Optimization of Goal Function Pseudogradient in the Problem of Interframe Geometrical Deformations Estimation

A.G. Tashlinskii
Ulyanovsk State Technical University
Russia

1. Introduction

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 which are characterized by space-time correlation and represent the sequence of framed of image to be changed (Gonzalez & Woods, 2002). Interframe geometrical deformations can be described by mathematical models of deformations of grids, on which images are specified.

Estimation of variable parameters of interframe deformations is required when solving a lot of problems, for example, at automate search of fragment on the image, navigation tracking of mobile object in the conditions of limited visibility, registration of multiregion images at remote investigations of Earth, in medical investigations. A large number of calls for papers are devoted to different problems of interframe deformations estimation (the bibliography is presented for example in (Tashlinskii, 2000)). This chapter is devoted to one of approaches, where the problems of optimization of quality goal function pseudogradient in pseudogradient procedures of interframe geometrical deformations parameters estimation are considered.

Let the model of deformations is determined with accuracy to a parameters vector $\overline{\alpha}$, frames $\mathbf{Z}^{(1)}=\left\{z_{\bar{j}}^{(1)}:\ \bar{j}\in\Omega\right\}$ and $\mathbf{Z}^{(2)}=\left\{z_{\bar{j}}^{(2)}:\bar{j}\in\Omega\right\}$ to be studied of images are specified on the regular sample grid $\Omega=\left\{\bar{j}=\left(j_x,j_y\right)\right\}$, and a goal function of estimation quality is formed in terms of finding extremum of some functional $\mathrm{J}(\overline{\alpha})$. However, it is impossible to find optimal parameters in the mentioned sense because of incompleteness of image observations description. But we can estimate parameters $\overline{\alpha}$ on the basis of analysis of specific images $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ realizations, between of which geometrical deformations are estimated. At that it is of interest to estimate $\overline{\alpha}$ directly on values $\mathrm{J}\left(\hat{\overline{\alpha}},\mathbf{Z}^{(1)},\mathbf{Z}^{(2)}\right)$ (Polyak & Tsypkin, 1984):

$$\hat{\overline{\alpha}}_t=\hat{\overline{\alpha}}_{t-1}-\mathbf{\Lambda}_t\nabla\mathrm{J}\left(\hat{\overline{\alpha}},\mathbf{Z}^{(1)},\mathbf{Z}^{(2)}\right), \quad (1)$$

Source: Pattern Recognition Techniques, Technology and Applications, Book edited by: Peng-Yeng Yin,
ISBN 978-953-7619-24-4, pp. 626, November 2008, I-Tech, Vienna, Austria

where $\hat{\bar{\alpha}}_t$ - the next after $\hat{\bar{\alpha}}_{t-1}$ approximation of the extremum point of $\mathrm{J}\left(\hat{\bar{\alpha}}, \mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}\right)$; $\mathbf{\Lambda}_t$ - positively determined matrix, specifying a value of estimates change at the t-th iteration; $\nabla \mathrm{J}(\cdot)$ - gradient of functional $\mathrm{J}((\cdot))$. The necessity of multiple and cumbrous calculations of gradient opposes to imply the procedure (1) in the image processing. We can significantly reduce computational costs if at each iteration instead of $\mathrm{J}\left(\hat{\bar{\alpha}}, \mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}\right)$ we use its reduction $\nabla \hat{\mathrm{J}}(\hat{\bar{\alpha}}_{t-1}, Z_t)$ on some part Z_t of realization which we call the local sample

$$Z_t = \left\{ z_{\bar{j}t}^{(2)}, \tilde{z}_{\bar{j}t}^{(1)} \right\},\ z_{\bar{j}t}^{(2)} \in \mathbf{Z}^{(2)},\ \tilde{z}_{\bar{j}t}^{(1)} = \tilde{z}^{(1)}(\bar{j}_t, \hat{\bar{\alpha}}_{t-1}) \in \tilde{Z}, \tag{2}$$

where $z_{\bar{j}t}^{(2)}$ – samples of a deformed image $\mathbf{Z}^{(2)}$, chosen to the local sample at the t-th iteration; $\tilde{z}_{\bar{j}t}^{(1)}$ – sample of a continuous image $\tilde{Z}^{(1)}$ (obtained from $\mathbf{Z}^{(1)}$ by means of some interpolation), the coordinates of which correspond to the current estimate of sample $z_{\bar{j}l}^{(2)} \in \mathbf{Z}^{(2)}$; $\bar{j}_t \in \Omega_t \in \Omega$ – coordinates of samples $z_{\bar{j}t}^{(2)}$; Ω_t – plan of the local sample at the t-th iteration. Let us call the number of samples $\left\{ z_{\bar{j}t}^{(2)} \right\}$ in Z_t through the local sample size and denote through μ.

At large image sizes pseudogradient procedures (Polyak & Tsypkin, 1973; Tashlinskii, 2005) give a solution satisfying to requirements of simplicity, rapid convergence and availability in different real situations.

For considered problem the pseudogradient $\bar{\beta}_t$ is any random vector in the parameters space, for which the condition $\left[\nabla \mathrm{J}(\hat{\bar{\alpha}}_{t-1}, Z_t)\right]^T \mathrm{M}\left\{\bar{\beta}_t\right\} \geq 0$ is fulfilled, где T - sign of transposition; $\mathrm{M}\{\cdot\}$ - symbol of the mathematical expectation.

Then pseudogradient procedure is (Tzypkin, 1995) :

$$\hat{\bar{\alpha}}_t = \hat{\bar{\alpha}}_{t-1} - \mathbf{\Lambda}_t \bar{\beta}_t, \tag{3}$$

where $t = \overline{0,T}$ - iteration number; T - total number of iterations.

Procedure (3) own indubitable advantages. It is applicable to image processing in the conditions of a priory uncertainty, supposes not large computational costs, does not require preliminary estimation of parameters of images to be estimated. The formed estimates are immune to pulse interferences and converge to true values under rather weak conditions. 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 (Tashlinskii, 2003).

However, pseudogradient procedures have disadvantages, in particular, the presence of local extremums of the goal function estimate at real image processing, that significantly reduces the convergence rate of parameters estimates. To the second disadvantage we can refer relatively not large effective range, where effective convergence of estimates is ensured. This disadvantage depends on the autocorrelation function of images to be estimated. A posteriori optimization of the local sample (Minkina et al., 2005; Samojlov et al., 2007; Tashlinskii et al., 2005), assuming synthesis of estimation procedures, when sample size automatically adapted at each iteration for some condition fulfillment is directed on the struggle with the first one. Relatively second disadvantage it is necessary to note that for increasing speed of procedures we tend to decrease local sample size, which directly influences on the convergence rate of parameters to be estimated to optimal values: as μ is larger, the convergence rate is higher. However on the another hand the increase of μ inevitably leads to increase of computational costs, that is not always acceptable. Let us note that at different errors of parameters estimates from optimal values at the same value of sample size the samples chosen in different regions of image ensure different estimate convergence rate. Thus, the problems of optimization of size and plan of local sample of samples used for goal function pseudogradient finding are urgent. The papers (Samojlov, 2006; Tashlinskii @ Samojlov, 2005; Dikarina et al., 2007) are devoted to solution of the problem of a priory optimization of local sample, in particular, on criteria of computational expenses minimum The problems of optimization of a plan of local sample samples choice are investigated weakly, that has determined the goal of this work.

