Effectiveness of correlation and information measures for synthesis of recurrent algorithms for estimating spatial deformations of video sequences

A.G. Tashlinskiy[1], A.V. Zhukova[1]

[1]Ulyanovsk State Technical University, Severnii Venets, 32, 432027, Ulyanovsk, Russia

Abstract

A comparative analysis of the efficiency of correlation (cross-correlation, Tanimoto coefficient and Kendall's rank correlation coefficient) and information (mutual information of Tsallis and Shannon, F-information measure and entropy of the joint probability distribution) measures of image similarity for the synthesis of recursive estimation algorithms is presented for the problem of estimating parameters of spatial deformations of a sequence of images. Unbiased additive Gaussian noise was used as an interfering factor in the experimental studies. It is shown that the potentially high convergence rate of the estimated parameters and the smaller variance of the estimation error from the investigated correlation measures are ensured by the Tonimoto coefficient, and from the I-information of the F-information among the information measures. According to these criteria, the Kendall's rank correlation coefficient and the M-measure of F-information are inferior, respectively.

Keywords: image; recurrent estimation; similarity measures; cross-correlation; Tanimoto coefficient; Kendall's rank correlation coefficient; Tsallis mutual information; Shannon mutual information; F-information measures

1. Introduction

The detection and evaluation of spatial changes (deformations) in a sequence of images $\mathbf{Z}^{(n)}$, $n = \overline{1, N}$ is one of the key tasks of processing video sequences. Various approaches to the solution of this problem are implemented in the frequency [1] and in the spatial [2] domains.

In this case, the solution is usually based on the search for an extremum of some similarity measure (SM) between each pair of adjacent images $\mathbf{Z}^{(n)} = \left\{ z_j^{(n)} \right\}$ and $\mathbf{Z}^{(n+1)} = \left\{ z_j^{(n+1)} \right\}$, where $\bar{j}$ − coordinates of the nodes of the grid of samples on which the images are defined. With the assumed model of interframe spatial deformations of images, estimates are sought for the strain parameters $\bar{\alpha}$ of one of the images at which the extremum of the SM is reached.

When recursively searching for the SM extremum at each iteration t of the estimation the current estimates $\bar{\alpha}$ of the spatial deformation parameters are corrected by a certain amount in the direction $\bar{\mathrm{d}}\left(\hat{\bar{\alpha}}, Z_t\right)$ [3]:

$$\hat{\bar{\alpha}}_{t+1} = \hat{\bar{\alpha}}_t - \mathbf{\Lambda}_t \, \bar{\mathrm{d}}_t\left(\hat{\bar{\alpha}}, Z_t\right), \tag{1}$$

where $\mathbf{\Lambda}_t$ — a positively defined matrix (usually diagonal); $t = \overline{1, T}$ — iteration number.

The direction $\bar{\mathrm{d}}(\cdot)$ is defined by the SM gradient estimated using a small subsample [4] $Z_t = \left\{ \tilde{z}_{\bar{j}_t}^{(n)}, z_{\bar{j}_t}^{(n+1)} \right\}$ of (usually random) points $\tilde{z}_{\bar{j}_t}^{(n)} \in \tilde{\mathbf{Z}}^{(n)}$ and $z_{\bar{j}_t}^{(n+1)} \in \mathbf{Z}^{(n+1)}$ drawn on t-th iteration, where $\tilde{\mathbf{Z}}^{(n)}$ - interpolated image obtained using current deformation parameters' estimates $\hat{\bar{\alpha}}_t$ of image $\mathbf{Z}^{(n)}$ [5]. In order to improve the accuracy of parameters' estimates $\hat{\boldsymbol{\alpha}}$ we need to increase a number of points in the subsamples but it will lead to increase in computational time.

There are many SM for images [6]. The choice of a particular SM is determined by the conditions and requirements of the applied problem, the nature of the possible spatial deformations of the video sequence, the properties of the images and the interfering factors.

