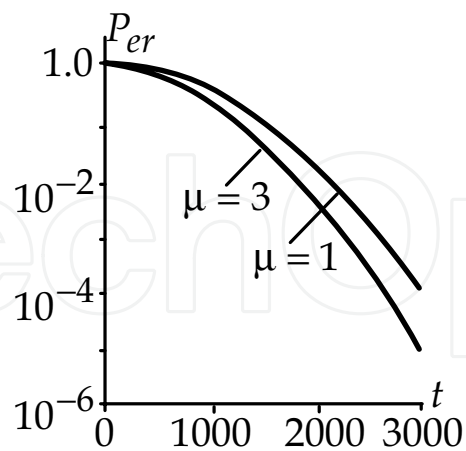
As an example, diagrams of the gain G in computational expenses at $N = 50, 200$ and 1000 obtained by means of the relation (33) for the problem of image fragment search are

presented in Fig. 10. Here, we assumed that $\tau_\psi = 0.1 \tau_r$. It is seen that the gain depends on the number N of subdomains of the parameter definitional domain. Thus, at $t=100$ for $N = 50$ the gain $G = 1.6$, for $N = 200 - G = 2.5$, for $N = 1000 - G = 5.3$.

Structural optimization of PGAs was used, for example, to solve the problem of small fragments location search on large-sized images with reference to a given standard fragment. Here, the sought and the standard fragments had reciprocal rotation angle, different scale and amplitude distortions. In Fig. 11,a) an example of image of 3048×1608 elements and the sought fragment of 48×48 elements is given. The number of subdomains of image decomposition has constituted $N = 31200$ and computational expense reduction in comparison with the traditional approach - more than 15 times as large. In Fig. 11,b) graphs of dependence of the probability P_{er} of fragment omission versus the number of iterations at the local sample size $\mu = 1$ and 3 (under the condition that the sought fragment does exist on the image) are shown.



a)



b)

Fig. 11. Image, fragment and error probability of fragment omission

5. Conclusion

The considered PGAs can be directly used in various applied problems of image processing. The algorithms of this class can be applied to image processing in the conditions of a priori uncertainty, they assume small computational expenses and do not require the preliminary estimation of the parameters of the image to be studied. The estimates formed through them are immune to impulse interference and converge to optimal values under rather weak conditions. At an unknown set of the parameters of geometrical deformations model PGAs enable to estimate shifts of each node of image sample grid. At a given IIGD model the processing of the image samples can be performed in an arbitrary order, for example, in order of scanning with decimation that is determined by the hardware speed, which facilitates obtaining a tradeoff between image entering rate and the speed of the available hardware. The mentioned properties make them attractive for usage in real time systems.

Unfortunately a limited size of this manuscript does not make it possible to consider some important aspects of this lead of investigations, in particular, the analysis of probabilistic properties and computational expenses at PGAs structural optimization for the situation when the goal function has several extremums. Let us note two more such aspects for further study in the form of the problem definition.

A disadvantage of the PGAs when performing the processing of real images is in the presence of local extremums of the goal function estimate characterizing the estimation quality which significantly reduces convergence speed or even may lead to its failure at some realizations in the process of estimate convergence. Besides, algorithms of this class have a comparably small range of operating.

At that the estimate convergence character and computational expenses in many respects depend on the image samples local sample size used on various iterations of estimation. Thus the development and study of the methods of a priori and a posteriori optimization of size and plan of the sample used to obtain the goal function pseudogradient is considered to be a vital problem. One of the trends of a posteriori optimization is planned in the part 3.3 in this work. Of works concerned with a priori optimization we can highlight (Samojlov, 2006; Minkina et al., 2005).

Modern information systems are characterized by increasing rate of the entering data. It gives rise to the vitality of pseudogradient procedures optimization on criteria of computational expense minimum and iterations number minimum at limitations on computational expenses. We should note that many scientists addressed to the study of precision potentiality of the pseudogradient procedures, in particular, (Albert & Gardner, 1967; Benveniste et al., 1990). Asymptotic rate of convergence of the estimates to be formed has been profoundly studied in works (Chung, 1954; Sacks, 1958), in works (Goodwin & Payne, 1977; Soderstrom, 1981) and others the conditions of asymptotic normality of various pseudogradient procedures have been found, the works (Kushner & Clark, 1978; Tsytkin & Polyak, 1974) are devoted to estimation of asymptotic rate of convergence. However, for practical application of these procedures the investigation of their precision potentiality at a finite number of iterations is of significant importance. Unfortunately, at present this issue has been studied insufficiently. It is due to the fact that at a finite number of iterations an analysis of interframe deformations parameter estimates probabilistic properties is complicated by a large number of factors whose effect cannot be ignored. These are the nature of probability densities and autocorrelation functions of images and interfering noise, the kind of goal function determining the quality of estimation, the parameters of the