Pseudogradient estimation of parameters (3) is recurrent, thus as a result of iteration the estimate $\hat{\alpha}_{i,t}$ of the parameter α_i changes discretely: $\hat{\bar{\alpha}}_t = \hat{\bar{\alpha}}_{t-1} + \Delta\hat{\bar{\alpha}}_t$. At that the following events are possible:

- If $\operatorname{sign}(\varepsilon_{i,t-1}) = \operatorname{sign}\Delta\alpha_{i,t}$, then change of the estimate $\hat{\bar{\alpha}}_t$ is directed backward from the optimal value α_i^*, where $\varepsilon_{i,t} = \hat{\alpha}_{i,t} - \alpha_i^*$ – the error of its optimal value of the parameter α_i^* and its estimate, $i = \overline{1,m}$. In accordance with (Tashlinskii & Tikhonov, 2001) let us denote the probability of such an event through $\rho_i^-(\bar{\varepsilon}_t)$.
- At $\Delta\alpha_{i,t} = 0$ the estimate $\hat{\bar{\alpha}}_t$ does not change with probability $\rho_i^0(\bar{\varepsilon}_t)$.
- If $-\operatorname{sign}(\varepsilon_{i,t-1}) = \operatorname{sign}\Delta\alpha_{i,t}$, the change of the estimate $\hat{\bar{\alpha}}_t$ is directed towards the optimal value of the estimate with some probability $\rho_i^+(\bar{\varepsilon}_t)$.

Let us note, that the probabilities $\rho_i^+(\bar{\varepsilon}_t)$, $\rho_i^0(\bar{\varepsilon}_{t-1})$ and $\rho_i^-(\bar{\varepsilon}_t)$ depend on the current errors $\bar{\varepsilon}_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \dots, \varepsilon_{m,t})^T$ of other parameters to be estimated, but because of divisible group of events we have $\rho_i^+(\bar{\varepsilon}_t) + \rho_i^-(\bar{\varepsilon}_t) = 1 - \rho_i^0(\bar{\varepsilon}_t)$. If the goal function is maximized and $\varepsilon_{i,t} > 0$, then $\rho_i^+(\bar{\varepsilon}_t)$ is the probability of the fact that projection β_i of the pseudogradient on the parameters α_i axes will be negative, and $\rho_i^-(\bar{\varepsilon}_t)$ – positive:

$$\rho_i^+(\bar{\varepsilon}_t) = P\{\beta_i < 0\} = \int_{-\infty}^{0} w(\beta_i)d\beta_i\,, \quad \rho_i^-(\bar{\varepsilon}_t) = P\{\beta_i > 0\} = \int_{0}^{\infty} w(\beta_i)d\beta_i\,, \tag{4}$$

where $w(\beta_i)$ – probability density function of the projection β_i on the axes α_i.

Probabilities $\rho_i^+(\bar{\varepsilon}_t)$, $\rho_i^0(\bar{\varepsilon}_t)$ and $\rho_i^-(\bar{\varepsilon}_t)$ will be used below for finding optimal region of local sample samples choice on some criterion.

2. Finding goal functions pseudorgadients with usage of finite differences

In the papers (Vasiliev & Tashlinskii, 1998; Vasiliev &Krasheninikov, 2007) it is shown, that when pseudogradient estimating of interframe deformations parameters as a goal function it is reasonable to use interframe difference mean square and interframe correlation coefficient. Pseudogradients of the mentioned functions are found through a local sample Z_t and estimates $\hat{\bar{\alpha}}_{t-1}$ of deformations parameters at the pervious iteration:

$$\bar{\beta}_t = \sum_{\bar{j}_t \in \Omega_t} \frac{\partial \tilde{z}^{(1)}_{\bar{j},t}}{\partial \bar{\alpha}} \left(\tilde{z}^{(1)}_{\bar{j},t} - z^{(2)}_{\bar{j},t} \right) \Bigg|_{\bar{\alpha} = \hat{\bar{\alpha}}_{t-1}} \text{ and } \bar{\beta}_t = -\sum_{\bar{j}_t \in \Omega_{\bar{j},t}} \frac{\partial \tilde{z}^{(1)}_{\bar{j},t}}{\partial \bar{\alpha}} z^{(2)}_{\bar{j}t} \Bigg|_{\bar{\alpha} = \hat{\bar{\alpha}}_{t-1}} .$$

However the direct usage of the obtained expressions for images, specified by discrete sample grids, is impossible, because they include analytic derivatives. Thus let us briefly consider approaches for goal functions pseudogradients calculation.

At the explicitly given function its estimate $\hat{J}$ at the current iteration can be found, using estimates $\hat{\bar{\alpha}}$ of deformations parameters, obtained by this iteration, information about brightness z and coordinates (x, y) of samples of the local sample, formed at the current iteration, and accepted deformations model. Thus the dependence of the goal function on parameters can be represented directly:

$$\hat{J} = f(\bar{\alpha}) , \tag{5}$$

and through intermediate brightness functions:

$$\hat{J} = f(z(\bar{\alpha})) ,\ z = u(\bar{\alpha}) \tag{6}$$

and coordinates:

$$\hat{J} = f(x(\bar{\alpha}), y(\bar{\alpha})) ,\ x = v_x(\bar{\alpha}) ,\ y = v_y(\bar{\alpha}) . \tag{7}$$

In accordance with rules of partial derivatives calculation different approaches of pseudogradient calculation correspond to the expressions (5)–(7):
for relation (5)

$$\bar{\beta} = \frac{\partial \hat{J}}{\partial \bar{\alpha}} = \frac{\partial f}{\partial \bar{\alpha}} ; \tag{8}$$

for relation (6)

$$\bar{\beta} = \frac{d f}{dz} \frac{\partial z}{\partial \bar{\alpha}} ; \tag{9}$$

for relation (7)

$$\bar{\beta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \bar{\alpha}} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \bar{\alpha}} . \tag{10}$$

Let us analysis the possibilities of finding derivatives $\frac{\partial f}{\partial \bar{\alpha}}$, $\frac{d f}{dz}$, $\frac{\partial z}{\partial \bar{\alpha}}$, $\frac{\partial x}{\partial \bar{\alpha}}$ and $\frac{\partial y}{\partial \bar{\alpha}}$. Since the sample grid is discrete then the true finding of derivative $\frac{\partial f}{\partial \bar{\alpha}}$ is impossible. We can find its estimate by means of finite differences of the goal function. At that each component β_i of the pseudogradient $\overline{\beta}$ is determined separately through increments $\Delta_{\alpha i}$ of the relative i-th parameter:

$$\beta_i = \frac{\partial \mathrm{f}}{\partial \alpha_i} \approx \frac{\hat{\mathrm{J}}(Z_t, \hat{\alpha}_1, \ldots, \hat{\alpha}_i + \Delta_{\alpha i}, \ldots, \hat{\alpha}_m) - \hat{\mathrm{J}}(Z_t, \hat{\alpha}_1, \ldots, \hat{\alpha}_i - \Delta_{\alpha i}, \ldots, \hat{\alpha}_{m,})}{2\Delta_{\alpha i}}, \; i = \overline{1, m} \tag{11}$$

where Z_t – the local sample. Let us note, that for forming elements $\tilde{z}^{(1)}_{\bar{j}t}$ of the local sample (2) it is necessary to specify the deformations model and the kind of the reference image interpolation. However the requirements to their first derivatives existence are not laid.

If the derivative of the goal function on variable z exists, then the derivative $\frac{d\mathrm{f}}{dz}$ can be found analytically (or calculated by numerous methods) for explicit and implicit representation of the function. The partial derivative $\frac{\partial z}{\partial \overline{\alpha}}$ can not be found analytically, because the sample grid of images is discrete. We can estimate it in coordinates of each sample $\tilde{z}^{(1)}_{\bar{j}t}$, $\bar{j}_t \in \Omega_t$, through increments $\Delta_{\alpha i}$ of the corresponding i-th deformation parameter. Then in accordance with (9):

$$\beta_i \approx \frac{d\mathrm{f}}{dz} \frac{\sum_{\Omega_t} \left(\hat{s}(\bar{j}_t, \hat{\alpha}_1, \ldots, \hat{\alpha}_i + \Delta_{\alpha i}, \ldots, \hat{\alpha}_m) - \hat{s}(\bar{j}_t, \hat{\alpha}_1, \ldots, \hat{\alpha}_i - \Delta_{\alpha i}, \ldots, \hat{\alpha}_{m,}) \right)}{2\Delta_{\alpha i}}. \tag{12}$$

Another approach for finding the derivative estimate $\frac{\partial z}{\partial \overline{\alpha}}$ is the representation of z in the combined functional form $z = \mathrm{s}(x(\overline{\alpha}), y(\overline{\alpha}))$, then (9) is:

$$\overline{\beta} = \frac{d\mathrm{f}}{dz} \left(\frac{\partial z}{\partial x} \frac{\partial x}{\partial \overline{\alpha}} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial \overline{\alpha}} \right). \tag{13}$$

If for the given deformations model the requirements to its first derivatives on parameters existence are fulfilled then the partial derivatives $\frac{\partial x}{\partial \overline{\alpha}}$ and $\frac{\partial y}{\partial \overline{\alpha}}$ can be found analytically, and derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ can be estimated through finite differences of samples brightness (Minkina et al., 2007).

In the expression (10) there are derivatives $\frac{\partial x}{\partial \overline{\alpha}}$ and $\frac{\partial y}{\partial \overline{\alpha}}$, which were considered above, and also derivatives $\frac{\partial \mathrm{f}}{\partial x}$ and $\frac{\partial \mathrm{f}}{\partial y}$, their estimates can be found through finite differences on coordinate axes. Then:

$$\overline{\beta} = \frac{\hat{\mathrm{J}}\left(Z_t(x + \Delta_x), \hat{\overline{\alpha}}_{t-1}\right) - \hat{\mathrm{J}}\left(Z_t(x - \Delta_x), \hat{\overline{\alpha}}_{t-1}\right)}{2\Delta_x} \frac{\partial x}{\partial \overline{\alpha}} + \frac{\hat{\mathrm{J}}\left(Z_t(y + \Delta_y), \hat{\overline{\alpha}}_{t-1}\right) - \hat{\mathrm{J}}\left(Z_t(y - \Delta_y), \hat{\overline{\alpha}}_{t-1}\right)}{2\Delta_y} \frac{\partial y}{\partial \overline{\alpha}}. \tag{14}$$

where $Z_t(x \pm \Delta_x)$ – the local sample, the samples $\left\{\tilde{z}^{(1)}_{\bar{j}t}\right\}$ coordinates of which are shifted on the axis x by a value Δ_x, $\bar{j}_t \in \Omega_t$. As well as estimates of derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$, they are identical for all parameters to be estimated.

Thus, four approaches to calculate pseudogradient of the goal function, which are defined by expressions (11), (12), (13) and (14), are possible. Let us note, when usage of different approaches different requirements are laid to features of the goal function and deformations model.

We can obtain the estimate of interframe difference mean square at the next iteration, using local sample (2) and estimates $\hat{\bar{\alpha}}_{t-1}$ of parameters to be estimated, obtained at the previous iteration:

$$\hat{\mathrm{J}}_t = \frac{1}{\mu}\sum_{l=1}^{\mu}\left(\tilde{z}^{(1)}_{\bar{j}l} - z^{(2)}_{\bar{j}l}\right)^2 . \tag{15}$$

The estimate of interframe correlation coefficient is determined by equation of sample correlation coefficient calculation:

$$\hat{\mathrm{J}}_t = \frac{1}{\mu\hat{\sigma}_{\tilde{z}1}\hat{\sigma}_{z2}}\left(\sum_{l=1}^{\mu}\tilde{z}^{(1)}_{\bar{j}l} z^{(2)}_{\bar{j}l} - \mu\,\tilde{z}^{(1)}_{av}\tilde{z}^{(2)}_{av}\right), \tag{16}$$

where $\hat{\sigma}^2_{\tilde{z}1}$, $\hat{\sigma}^2_{z2}$ и $\tilde{z}^{(1)}_{av}$, $z^{(2)}_{av}$ - estimates of variances and mean values of $z^{(2)}_{\bar{j}t}$ and $\tilde{z}^{(1)}_{\bar{j}t}$, $\bar{j}_t \in \Omega_t$.

As an example let us find design expressions for calculation of the pseudogradient of interframe difference mean square through finite differences. At that for definition let us suppose, that the affine deformations model, containing parameters of rotation angle φ, scale coefficient κ and parallel shift $\bar{h} = \left(h_x, h_y\right)$ is used. Then coordinates (x, y) of the point on the image $\mathbf{Z}^{(1)}$ at vector $\bar{\alpha} = \left(h_x, h_y, \varphi, \kappa\right)^T$ of deformations transform to coordinates:

$$\left(\tilde{x} = x_0 + \kappa((x - x_0)\cos\varphi - (y - y_0)\sin\varphi) + h_x,\ \tilde{y} = y_0 + \kappa((x - x_0)\sin\varphi + (y - y_0)\cos\varphi) + h_y\right), \tag{17}$$

where (x_0, y_0) – rotation center coordinates. We use bilinear interpolation for a forecast of brightness in the point $(\tilde{x}, \tilde{y})$ from the image $\tilde{Z}^{(1)}$. Subject to accepted limitations let us concretize the methods for pseudogradient calculation. Let us note that these limitations are introduced for concretization of the obtained expressions and do not reduce consideration generality.

The first method. It is the least laborious way in calculus, where differentiation of the deformations model and the goal function is not used. The component β_{it} of the pseudogradient is calculated as normalized difference of two estimates of the goal function:

$$\beta_{it} = \frac{\sum_{l=1}^{\mu}\left(\tilde{z}^{(1)}_{\bar{j}l}\left(\hat{\alpha}_{i,t-1} + \Delta_{\alpha i}\right) - z^{(2)}_{\bar{j}l}\right)^2 - \sum_{l=1}^{\mu}\left(\tilde{z}^{(1)}_{\bar{j}l}\left(\hat{\alpha}_{i,t-1} - \Delta_{\alpha i}\right) - z^{(2)}_{\bar{j}l}\right)^2}{2\mu\Delta_{\alpha i}}, \tag{18}$$

where $\tilde{z}^{(1)}_{\bar{j}l}(\hat{\alpha}_{i,t-1} \pm \Delta_{\alpha i}) \in Z_t$ - brightness of the interpolated image in the point with coordinates $(\tilde{x}_l, \tilde{y}_l)$, determined by deformations model and current parameters estimates $\hat{\bar{\alpha}}_{t-1}$; $\bar{j}_l \in \Omega_t$ – samples coordinates $z^{(2)}_{\bar{j}l}$; $\Delta_{\alpha i}$ – increment of a parameter α_i to be estimated. In particular, for affine model (17) for shifts on the axis x, y, scale coefficient and rotation angle we obtain correspondingly:

$$
\begin{gathered}
\tilde{x}_l = x_0 + \hat{\kappa}_{t-1}((x_l - x_0)\cos\hat{\varphi}_{t-1} - (y_l - y_0)\sin\hat{\varphi}_{t-1}) + \hat{h}_{x,t-1} + \Delta_h, \\
\tilde{y}_l = y_0 + \hat{\kappa}_{t-1}((x_l - x_0)\sin\hat{\varphi}_{t-1} + (y_l - y_0)\cos\hat{\varphi}_{t-1}) + \hat{h}_{y,t-1}, \\
\tilde{x}_l = x_0 + \hat{\kappa}_{t-1}((x_l - x_0)\cos\hat{\varphi}_{t-1} - (y_l - y_0)\sin\hat{\varphi}_{t-1}) + \hat{h}_{x,t-1}, \\
\tilde{y}_l = y_0 + \hat{\kappa}_{t-1}((x_l - x_0)\sin\hat{\varphi}_{t-1} + (y_l - y_0)\cos\hat{\varphi}_{t-1}) + \hat{h}_{y,t-1} + \Delta_h, \\
\tilde{x}_l = x_0 + (\hat{\kappa}_{t-1} + \Delta_\kappa)((x_l - x_0)\cos\hat{\varphi}_{t-1} - (y_l - y_0)\sin\hat{\varphi}_{t-1}) + \hat{h}_{x,t-1}, \\
\tilde{y}_l = y_0 + (\hat{\kappa}_{t-1} + \Delta_\kappa)((x_l - x_0)\sin\hat{\varphi}_{t-1} + (y_l - y_0)\cos\hat{\varphi}_{t-1}) + \hat{h}_{y,t-1}, \\
\tilde{x}_l = x_0 + \hat{\kappa}_{t-1}((x_l - x_0)\cos(\hat{\varphi}_{t-1} + \Delta_\varphi) - (y_l - y_0)\sin(\hat{\varphi}_{t-1} + \Delta_\varphi)) + \hat{h}_{x,t-1}, \\
\tilde{y}_l = y_0 + \hat{\kappa}_{t-1}((x_l - x_0)\sin(\hat{\varphi}_{t-1} + \Delta_\varphi) + (y_l - y_0)\cos(\hat{\varphi}_{t-1} + \Delta_\varphi)) + \hat{h}_{y,t-1}.
\end{gathered}
\tag{19}
$$

Brightness of the sample $\tilde{z}^{(1)}_{\bar{j}l}(\hat{\alpha}_{i,t-1} \pm \Delta_{\alpha i})$ in the point $(\tilde{x}_l, \tilde{y}_l)$ is found, for example, by means of bilinear interpolation:

$$
\begin{aligned}
\tilde{z}^{(1)}_{\tilde{x}_l,\tilde{y}_l} = {} & z^{(1)}_{jx-,jy-} + (\tilde{x}_l - j_{x-})\left(z^{(1)}_{jx+,jy-} - z^{(1)}_{jx-,jy-}\right) + (\tilde{y}_l - j_{y-})\left(z^{(1)}_{jx-,jy+} - z^{(1)}_{jx-,jy-}\right) + \\
& + (\tilde{x}_l - j_{x-})(\tilde{y}_l - j_{y-})\left(z^{(1)}_{jx+,jy+} + z^{(1)}_{jx-,jy-} - z^{(1)}_{jx+,jy-} - z^{(1)}_{jx-,jy+}\right).
\end{aligned}
\tag{20}
$$

where $j_{x-} = \operatorname{int}\tilde{x}_l$, $j_{x+} = j_{x-} + 1$, $j_{y-} = \operatorname{int}\tilde{y}_l$, $j_{y+} = j_{y-} + 1$ – coordinates of nodes of the image $\mathbf{Z}^{(1)}$, nearby to the point $(\tilde{x}_l, \tilde{y}_l)$; $z^{(1)}_{jx\pm,jy\pm}$ – brightness in the corresponding nodes of thesample grid. Let us note that the expression (18) can be written in more handy form for calculations.

The second method is based on the analytical finding of derivative $\frac{d\mathrm{f}}{dz}$ and estimation through limited differences of derivative $\frac{\partial z}{\partial \bar{\alpha}}$. Subject to (12) and (15) we obtain:

$$
\beta_{it} \approx \frac{\sum_{l=1}^{\mu}\left(\tilde{z}^{(1)}_{\bar{j}l}(\hat{\alpha}_{i,t-1}) - z^{(2)}_{\bar{j}l}\right)\left(\tilde{z}^{(1)}_{\bar{j}l}(\hat{\alpha}_{i,t-1} + \Delta_{\alpha i}) - \tilde{z}^{(1)}_{\bar{j}l}(\hat{\alpha}_{i,t-1} - \Delta_{\alpha i})\right)}{\mu\Delta_{\alpha i}},
$$

where coordinates of interpolated samples $\tilde{z}^{(1)}_{\bar{j}l}(\hat{\alpha}_{i,t-1} \pm \Delta_{\alpha i})$ are found on the equations (19), and their brightness at bilinear interpolation – on the equation (20).

The third method assumes the existence derivative $\frac{d\mathrm{f}}{dz}$ and particular derivatives $\frac{\partial x}{\partial \overline{\alpha}}$ and $\frac{\partial y}{\partial \overline{\alpha}}$. Derivatives of brightness $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ on the base axis are estimated through finite differences. Then in accordance with (13):

$$\overline{\beta}_{it} \approx \frac{1}{\mu}\sum_{l=1}^{\mu}\left(\tilde{z}_{\tilde{j}l}^{(1)} - z_{\tilde{j}l}^{(2)}\right)\left(\frac{\tilde{z}^{(1)}_{\tilde{x}l+\Delta x,\,\tilde{y}l} - \tilde{z}^{(1)}_{\tilde{x}l-\Delta x,\,\tilde{y}l}}{\Delta_x}\frac{\partial x}{\partial \alpha_i} + \frac{\tilde{z}^{(1)}_{\tilde{x}l,\,\tilde{y}l+\Delta y} - \tilde{z}^{(1)}_{\tilde{x}l,\,\tilde{y}l-\Delta y}}{\Delta_y}\frac{\partial y}{\partial \alpha_i}\right), \tag{21}$$

where coordinates of samples $\tilde{z}^{(1)}_{\tilde{x}l\pm\Delta x,\,\tilde{y}l}$, $\tilde{z}^{(1)}_{\tilde{x}l,\,\tilde{y}l\pm\Delta y}$ are found in points $(\tilde{x}_l \pm \Delta_x, \tilde{y}_l)$, $(\tilde{x}_l, \tilde{y}_l \pm \Delta_y)$ of the image $\tilde{\mathrm{Z}}^{(1)}$. Derivatives $\frac{\partial x}{\partial \overline{\alpha}}$ and $\frac{\partial y}{\partial \overline{\alpha}}$ depend on the accepted deformations model. At affine model in the point $(\tilde{x}_l, \tilde{y}_l)$:

$$\begin{gathered}
\frac{\partial \tilde{x}_l}{\partial h_x} = 1,\ \frac{\partial \tilde{y}_l}{\partial h_x} = 0;\ \frac{\partial \tilde{x}_l}{\partial h_y} = 0,\ \frac{\partial \tilde{y}_l}{\partial h_y} = 1; \\
\frac{\partial \tilde{x}_l}{\partial \varphi} = \hat{\kappa}_{t-1}((x_l - x_0)\sin\hat{\varphi}_{t-1} - (y_l - y_0)\cos\hat{\varphi}_{t-1}), \\
\frac{\partial \tilde{y}_l}{\partial \varphi} = \hat{\kappa}_{t-1}((x_l - x_0)\cos\hat{\varphi}_{t-1} + (y_l - y_0)\sin\hat{\varphi}_{t-1}); \\
\frac{\partial \tilde{x}_l}{\partial \kappa} = (x_l - x_0)\cos\hat{\varphi}_{t-1} - (y_l - y_0)\sin\hat{\varphi}_{t-1}, \\
\frac{\partial \tilde{y}_l}{\partial \kappa} = (x_l - x_0)\sin\hat{\varphi}_{t-1} + (y_l - y_0)\cos\hat{\varphi}_{t-1}.
\end{gathered} \tag{22}$$

Having introduced denotations $\frac{\partial \tilde{x}_l}{\partial \alpha_i} = c_{il}$ and $\frac{\partial \tilde{y}_l}{\partial \alpha_i} = d_{il}$, for the i-th component of pseudogradient we can write:

$$\beta_{it} = \frac{1}{\mu}\sum_{l=1}^{\mu}\left(\tilde{z}_{\tilde{j}l}^{(1)} - z_{\tilde{j}l}^{(2)}\right)\left(\frac{\tilde{z}^{(1)}_{\tilde{x}l+\Delta x,\,\tilde{y}l} - \tilde{z}^{(1)}_{\tilde{x}l-\Delta x,\,\tilde{y}l}}{\Delta_x}c_{il} + \frac{\tilde{z}^{(1)}_{\tilde{x}l,\,\tilde{y}l+\Delta y} - \tilde{z}^{(1)}_{\tilde{x}l,\,\tilde{y}l-\Delta y}}{\Delta_y}d_{il}\right). \tag{23}$$

In the case if increments on coordinates are equal to the step of sample grid $\Delta_x = \Delta_y = 1$, then (23) takes a form

$$\beta_{it} = \frac{1}{\mu}\sum_{l=1}^{\mu}\left(\tilde{z}_{\tilde{j}l}^{(1)} - z_{\tilde{j}l}^{(2)}\right)\left(\left(\tilde{z}^{(1)}_{\tilde{x}l+1,\,\tilde{y}l} - \tilde{z}^{(1)}_{\tilde{x}l-1,\,\tilde{y}l}\right)c_{il} + \left(\tilde{z}^{(1)}_{\tilde{x}l,\,\tilde{y}l+1} - \tilde{z}^{(1)}_{\tilde{x}l,\,\tilde{y}l-1}\right)d_{il}\right).$$