2. Problem formulation

The chosen SM largely determines the potential effectiveness of the recursive algorithms for estimating spatial deformations in a sequence of images synthesized on its basis. However, criteria that allow a priori evaluation of the potential efficiency of algorithms by SM are poorly investigated. A variant of the solution of this problem was considered in [7,8]. In this work several correlation correlations have been selected for the study: cross-correlation (the coefficient of interframe correlation) [9], the Tanimoto coefficient [10] the Kendall's rank correlation coefficient [11], and a number of informational SMs: Tsallis [12] and Shannon mutual information, F-information measures, and the entropy of the joint probability distribution [13]. Unbiased additive Gaussian noise was used as an interfering factor in the studies. The relation between SM characteristics and the probabilistic properties of estimates of the parameters of spatial deformations formed by recurrent algorithms synthesized on the basis of these measures was also investigated.

A method for calculating the probabilistic properties of estimates of the parameters of inter-frame spatial deformations of images for a finite number of iterations of recurrent estimation [14] was proposed in [16]. It is based on finding the probabilities of demolition at each iteration (the probabilities of changing the estimates of the investigated parameters $\hat{\bar{\alpha}}$ towards optimal

CEUR-WS.org/Vol-1901/paper37.pdf

values). With allowance for (1) the probability of demolishing of the estimates at the next iteration can be interpreted as the probability that the projection of the gradient of the SM gradient on the axis of this parameter will be negative. In [15] this characteristic was used to find the error in estimates of the parameters of inter-frame spatial deformations of images formed by relay procedures of the form (1). However, this approach can be used for recurrent procedures of other classes.

Since procedure (1) is based on the use of SM gradient estimates its capabilities are largely determined by the nature of the slope K of the SM used in the synthesis of the procedure. Fig. 1 shows examples of the dependence of the slope of the SMs chosen for the study on the magnitude of the parallel shift h of the images, where $h = \overline{0 - 50}$ is the shift in the grid steps of the image counts. Curves of the curvature module of the SM for correlation measures are shown on the left graph, where curve 1 corresponds to the Kendall rank correlation coefficient, curve 2 - to the Tanimoto coefficient, curve 3 - to the interframe correlation coefficient (CC). The right graph shows the dependencies for information SMs, here curve 1 corresponds to Shannon MI, curve 2 - to Tsallis MI, curve 3 - to I-measure of F-information, curve 4 - to M-measure of F-information, curve 5 - to excluding F-information, curve 6 - the entropy of the joint probability distribution (JPD). The legend is the same for all other figures in this work.

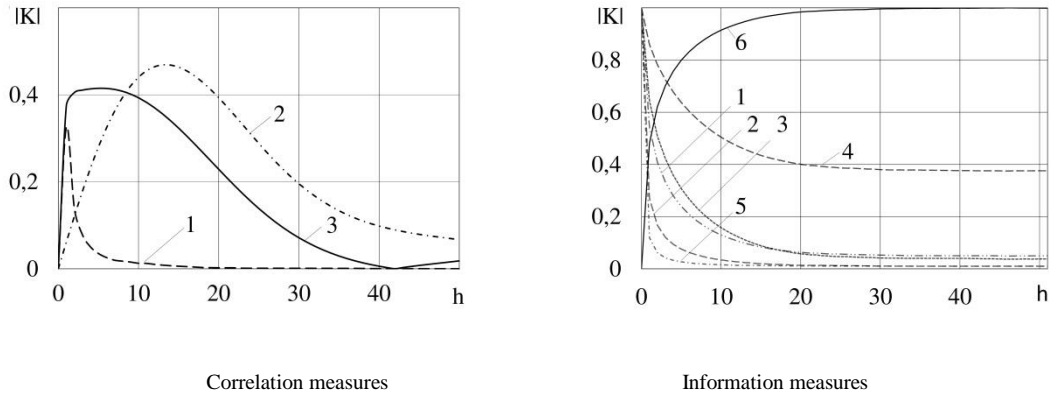Correlation measures　　　　　　　　Information measures

Fig. 1. Slope of the studied SM.

From Fig. 1 it can be seen that the nature of the slope of different SMs differs significantly, which must affect the potential characteristics of the recurrent evaluation procedures synthesized on their basis. A number of efficiency criteria based on an analysis of the nature of the SM slope was proposed in [8]. We will consider these criteria with reference to the correlation and information SMs chosen for the study.

3. Studied image similarity measures

3.1. Probability-based measures

CC (cross-correlation) [9] is one of the most popular similarity measures. CC can be defined as

$$r = \frac{1}{\mu \hat{\sigma}_{z_n} \hat{\sigma}_{z_{n+1}}} \sum_{\bar{j}_t \in \Omega_t} \left(\widetilde{z}_{\bar{j}_t}^{(n)} - M[\widetilde{\mathbf{Z}}^{(n)}] \right) \left(z_{\bar{j}_t}^{(n+1)} - M[\mathbf{Z}^{(n+1)}] \right),$$

where μ — the number of points in a subsample; $M[\mathbf{Z}] = \sum_{\bar{j}_t \in \Omega_t} z_{\bar{j}} \big/ \mu$ — image mean estimation $\mathbf{Z}$; $\hat{\sigma}^2_z = \sum_{\bar{j}_t \in \Omega_t} \left(z_{\bar{j}} - M[\mathbf{Z}] \right)^2 \big/ \mu$ — estimation of image $\mathbf{Z}$ variance. Correlation coefficient r changes between -1 and $+1$. The value $r = +1$ means the full linear relation, $r = -1$ – negative linear relation. If r is different to ± 1 then the relation of corresponding pixels $\widetilde{z}^{(n)}$ and $z^{(n+1)}$ can be described as: $z_{\bar{j}_t}^{(n+1)} = \frac{\hat{\sigma}_{z_n}}{\hat{\sigma}_{z_{n+1}}} \left(\widetilde{z}_{\bar{j}_t}^{(n)} - M[\widetilde{\mathbf{Z}}^{(n)}] \right) + M[\mathbf{Z}^{(n+1)}]$ the measure of relation linearity between $\mathbf{Z}^{(n)}$ and $\mathbf{Z}^{(n+1)}$ which allows us to efficiently use correlation coefficient in case of additive noise or linear intensity distortions. In terms of computational complexity, CC is very effective, as it requires a small number of additions and multiplications for each element in the subsample. It is in the order of μ.

Tanimoto coefficient [10] between two images $\widetilde{\mathbf{Z}}^{(n)}$ and $\mathbf{Z}^{(n+1)}$ is defined as:

$$S_T = \frac{\sum\limits_{\bar{j}_t \in \Omega_t} \widetilde{z}_{\bar{j}_t}^{(n)} z_{\bar{j}_t}^{(n+1)}}{\sum\limits_{\bar{j}_t \in \Omega_t} \widetilde{z}_{\bar{j}_t}^{(n)} z_{\bar{j}_t}^{(n+1)} + \sum\limits_{\bar{j}_t \in \Omega_t} \left(\widetilde{z}_{\bar{j}_t}^{(n)} - z_{\bar{j}_t}^{(n+1)} \right)^2} . \tag{2}$$

In comparison with CC in Tanimoto coefficient normalization of intensity multiplication with respect to their standard deviations is replaced by the normalization with respect to the sum of squared differences between corresponding sample counts, which effects in the same way. Using the inner product of intensities in the denominator of Tanimoto coefficient gives

the same effect as the normalization with respect to mean values of images. Computational complexity of Tanimoto coefficient is on the same order as for correlation coefficient.

Kendall rank correlation coefficient [11]. If $\widetilde{z}_j^{(n)}$ and $z_j^{(n+1)}$ are intensities of corresponding image pixels then for their differences $(\widetilde{z}_{\overline{i}}^{(n)} - \widetilde{z}_{\overline{j}}^{(n)})$ and $(z_{\overline{i}}^{(n+1)} - z_{\overline{j}}^{(n+1)})$ when $\overline{i} \neq \overline{j}$ there are two possible situations: concordance - when $\text{sign}(\widetilde{z}_{\overline{i}}^{(n)} - \widetilde{z}_{\overline{j}}^{(n)}) = \text{sign}(z_{\overline{i}}^{(n+1)} - z_{\overline{j}}^{(n+1)})$ and mismatch - when $\text{sign}(\widetilde{z}_{\overline{i}}^{(n)} - \widetilde{z}_{\overline{j}}^{(n)}) = -\text{sign}(z_{\overline{i}}^{(n+1)} - z_{\overline{j}}^{(n+1)})$. If we take a large sample from images $\widetilde{\mathbf{Z}}^{(n)}$ and $\mathbf{Z}^{(n+1)}$ and the number of concordances is greater than the number of mismatches then we can conclude that image intensities are bounded. Let assume that from $\mu/2$ pixel pairs are concordances and N_d are mismatches then Kendall correlation coefficient can be defined as follows [9]:

$$\tau = \frac{2(N_c - N_d)}{\mu(\mu - 1)}. \tag{3}$$

Kendal correlation coefficient is one of the most complex measures in terms of computational complexity. It requires concordance and mismatch computations for $\mu(\mu - 1)/2$ corresponding pixel pairs. Therefore, its computational complexity of τ is on the order of μ^2 operations.

3.2. Information measures

Shannon mutual information (MI) is one of the most widely used similarity measure in image registration [12,13] as it provides an extremely high accuracy when images have linear and non-linear intensity distortions, occlusions and also in case of additive noise and multimodal images. Generalized Shannon MI can be defined in terms of entropy as:

$$\text{J}\left(\hat{\overline{\alpha}}, \widetilde{\mathbf{Z}}^{(n)}, \mathbf{Z}^{(n+1)}\right) = H\left(\widetilde{\mathbf{Z}}^{(n)}\right) + H\left(\mathbf{Z}^{(n+1)}\right) - H\left(\widetilde{\mathbf{Z}}^{(n)}, \mathbf{Z}^{(n+1)}\right), \tag{4}$$

where $H(\mathbf{Z}) = -\sum_i p_z(z_i) \log p_z(z_i)$ – image $\mathbf{Z}$ entropy estimation; p_z – marginal PDF estimation of the image sample; $H(\widetilde{\mathbf{Z}}^{(n)}, \mathbf{Z}^{(n+1)}) = -\sum_i \sum_k p_{z_n, z_{n+1}}(z_i, z_k) \log p_{z_n, z_{n+1}}(z_i, z_k)$ – joint entropy estimation; $p_{z_n, z_{n+1}}$ – joint PDF estimation of intensities.

Tsallis MI is defined as [12]:

$$R_g = S_g(\widetilde{\mathbf{Z}}^{(n)}) + S_g(\mathbf{Z}^{(n+1)}) + (1-q)S_g(\widetilde{\mathbf{Z}}^{(n)})S_g(\mathbf{Z}^{(n+1)}) - S_q(\widetilde{\mathbf{Z}}^{(n)}, \mathbf{Z}^{(n+1)}), \tag{5}$$

where $S_g(\widetilde{\mathbf{Z}}^{(n)}, \mathbf{Z}^{(n+1)}) = (g-1)^{-1}\left(1 - \sum_{i=0}^{N}\sum_{j=0}^{N} p_{i,j}^g\right)$ – Tsallis entropy of the order g, g - a real number. When $g = 1$ Tsallis entropy approaches Shannon entropy.

F-information measures are based on divergence or distance between joint probability distributions and multiplication of marginal distribution of image pair. A class of divergence measures that uses MI is the F-information measures. *F*-information measures are [13]:

$$I\text{-measure:} \qquad I_\alpha = \frac{1}{\alpha(\alpha-1)}\left(\sum_{i=0}^{N}\sum_{j=0}^{N} \frac{p_{i,j}^\alpha}{(p_i p_j)^{\alpha-1}} - 1\right),$$

$$M\text{-measure:} \qquad M_\alpha = \sum_{i=0}^{N}\sum_{j=0}^{N} \left|p_{i,j}^\alpha - (p_i p_j)^\alpha\right|^{\frac{1}{\alpha}}, \tag{6}$$

$$\chi\text{-measure:} \qquad \chi_\alpha = \sum_{i=0}^{N}\sum_{j=0}^{N} \frac{|p_{i,j} - p_i p_j|^\alpha}{(p_i p_j)^{\alpha-1}}.$$

I-measure is defined for $\alpha \neq 0$, $\alpha \neq 1$, and when $\alpha = 1$ it approaches Shannon MI. M-measure is defiened for $0 \leq \alpha \leq 1$, and χ-measure – for $\alpha > 1$.

Exclusive F-information is related to entropy of JPD and Shannon MI:

$$D_f(\widetilde{Z}^{(n)}, \mathbf{Z}^{(n+1)}) = 2H\left(\widetilde{Z}^{(n)}, \mathbf{Z}^{(n+1)}\right) - H(\widetilde{Z}^{(n)}) - H(\mathbf{Z}^{(n+1)}). \tag{7}$$

Entropy of JPD is defined as:

$$E = \sum_{i=0}^{N}\sum_{j=0}^{N} p_{i,j}^2, \tag{8}$$

where $p_{i,j}$ – an element of JPD which can be estimated, for example using histograms. The stronger the relationship between two variables the smaller the entropy of JPD.

4. Effectiveness analysis of similarity measures

4.1. Effectiveness criteria

As already noted, a number of criteria for the effectiveness of SM's use in the synthesis of recurrent procedures for estimating spatial deformations of images were proposed in [8]. They are based on an analysis of the slope of SM dependening on the Euclidean mismatch distance [16]. Among the characteristics that determine efficiency, the maximum slope K_{max} in the area of interest of the parameters being evaluated, the effective range of the parameters P, and the region of curvature growth S can be attributed.

It is shown that the maximum slope of SM corresponds to the maximum probability of improving the estimates of the parameters of interframe spatial deformations toward optimal values, and this extremum determines the potential rate of convergence of the estimation vector [7].

The effective range in this work refers to as subdomain of the domain of definition of registration parameters in which the required accuracy values are attained under given constraints, e.g. computational complexity, number of iteration etc. A condition under which a point of the space of estimated parameters falls into the effective range is that the slope of the SM at this point should exceed a certain critical value which does not guarantee the required convergence rate of the estimate vector $\bar{\alpha}$ any longer. Note that the actual effective range also depends on many other factors, in particular, the parameters of the estimation algorithm, the type of spatial deformation, so the critical value of the slope here is considered, in fact, as a criterion for determining the range of values of the discrepancies of the estimated parameters, for which the probabilitie of demolition of estimates exceeds the specified threshold value.

4.2. Experimental results

For SM efficiency analysis it is reasonable to use simulated images whose intensity PDF and correlation function can be priori defined during their synthesis. In conducted experiments simulated images based on wave model [17] with intensity PDF and correlation function close to Gaussian were used. In addition, an unbiased white noise was used as a disturbing factor.

For example, Fig. 2 - Fig. 3 show the graphs of the dependence of the above efficiency indicators on the noise/signal ratio q (in terms of variances), where the left graph corresponds to the data for the correlation ones, and the right graph for the information SMs. As a mismatch, for simplicity, a parallel image shift was selected. We also note that the choice of parallel shift of images as frame-wise deformations doesn't weakens the generalization of the examination, since for any set of strain parameters the result of their effect on each point of the image can be recalculated through the Euclidean mismatch distance [16] into the vector of its shift relative to the initial position. That is, the result of any deformations can be represented by the PDF of such shifts.
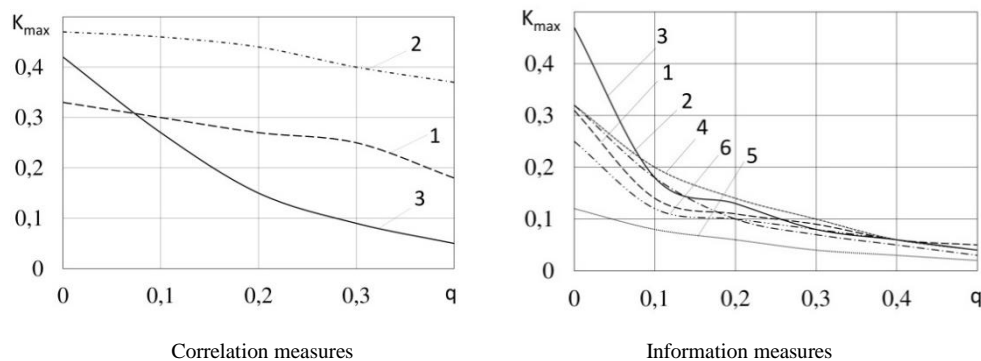

Fig. 2. Maximum slope of SM.

Analysis of fig. 2 shows that by the criterion of the maximum slope among the correlation SM the best result is provided by the Tonimoto coefficient whose maximum slope decreases most slowly with increasing noise intensity. The next by the result is the Kendall coefficient. CC at large noise is much inferior to them, however, with $q < 0.07$ the maximum slope CC exceeds the K_{max} of Kendall coefficient. Investigations of information SMs showed that the potentially high convergence rate of the parameters being evaluated is provided by the I-measure of the F-information. The next most effective M-measure of F-information, which in large noises is more effective than the I-measure. Shannon and Tsallis MI, as well as the entropy of the JPD give close characteristics. Essentially they are inferior to the excluding F-information.

According to the effective range criterion (fig. 3), the best results among the correlation SMs also gave the use of the Kendall rank correlation coefficient. Then, with a margin of up to 40%, the coefficient of Tanimoto and CC, respectively. Among the information measures for small noise ($q < 0.15$) the best results were shown by the I-measure of the F-information and the netropy of the joint JPD. This is followed by Tsallis and Shannon MI and the M-measure of F-information. For large noise ($q > 0.25$), the M-measure and the I-measure of the F-information provide a larger effective range. Approximately the same parameters show the entropy of the JPD, Shannon and Tsallis MI. We note that the slope of the I-measure of the F-

information decreases significantly more rapidly with increasing noise than in the M-measure of the F-information. The worst results were shown by the excluding F-information.
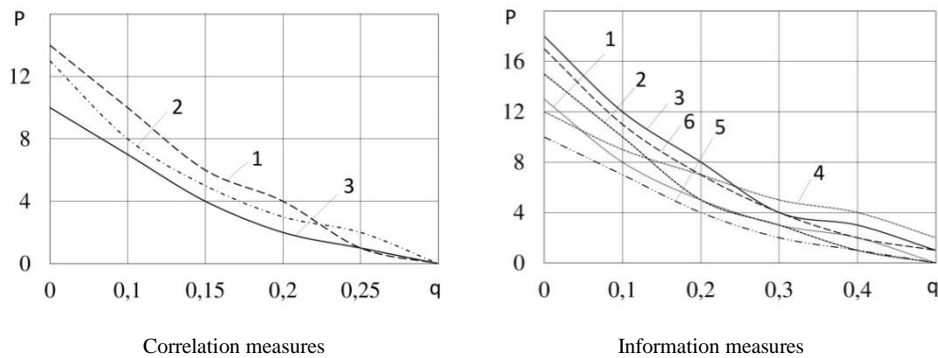

Correlation measures   Information measures

Fig. 3. Effective range of SM.

5. Conclusion

There are many SM of images on the basis of which recursive algorithms for estimating spatial deformations in a sequence of images can be synthesized. However, criteria that allow a priori evaluation of the potential efficiency of algorithms by the SM function are not well studied. A number of correlation (CC, Tanimoto coefficient and Kendall rank correlation coefficient) and informational (Tsallis and Shannon mutual information, F-information measures and joint probability distribution energy) SM are investigated based on the criteria of maximum slope that determines the potential rate of convergence of deformation parameter estimates and effective range, which refers to the subdomain of the domain of deformation parameters, in which the required indicators of accuracy and reliability of estimation are provided.