procedures and the number of iterations. Besides, when estimating the image parameters we have to deal with complex assemblage of hindering factors such as time and spatial inhomogeneity of characteristics of the desired signals and noise, inhomogeneity of sensitivity and the faults in sensors, pulse interference, etc. These above-mentioned factors are of random nature, so when describing real images both parametric and non-parametric a priori uncertainty nearly always takes place. One of the techniques of analysis of probabilistic characteristics of the estimates, formed by PGAs at a finite number of iterations has been proposed in works (Tashlinskii & Tikhonov, 2001; Tashlinskii, 2004) and was used when developing the chapters 3.1, 4.3 and 4.4 of the present manuscript. However, this lead requires further serious investigations.

6. References

- Albert, A. & Gardner, L. (1967). *Stochastic approximation and nonlinear regression*, MIT-Press, Cambridge, Massachusetts
- Benveniste, A.; Metivier, M. & Priouret, P. (1990). *Adaptive Algorithms and Stochastic Approximations*, Springer-Verlag, Berlin
- Chung, K. L. (1954). On a stochastic approximation method. *The Annals of Mathematical Statistics*, Vol. 25, No. 3, pp. 463–483
- Gonzalez, R. C. & Woods, R. E. (2002). *Digital Image Processing*. Prentice Hall, New Jersey, ISBN 0-201-18075-8
- Goodwin, G. C. & Payne, R.L. (1977). *Dynamic system identification: Experimental design and data analysis*. Academic Press, New York
- Korn, G. A. & Korn, T. M. (1968). *Mathematical handbook for scientists and engineers*. McGraw-Hill Book Company, New York, San Francisco, Toronto, London, Sydney
- Krasheninikov, V. R. (2003). *Osnovi teorii obrabotki izobrazhenii [The Fundamentals of image processing theory]*, UIGTU, ISBN 5-89146-434-9, Uljanovsk [in Russian]
- Kushner, H.J. & Clark, D.S. (1978). *Stochastic approximation methods for constrained and unconstrained systems*. Springer-Verlag, New York
- Minkina, G. L. & Samojlov, M. U. (2005). Choice of Values Characterizing Estimate Convergence for Pseudogradient Estimation of Image Interframe Deformation Parameters, *Vestnik UIGTU [Herald of the UISTU]*, No. 4, pp. 32–37, ISSN 1674-7016 [in Russian]
- Minkina, G. L.; Samojlov, M. U. & Tashlinskii, A. G. (2007). Choice of The Objective Function for Pseudogradient Measurement of Image Parameters. *Pattern Recognition and Image Analysis*, Vol. 17, No. 1, pp. 136–139, ISSN 1054-6618
- Polyak, B. T. & Tsypkin, Ya. Z. (1973). Pseudogradient Algorithms of Adaptation and Education. *Avtomatika i telemekhanika [Automation and Telemechanics]*, No. 3, pp. 45–68, ISSN 0005-2310 [in Russian]
- Polyak, B. T. (1976). Convergence and Convergence Speed of Iterative Stochastic Algorithms: a General Case. *Avtomatika i telemekhanika [Automation and Telemechanics]*, No 2, pp. 83–94, ISSN 0005-2310 [in Russian]
- Polyak, B. T. & Tsypkin, Ya. Z. (1984). Criterion Algorithms of Stochastic Optimization. *Avtomatika i telemekhanika [Automation and Telemechanics]*, No. 6, pp. 95–104, ISSN 0005-2310 [in Russian]
- Sacks, J. (1958). Asymptotic distribution of stochastic approximation. *The Annals of Mathematical Statistics*, Vol. 29, No. 2, pp. 373–405