Let us note, that the number of computational operations in the last expression can be reduced in the assumption of equality of derivatives on coordinates for the sample $\tilde{z}_{\tilde{j}l}^{(1)}$ of the image $\tilde{Z}^{(1)}$ and the sample $z_{\tilde{j}l}^{(2)}$ of the image $\mathbf{Z}^{(2)}$. This assumption is approximately fulfilled at small deviations $\hat{\overline{\alpha}}$ from the optimal value $\overline{\alpha}$. Тогда :

$$\beta_{it} = \frac{1}{\mu}\sum_{l=1}^{\mu}\left(\tilde{z}_{\tilde{j}l}^{(1)} - z_{jl}^{(2)}\right)\left(\left(z_{j_{xl}+1,\,j_{yl}}^{(2)} - z_{j_{xl}-1,\,j_{yl}}^{(2)}\right)c_{il} + \left(z_{j_{xl},\,j_{yl}+1}^{(2)} - z_{j_{xl},\,j_{yl}-1}^{(2)}\right)d_{il}\right).$$

The fourth method is based on the estimation of derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ through finite differences at analytic finding derivatives $\frac{\partial x}{\partial \overline{\alpha}}$ and $\frac{\partial y}{\partial \overline{\alpha}}$:

$$\beta_{it} = \frac{1}{2\mu}\left[\frac{1}{\Delta_x}\sum_{l=1}^{\mu}\left(\left(\tilde{z}_{\tilde{x}l+\Delta x,\,\tilde{y}l}^{(1)} - z_{jl}^{(2)}\right)^2 - \left(\tilde{z}_{\tilde{x}l-\Delta x,\,\tilde{y}l}^{(1)} - z_{jl}^{(2)}\right)^2\right)\frac{\partial x}{\partial \alpha_i} + \frac{1}{\Delta_y}\sum_{l=1}^{\mu}\left(\left(\tilde{z}_{\tilde{x}l,\,\tilde{y}l+\Delta y}^{(1)} - z_{jl}^{(2)}\right)^2 - \left(\tilde{z}_{\tilde{x}l,\,\tilde{y}l-\Delta y}^{(1)} - z_{jl}^{(2)}\right)^2\right)\frac{\partial y}{\partial \alpha_i}\right].$$

3. Improvement coefficient of parameters estimates

The convergence of parameters estimates of interframe deformations depends on a large number of influencing factors. We can divide them into a priory factors, which can be defined by probability density functions and autocorrelation functions of images and interfering noises, and a posteriori factors, determined by procedure (3) characteristics: pseudogradient calculation method, the kind of a gain matrix and number of iterations. As a rule, we can refer a goal function to the first group. For analysis it is desirable to describe the influence of the factors from the first group by a small number of values as far as possible. In the papers (Tashlinskii & Tikhonov, 2001) as such values it is proposed to use probabilities (4) of estimates change in parameters space. On their basis in the paper (Samojlov, 2006) a coefficient characterizing probabilistic characteristics of parameters change in the process of convergence is proposed. Let us consider it in details. If a value of parameter estimate at the ($t-1$)-th iteration is $\hat{\alpha}_{i,t-1}$, then the mathematical expectation of the estimate at the t-th iteration can be expressed through probabilities $\rho^+(\overline{\varepsilon})$ and $\rho^-(\overline{\varepsilon})$:

$$\mathrm{M}[\hat{\alpha}_{i,t}] = \hat{\alpha}_{i,t-1} - \lambda_{i,t}\left(\rho^+(\overline{\varepsilon}_{t-1}) - \rho^-(\overline{\varepsilon}_{t-1})\right).$$

If $\rho^+(\overline{\varepsilon}_{t-1}) > \rho^-(\overline{\varepsilon}_{t-1})$, then the estimate is improved, if not – is deteriorated. Thus the characteristic

$$\Re_i = \rho_i^+(\overline{\varepsilon}) - \rho_i^-(\overline{\varepsilon}) \tag{24}$$

let us call the estimate improvement coefficient. The range of its change is from –1 to +1. At that a value +1 means that the mathematical expectation $\mathrm{M}[\hat{\alpha}_{i,t}]$ of the estimate is improved at the *t*-th iteration by $\lambda_{i,t}$.

The improvement coefficient can be the generalized characteristic of images to be estimated, effecting noises and also chosen goal function. Having used for its calculation the equations (4), we obtain

$$\Re_i = \int_{-\infty}^{0} w(\beta_i)d\beta_i - \int_{0}^{\infty} w(\beta_i)d\beta_i . \tag{25}$$

Let us analyze possibilities for improvement coefficient calculation for the cases of usage as a goal function interframe difference mean square and interframe correlation coefficient. At that let us assume that $\rho_i^0(\bar{\varepsilon})=0$. The last assumption is true at unquantified samples of images to be studied. Subject to divisible group of events $\rho_i^+(\bar{\varepsilon})=1-\rho_i^-(\bar{\varepsilon})$, then

$$\Re_i = 2\rho_i^+(\bar{\varepsilon})-1 = 2\int_{-\infty}^{0} w(\beta_i)d\beta_i - 1 .$$

Interframe difference mean square

The estimate of interframe difference mean square at each iteration of estimation can be found on the relation (15). Let us assume the images to be studied have Gaussian distribution of brightness with zero mean and unquantified samples and the model of images $\tilde{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ is additive :

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

where $\tilde{S}^{(1)} = \left\{\tilde{s}_{\bar{j}}^{(1)}\right\}$, $\mathbf{S}^{(2)} = \left\{s_{\bar{j}}^{(2)}\right\}$ – desired random fields with identical variances σ_s^2, at that the field $\left\{s_{\bar{j}}^{(2)}\right\}$ has autocorrelation function $R(\ell)$; $\mathbf{\Theta}^{(1)} = \left\{\theta_{\bar{j}}^{(1)}\right\}$, $\mathbf{\Theta}^{(2)} = \left\{\theta_{\bar{j}}^{(2)}\right\}$ – independent Gaussian random fields with zero mean and equal variances σ_θ^2. Let us accept the affine model of deformations (17): $\bar{\alpha} = \left(h_x, h_y, \varphi, \kappa\right)^T$.

In accordance with (25) for calculation of the estimate improvement coefficient $\Re_i$ it is necessary to find probability density function $w(\beta_i)$ of projection β_i of pseudogradient $\bar{\beta}$ on the parameter α_i axis. For this purpose let us use the third way (21) for interframe difference mean square pseudogradient calculation.

Analytic finding probability distribution (23) as a function of σ_s^2, σ_θ^2 and $R(\ell)$ is a difficult problem. However the approximate solution can be found (Tashlinskii & Tikhonov, 2001), if we use the circumstance, that as μ increases the component β_i normalizes quickly. At $\mu=1$ (23) includes from four to eight similar summands, at $\mu=2$ – from eight to sixteen, etc. Thus the distribution of probabilities β_i can be assumed to be close to Gaussian. Then:

$$\Re_i(\bar{\varepsilon}) = 2F\left(\frac{M[\beta_i]}{\sigma[\beta_i]}\right) - 1 , \; i=\overline{1,m} , \tag{26}$$

where $F(\cdot)$ – Laplace function; $M[\beta_i]$ and $\sigma[\beta_i]$ – mathematical expectation and standard deviation of the component β_i. Thus the problem can be reduced to finding the mathematical expectation and variance of β_i. For relation (23) we obtain:

$$M[\beta_i] = -\sigma_s^2 \sum_{i=1}^{\mu}\left(\left(R\left(\ell_{a-1,b}^{(l)}\right) - \ell_{a+1,b}^{(l)}\right)c_{il} + \left(R\left(\ell_{a,b-1}^{(l)}\right) - R\left(\ell_{a,b+1}^{(l)}\right)\right)d_{il}\right); \tag{27}$$

$$\sigma^2[\beta_i] = \sigma_s^4 \sum_{l=1}^{\mu}\left(4\left(\left(c_{il}^2 + d_{il}^2\right)\left(\left(1 - R\left(\ell_{a,b}^{(l)}\right)\right)(1 - R(2)) + g^{-1}\left(2 - R\left(\ell_{a,b}^{(l)}\right) - R(2) + g^{-1}\right)\right) + \right.\right.$$

$$\left. + \left(c_{il}\left(R\left(\ell_{a-1,b}^{(l)}\right) - R\left(\ell_{a+1,b}^{(l)}\right)\right) + \; d_{il}\left(R\left(\ell_{a,b-1}^{(l)}\right) - R\left(\ell_{a,b+1}^{(l)}\right)\right)\right)^2\right), \tag{28}$$

where $\ell_{a,b}^{(l)}$ – Euclidian distance between point with coordinates (a_l, b_l) and point with coordinates (x_l, y_l), $l = \overline{1,\mu}$; $R(\ell_{a,b})$ – normalized autocorrelation function of the image; c_{il} and d_{il} – functions c and d for the i-ro parameter in the point (a_l, b_l) (Tashlinskii @ Minkina, 2006). For finding $\Re_i$ it is necessary to substitute (27) and (28) into (26).

As it is seen from (27) and (28) $\Re_i$ does not depend only on σ_s^2, σ_θ^2 and $R(\ell)$, it also depends on a plan of the local sample Z_t, namely on reciprocal location of samples (a_l, b_l), of the deformed image which are in the local sample at the t-th iteration.
In Fig. 1,a as an example the plots of the improvement coefficient for rotation angle ($\Re_\varphi$) as a function of error $\varepsilon_\varphi = \hat{\varphi} - \varphi^*$, where φ^* is the sought value of parameter are presented. The results are obtained for images with Gaussian autocorrelation function with correlation radius equal to 5 at signal/noise ration $g = 20$ and local sample size $\mu = 3$. At that it is supposed that coordinates of points of the local sample are chosen on the circle with radius $L = 20$ (curve 1) and $L = 30$ (curve 2) with the center, coinciding with rotation center.

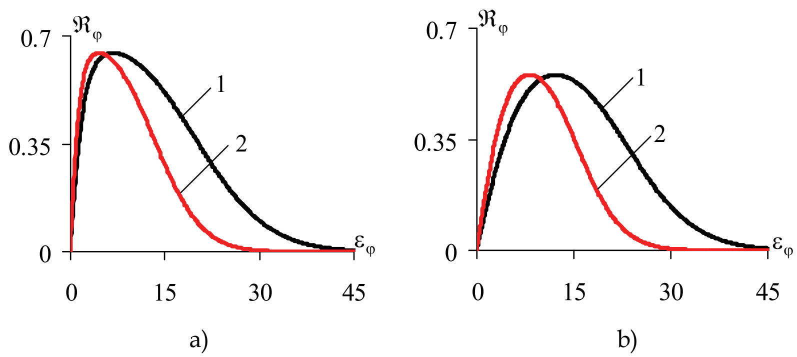

Fig. 1. The dependence of estimate improvement coefficient of rotation angle versus error

Interframe correlation sample coefficient

When choosing as a goal function interframe correlation coefficient its estimate at each iteration can be found on the relation (16). Having accepted for image to be studied the same assumptions as in the previous case for finding $w(\beta_i)$ let us use the expression:

$$\begin{aligned}\beta_{it} = \frac{1}{2\mu^2\hat{\sigma}_{z1}^3\hat{\sigma}_{z2}}\Bigg[\mu\hat{\sigma}_{z1}^2\left(\sum_{l=1}^{\mu}\left(z_{\bar{j}l}^{(2)} - z_{cp}^{(2)}\right)\times\right.\\ \left.\times\left(\frac{\tilde{z}_{\tilde{x}l+\Delta x,\tilde{y}l}^{(1)} - \tilde{z}_{\tilde{x}l-\Delta x,\tilde{y}l}^{(1)}}{\Delta_x}\frac{\partial\tilde{x}}{\partial\alpha_i} + \frac{\tilde{z}_{\tilde{x}l,\tilde{y}l+\Delta y}^{(1)} - \tilde{z}_{\tilde{x}l,\tilde{y}l-\Delta y}^{(1)}}{\Delta_y}\frac{\partial\tilde{y}}{\partial\alpha_i}\right)\right) - \\ -\left(\sum_{l=1}^{\mu}\left(z_{\bar{j}l}^{(2)} - z_{cp}^{(2)}\right)\tilde{z}_{\tilde{x}l,\tilde{y}l}^{(1)}\right)\sum_{l=1}^{\mu}\left(\tilde{z}_{\tilde{x}l,\tilde{y}l}^{(1)} - \frac{1}{\mu-1}\sum_{k=1,k\neq l}^{\mu}\tilde{z}_{\tilde{x}k,\tilde{y}k}^{(1)}\right)\times\\ \times\left(\frac{\tilde{z}_{\tilde{x}l+\Delta x,\tilde{y}l}^{(1)} - \tilde{z}_{\tilde{x}l-\Delta x,\tilde{y}l}^{(1)}}{\Delta_x}\frac{\partial\tilde{x}}{\partial\alpha_i} + \frac{\tilde{z}_{\tilde{x}l,\tilde{y}l+\Delta y}^{(1)} - \tilde{z}_{\tilde{x}l,\tilde{y}l-\Delta y}^{(1)}}{\Delta_y}\frac{\partial\tilde{y}}{\partial\alpha_i}\right)\Bigg],\end{aligned} \qquad (29)$$

where coordinates and brightness of samples $\tilde{z}^{(1)}_{\tilde{x}l\pm\Delta x,\,\tilde{y}l}$, $\tilde{z}^{(1)}_{\tilde{x}l,\,\tilde{y}l\pm\Delta y}$ are determined by equations (19) and (20) correspondingly. The derivatives $\dfrac{\partial\tilde{x}}{\partial\alpha_i}$ and $\dfrac{\partial\tilde{y}}{\partial\alpha_i}$ can be found on equations (22).

Let us consider several cases. At first let us suppose that mean values $z^{(2)}_{av}$ and $\tilde{z}^{(1)}_{av}$ are equal to zero, and estimates of standard deviation $\hat{\sigma}_{z1}$ and $\hat{\sigma}_{z2}$ are known a priory. In this case pseudogradient of interframe correlation coefficient differs from pseudogradient of covariation estimate of images $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ only by constant factor $(\sigma_{z1}\sigma_{z2})^{-1}$, and $M[\beta_i]$ from expression (27) – by factor $-\left(2\mu\sigma_s^2\right)^{-1}$. The variance of β_i :

$$\sigma^2[\beta_i]=\frac{1}{\mu^2}\sum_{l=1}^{\mu}\left(\frac{1}{2}\left(c_{il}^2+d_{il}^2\right)\left(1+g_1^{-1}\right)\left(1-R(2)+g_2^{-1}\right)+\right.$$
$$\left.+\frac{1}{4}\left(c_{il}\left(R\left(\ell^{(l)}_{a-1,b}\right)-R\left(\ell^{(l)}_{a+1,b}\right)\right)+d_{il}\left(R\left(\ell^{(l)}_{a,b-1}\right)-R\left(\ell^{(l)}_{a,b+1}\right)\right)\right)^2\right),$$

where $g_1=\sigma_{s1}^2/\sigma_{\theta1}^2$, $g_2=\sigma_{s2}^2/\sigma_{\theta2}^2$ – signal/noise ratio for images $\tilde{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ correspondingly.

If only variances σ_{z1}^2, σ_{z2}^2 are known a priory, then

$$\beta_i=\frac{1}{\mu\sigma_{z1}\sigma_{z2}}\sum_{l=1}^{\mu}\frac{\partial z^{(1)}_{\tilde{x}_l,\tilde{y}_l}}{\partial\alpha_i}\left(\frac{\mu-1}{\mu}z^{(2)}_{jl}-\frac{1}{\mu}\sum_{k=1,k\neq l}^{\mu}z^{(2)}_{jk}\right).$$

For simplification of finding the mathematical expectation and the variance of β_i summands in the sum $\sum\limits_{k=1,k\neq l}^{\mu}z^{(2)}_{jk}$ let us assume to be noncorrelated with a value $z^{(2)}_{jl}$. The assumption aboud noncorrelatedness of $z^{(2)}_{jk}$, $k=\overline{1,\mu}$, $k\neq l$ and $z^{(2)}_{jl}$ is not rigid, because samples to the local sample are chosen as a rule to be weakly correlated (Tashlinskii @ Minkina, 2006). Then:

$$M[\beta_i]=\frac{1}{2\mu}\left(1-\frac{1}{\mu}\right)\sum_{l=1}^{\mu}\left(c_{il}\left(R\left(\ell^{(l)}_{a-1,b}\right)-R\left(\ell^{(l)}_{a+1,b}\right)\right)+d_{il}\left(R\left(\ell^{(l)}_{a,b-1}\right)-R\left(\ell^{(l)}_{a,b+1}\right)\right)\right), \tag{30}$$

$$\sigma^2[\beta_i]=\frac{1}{\mu^2}\sum_{l=1}^{\mu}\left(\frac{1}{2}\left(c_{il}^2+d_{il}^2\right)\left(\left(1-\frac{1}{\mu}\right)(1-R(2))+g_1^{-1}\left(1-R(2)+g_2^{-1}\right)+g_2^{-1}\right)+\right.$$
$$\left.+\frac{1}{4}\left(1-\frac{1}{\mu}\right)^2\left(c_{il}\left(R\left(\ell^{(l)}_{a-1,b}\right)-R\left(\ell^{(l)}_{a+1,b}\right)\right)+d_{il}\left(R\left(\ell^{(l)}_{a,b-1}\right)-R\left(\ell^{(l)}_{a,b+1}\right)\right)\right)^2\right). \tag{31}$$

As an example in Fig. 1,b the plots of the improvement coefficient for rotation angle ($\Re_\varphi$) as the function of errors ε_φ are presented. Image parameters, signal/noise ration and sample size correspond to the example (Fig. 1,a). From the plot it is seen, that at similar conditions $\Re_\varphi$ for interframe correlation coefficient is less, than for interframe difference mean square.

Similarly the case when σ_{z1}^2, σ_{z2}^2 of $z_{av}^{(2)}$ and $\tilde{z}_{av}^{(1)}$ are a priory known can be considered.

4. Optimization of samples choice region on criterion of estimate improvement coefficient maximum

Let one parameter is estimated. Then for finding optimal region of samples of a local sample we can use results, obtained in the previous part, choosing the estimate improvement coefficient maximum of a parameter to be estimated as an optimality criterion. At the affine deformations model the improvement coefficient for parameters h_x and h_y depends only on their errors from optimal values and does not depend on the location of samples of the local sample. Thus when estimating parallel shift parameters it is impossible to improve estimates convergence at the expense of choice of samples of the local sample. For parameters of rotation φ and scale κ the improvement coefficient depends on the samples coordinates. Correspondingly, the region of image, where the improvement coefficient maximum is ensured, can be found.

As an example let us find the image region at the known error ε_φ of rotation angle φ. It is not difficult to show, that initial region at the given image parameters is determined by the distance L_{op} from the rotation centre (x_0, y_0). At that for each error ε_φ a value L_{op} will be individual. In Fig. 2 for $\varepsilon_\varphi = 5^o$ the dependences of $\Re_\varphi$ as the function versus distance L from the rotation centre when using as the goal function interframe difference mean square (рис. 2,a) and interframe correlation coefficient (рис. 2,b), calculated by equations (26), (28) and (30), (31) correspondingly. The image was assumed to be Gaussian with autocorrelation function with correlation radius equal to 13 steps of the sample grid and signal/noise ratio $g = 50$. From the plot it is seen that for interframe difference mean square maximum of estimation coefficient attainess at $L_{op} = 58$, for interframe correlation coefficient– at $L_{op} = 126.7$.

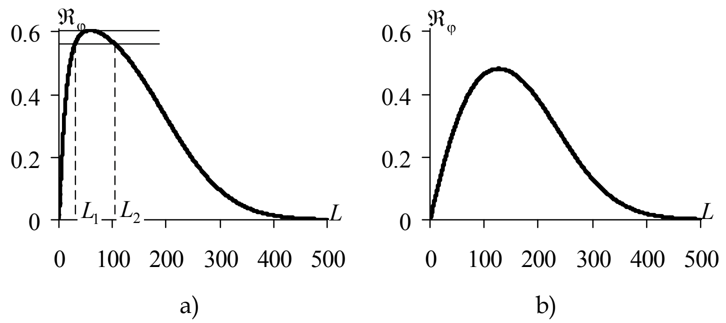

Fig. 2. The dependence of estimate improvement coefficient of rotation angle versus the distance from rotation center

If we know the dependence of ε_φ change versus the number of iterations, we can find L_{op} for each iteration. The rule of forming the dependence ε_φ on the number of iterations can be different and can depend on conditions of the problem to be solved. For instance, if we use a minimax approach, then it is enough to find the dependence beginning from maximum possible parameter error (for the worst situation), and to find the number of iterations which is necessary for the given accuracy attainment. In the sequel the obtained rule of L_{op} change on the number of iterations is applied for any initial parameter error, ensuring the estimation accuracy which is not worse than the given one. At that the dependence of change of error ε_φ versus the number of iterations can be found either theoretically for the given autocorrelation function and probability density function of image brightness by the method of pseudogradient procedures simulation at finite number of iterations (Tashlinskii & Tikhonov, 2001), or experimentally on the current estimates, averaged on the given realizations assemblage. At the last approach the following algorithm can be used.

1^0. To specify the initial error $\varepsilon_{\varphi 0}$ of rotation angle.

2^0. To find L_{op1} for the first iteration.

3^0. To perform the iteration u times. On the obtained estimates to find the average error $\varepsilon_{\varphi 1} = \frac{1}{u}\sum_{r=1}^{u}\varepsilon_{\varphi 0,r}$, where u – given number of realizations.

4^0. For obtained $\varepsilon_{\varphi 1}$ to repeat operators 1^0–3^0 T times until the next $\varepsilon_{\varphi T}$ is less the required estimation error ε_{nop}, T – total number of iterations.

Let us notice, that for digital images the circle with radius L_{op} can be considered as an optimal region only conditionally, because the probability of its intersection with nodes of sample grid is too small. To obtain the suboptimal region we can specify some range of acceptable values for the improvement coefficient from $\gamma\Re_{\max}$ to $\Re_{\max}$ (Fig. 2,a), where γ – threshold coefficient. Then values L_1 and L_2 specify region bounds, where the improvement coefficient does not differ from maximum more than, for example, 10%. At that the suboptimal region is ring. As an example in Fig. 3 the dependences L_{op} versus error ε_φ for interframe difference mean square (curve 1) and interframe correlation coefficient (curve 2) are shown. From figure it is seen that values of L_{op} for correlation coefficient exceed values of L_{op} for difference mean square.

In Fig. 4 the dependence ε_φ versus the number of iterations (curve 1), obtained at usage of pseudogradient procedure with parameter $\lambda_\varphi = 0.15^o$, initial error $\varepsilon_{\varphi 0} = 45^o$ and choice of samples of the local sample from 10 % suboptimal region on the image of size 1024×1024 pixels are presented. On the same figure the dependences for $\varepsilon_{\varphi 0} = 25^o$ (curve 2) and $\varepsilon_{\varphi 0} = 15^o$ (curve 3), obtained at the same rule of suboptimal region change and dependence obtained for $\varepsilon_{\varphi 0} = 45^o$ without optimization (curve 4) are shown. All curves are averaged on 200 realizations. It is seen that optimization increases the convergence rate of rotation angle about several times. At initial errors which are less the maximum errors(curves 2 and 3), the convergence rate of estimate is a little less, than at maximum one, but the number of iterations which is necessary for the given error attainment does not exceed the number of iterations at maximum error. The behavior of estimates for scale coefficient estimation is similar.

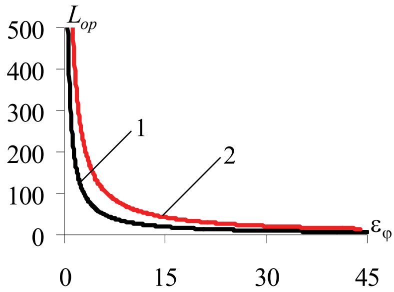

Fig. 3. The dependence of L_{op} versus ε_φ

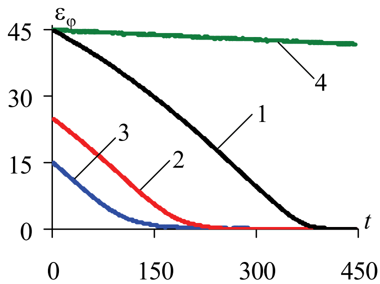

Fig. 4. The dependence of ε_φ versus the number of iterations

Let us note, that the considered method of optimization of the local sample samples choice region is unacceptable when estimating a parameters vector. It is due to the fact that the improvement coefficient of parameters vector can not be found on estimates improvement coefficient of separate parameters. Thus let us consider another approach for the case when estimating a vector of parameters.

5. Optimal Euclidian error distance of deformations parameters estimates

For any set of deformations model parameters as a result of the next iteration performing for the sample $z_{jk}^{(2)}$ with coordinates $\left(j_{xk}, j_{yk}\right)$ its estimate $\tilde{z}_k^{(1)}$ is found on the reference image with coordinates $(\tilde{x}_k, \tilde{y}_k)$. At that the location of the point $(\tilde{x}_k, \tilde{y}_k)$ relatively the point $\left(j_{xk}, j_{yk}\right)$ can be defined through Euclidian error distance (EED) $\Re = \sqrt{\left(j_{xk} - \tilde{x}_k\right)^2 + \left(j_{yk} - \tilde{y}_k\right)^2}$ and angle $\phi = \arg tg \dfrac{j_{yk} - \tilde{y}_k}{j_{xk} - \tilde{x}_k}$ (Fig. 5). It can be shown that if only rotation angle is estimated then in different regions maximum EED attains at different values of estimate error, but at the same EED value. It is explained in Fig. 6. What is more when estimating any another parameter (scale, shift on one of the axis) or their set maximum EED attains at the same EED. We can suppose, that this optimal value of EED depends only on the goal function and characteristics of images to be studied and does not depend on the model of deformations.

On the other hand the optimal value of EED at the known error of parameters estimates determines the optimal region of samples for the local sample. Thus the solution of the problem of finding of optimal (suboptimal) region of samples of local sample can be divided into two steps:
1) finding for the chosen goal function of estimation quality optimal EED as a function of image parameters (probability density function of brightness, autocorrelation function of desired image and signal/noise ratio);
2) determining on the deformations model and a vector of parameters estimates error the optimal region of choice of samples of the local sample as a region in which the optimal EED is ensured.

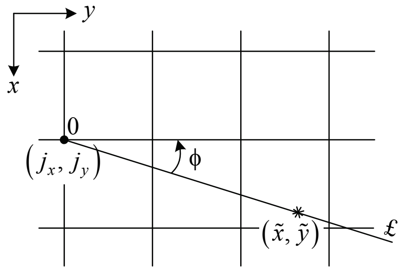

Fig. 5. Illustration of points (j_x, j_y) and $(\tilde{x}, \tilde{y})$ location

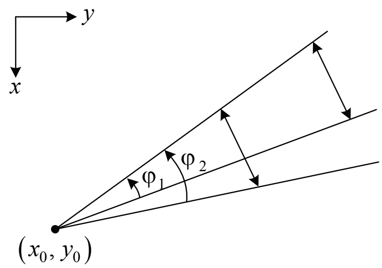

Fig. 6. Location of optimal EED in dependence on rotation angle

Let us consider the solution of the first mentioned problems. Let us the goal function of estimation quality is given. It is required to find value of EED, when maximum information about reciprocal deformation of images $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ is extracted. Let us understand the quantity of information in the sense of information, contained in one random value respectively another random value.
The estimate of the goal function gradient is calculated on the local sample, containing μ samples pairs. Each pair of samples $z^{(2)}_{\bar{j}k}$ and $\tilde{z}^{(1)}_k$, $k=\overline{1,\mu}$, of the local sample has desired information about contact degree of these samples. At that all pairs of samples are equal on average, thus bellow we will consider one pair.
Assuming the image to be isotropic, for simplification of analysis of influence of distance between samples $z^{(2)}_{\bar{j}k}$ and $\tilde{z}^{(1)}_k$ on the features of the goal function estimate it is reasonable to amount the problem to one-dimensional problem. For that it is enough to specify the

coordinate axis $0-l$, passing through coordinates of samples with the centre in the point $\left(j_{xk}, j_{yk}\right)$ (Fig. 5). Correspondingly the literal notations for samples are simplified: $z=z_{\bar{j}k}^{(2)}$, $\tilde{z}_k^{(1)}=\tilde{z}_£$, where £ – the distance between samples.

As it was already noticed, the information about contact degree of samples z and $\tilde{z}_£$ is noisy. For the additive model of image observations: $z=s+\theta$, $\tilde{z}_£=\tilde{s}_£+\tilde{\theta}_£$, the noise component is caused by two factors: additive noises θ, $\tilde{\theta}_£$ and sample correlatedness. The influence of noncorrelated noises is equal for any sample location. As the distance between them increases the random component increases too. Thus the noise component is minimum if the coordinates of samples coincide, correspondingly in this case the correlatedness is maximum. Actually, let us assume that variances of the samples $z=s+\theta$ and $\tilde{z}=\tilde{s}+\tilde{\theta}$ are equal, and

$$\sigma_s^2=\sigma_{\tilde{s}}^2,\ \sigma_\theta^2=\sigma_{\tilde{\theta}}^2, \tag{32}$$

for the mathematical expectation and variance of difference $z-\tilde{z}_£$ square we obtain:

$$\mathrm{M}\left[(z-\tilde{z}_£)^2\right]=\mathrm{M}\left[\left(s+\theta-\tilde{s}_£-\tilde{\theta}_£\right)^2\right]=2\sigma_s^2\left(1-R(£)+g^{-1}\right),$$

$$\mathrm{D}\left[(z-\tilde{z}_£)^2\right]=\mathrm{M}\left[(z-\tilde{z}_£)^4\right]-\mathrm{M}^2\left[(z-\tilde{z}_£)^2\right]=8\sigma_s^4\left(1-R(£)+g^{-1}\right)^2,$$

where $R(£)$ – normalized autocorrelation function of images to be studied; $g=\dfrac{\sigma_s^2}{\sigma_\theta^2}$ – signal/noise ratio. The plots normalized to σ_s^2 for the mathematical expectation and the mean-square distance of $(z-\tilde{z}_£)^2$ as the function of £ at $g=20$ and Gaussian $R(£)$ with correlation radius, equal to 5 steps of the sample grid, are given in Fig. 7 and Fig. 8 correspondingly.

For the mathematical expectation and variance of the product $z\tilde{z}_£$ correspondingly we obtain:

$$\mathrm{M}[z\tilde{z}_l]=\mathrm{cov}\left[\left(s+\theta\right)\left(\tilde{s}_£+\tilde{\theta}_£\right)\right]=\sigma_s^2 R(£),$$

$$\mathrm{D}[z\tilde{z}_£]=\sigma_s^4\left(\left(1+g^{-1}\right)^2+R^2(£)\right).$$

The normalized plots of covariation and mean-square distance $(z\tilde{z}_£)$ as a function of £ at the same image parameters are shown in Fig. 9 and Fig. 10 correspondingly. Let us notice that according to the assumptions (32) the normalized covariation, namely the correlation coefficient between samples z and $\tilde{z}$, is determined by the expression:

$$r(£)=\frac{\mathrm{cov}\left[(s+\theta)\left(\tilde{s}_£+\tilde{\theta}_£\right)\right]}{\mathrm{D}[s+\theta]}=\frac{R(£)}{1+g}.$$