Studies have shown that, according to the criterion of maximum slope from correlation SM, the best result is provided by the Tonimoto coefficient, the maximum slope of which decreases most slowly with increasing noise intensity, and among the information I-measure of F-information. The worst results were shown by the CC and the excluding F-information, respectively.

By the criterion of the effective range, the best results among the correlation SMs were provided by the Kendall rank correlation coefficient. Then, respectively, the Tanimoto and CC coefficients. Among the information measures, the I-measure and the M-measure of the F-information showed greater resistance to noise. In this case, the slope of the I-measure of the F-information decreases significantly faster with increasing noise than in the M-measure.

Acknowledgement

References

[1] Gonzalez RC, Woods RE. Digital image processing. New Jersey: Prentice Hall, 2002; 793 p.

[2] Goshtasby AA. Image registration. Principles, tools and methods: Advances in Computer Vision and Pattern Recognition. Springer, 2012; 441 p. DOI: 10.1007/978-1-4471-2458-0.

[3] Theodoridis S, Koutroumbas K. Pattern Recognition, 4th edn. New York: Academic Press, 2009; 984 p.

[4] Vasiliev, K.K. Statistical analysis of multidimensional images / K.K. Vasiliev, V.R. Krasheninnikov. - Ulyanovsk: UlGTU, 2007. – 170 p.

[5] Krasheninnikov VR. The fundamentals of image processing theory. Ulyanovsk: UlGTU, 2003; 152 p.

[6] Tashlinskiy AG. Estimation of image matching parameters for image sequences. Ulyanovsk: UlGTU, 2000; 131 p.

[7] Brown LG. A survey of image registration techniques. ACM Computing surveys 1992; 24: 325–376. DOI: 10.1145/146370.146374.

[8] D'Agostino E, Maes F, Vandermeulen D, Suetens P. An information theoretic approach for non-rigid image registration using voxel class probabilities. Med Image Anal. 2006; 6(3): 413–431. DOI: 10.1007/978-3-540-39701-4_13.

[9] De Castro E, Morandi C. Registration of translated and rotated images using finite Fourier transform. IEEE Transactions on Pattern Analysis and Machine Intelligence 1987; 9(5): 700–703. DOI: 10.1109/TPAMI.1987.4767966/.

[10] Kendall MG. A new measure of rank correlation. Biometrika 1938; 30: 81–93. DOI: 10.2307/2332226.

[11] Sevim YA, Atasoy A . Performance comparison of new nonparametric independent component analysis algorithm for different entropic indexes. Turkish Journal of Electrical Engineering & Computer Sciences 2012; 20: 287–297. DOI:10.3906/elk-1004-1.

[12] Tashlinskii AG. Computational expenditure reduction in pseudo-gradient image parameter estimation. Lecture Notes in Computer Science 2003; 2658: 456–462. DOI: 10.1007/3-540-44862-4_48.

[13] Tashlinskii AG. Pseudogradient Estimation of Digital Images Interframe Geometrical Deformations. Vision Systems: Segmentation & Pattern Recognition. Vienna, Austria: I-Tech Education and Publishing, 2007: 465–494. DOI: 10.5772/4975.

[14] Taslinskii AG. Optimization of goal function pseudogradient in the problem of interframe geometrical deformations estimation. Pattern Recognition Techniques, Technology and Applications. Vienna, Austria: I-Tech. 2008: 249–280. DOI: 10.5772/4975.

[15] Voronov SV, Taslinskii AG. Efficiency analysis of information theoretic measures in image registration. Pattern recognition and image analysis 2016; 26(3): 502–505. DOI: 10.1134/S1054661816030226.

[16] Voronov SV. The use of mutual information as objective function for image parameters' estimation. Radiotehnika 2014: 7: 88–94.

[17] Tashlinskiy AG, Tikhonov VO. Method for error estimation for stochastic gradient estimation of multidimensional processes. Izvestiya vuzov, seriya "Radioelektronika" 2001; 44(9): 75–80.