- Samojlov, M. U. (2006). Procedure Optimization of Image Interfarme Geometrical Deformation Parameter Pseudogradient Estimation. *Radiolocation, Navigation, Connection: Proc. of The XII Inter. Sci.-Tech. Conf.*, Vol. 1, pp.162-167, ISBN 5-9900094-8-8, Voronezh, March, 2006, Sakvoee, Voronezh
- Soderstrom, T. (1981). *On a method for model selection in system identification*. *Automatica*, Vol. 13, No. 2, pp. 387-388
- Tashlinskii, A. G. (2000). *Otsenivanie parametrov prostranstvennih deformatsii posledovatelnoitei izobrazhenii [Image Sequence Spatial Deformation Parameters Estimation]*, UIGTU, ISBN 5-89146-204-4, Uljanovsk [in Russian]
- Tashlinskii, A. G. & Tikhonov, V. O. (2001). The Method of Multidimensional Process Parameter Pseudogradient Measuring Error Analysis. *Izvestija Vuzov: Radioelektronika, [Proceedings of institutes of higher education: Radio Electronics]*, Vol. 44, No. 9, pp. 75-80, ISSN 0021-3470 [in Russian]
- Tashlinskii, A. G. (2002). Pseudogradient estimation of image sequence spatial deformations. *A Publication of The International Association of Science and Technology for Development - IASTED*, pp. 382-385, ACTA Press, ISBN 0-889860342-3, Anaheim, Calgary, Zurich
- Tashlinskii, Alexandr (2003). Computational Expenditure Reduction in Pseudo-Gradient Image Parameter Estimation. *Computational Science*, Vol. 2658, No. 2, pp. 456-462, ISSN 0302-9743
- Tashlinskii, A. G. (2004). Probabilistic Modeling of Image Parameters Pseudogradient Measuring Process. *7-th International Conference PRIA-7-2004*, Vol. II, pp. 402-403, ISBN 5-7629-0631-0, St. Peterburg, October, 2004, Nauka/Interperiodica, St. Peterburg
- Tashlinskii, A. G. (2006). Structural Optimization of Pseudogradient Algorithms for Estimating Image Parameters. *Pattern Recognition and Image Analysis*, Vol. 16, No. 2, pp. 218-222, ISSN 1054-6618
- Tsyarkin, Ya. Z. & Polyak, B. T. (1974). Attainable Accuracy of Adaptation Algorithms. *Dokladi AN SSSR [Proceedings of USSR Academy of Sciences]*, Vol. 218, No. 3, pp. 532-535, ISSN 0002-3264 [in Russian]
- Tsyarkin, Ya. Z. (1995). *Informatsionnaya teoriya identifikatsii [Information theory of Identification]*, Nauka/ Fizmatlit, ISBN 5-02-015071-1, Moskow [in Russian]
- Vasiliev, K. K. & Tashlinskii, A. G. (1998). Estimation of Deformation Parameters of Multidimensional Images to Be Observed on The Background of Interference. *Pattern Recognition and Image Analysis: New Information Technologies, Proceedings of the IV Inter. Conf.*, Vol. 1, pp. 261-264, Novosibirsk, June, 1998, SO RAN, Novosibirsk, ISBN 5-89896-189-5



Vision Systems: Segmentation and Pattern Recognition

Edited by Goro Obinata and Ashish Dutta

ISBN 978-3-902613-05-9

Hard cover, 536 pages

Publisher I-Tech Education and Publishing

Published online 01, June, 2007

Published in print edition June, 2007

Research in computer vision has exponentially increased in the last two decades due to the availability of cheap cameras and fast processors. This increase has also been accompanied by a blurring of the boundaries between the different applications of vision, making it truly interdisciplinary. In this book we have attempted to put together state-of-the-art research and developments in segmentation and pattern recognition. The first nine chapters on segmentation deal with advanced algorithms and models, and various applications of segmentation in robot path planning, human face tracking, etc. The later chapters are devoted to pattern recognition and covers diverse topics ranging from biological image analysis, remote sensing, text recognition, advanced filter design for data analysis, etc.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

A. G. Tashlinskii (2007). Pseudogradient Estimation of Digital Images Interframe Geometrical Deformations, Vision Systems: Segmentation and Pattern Recognition, Goro Obinata and Ashish Dutta (Ed.), ISBN: 978-3-902613-05-9, InTech, Available from:

http://www.intechopen.com/books/vision_systems_segmentation_and_pattern_recognition/pseudogradient_estimation_of_digital_images_interframe_geometrical_deformations

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2007 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen