Internal and External Salient Points under Affine Transformations. Comparative Study

S. Mashtalir
Department of Computer Science
Kharkov National University of Radio Electronics
Lenin ave. 14
61166, Kharkov, Ukraine
mashtalir_s@kture.kharkov.ua

K. Shcherbinin
Department of Computer Science
Kharkov National University of Radio Electronics
Lenin ave. 14
61166, Kharkov, Ukraine
konstantin@inetvisor.com

E. Yegorova
Department of Computer Science
Kharkov National University of Radio Electronics
Lenin ave. 14
61166, Kharkov, Ukraine
yegorova@kture.kharkov.ua

ABSTRACT

Permanently a special emphasis is given to image processing considering geometric distortions for remote sensing, and image tracking especially. We discuss an approach to affine transformation parameters search on the base of salient points which can be external and internal relative to shapes. To find affinity we use sets of image binary cuts and one-to-one point correspondence. This correspondence is determined by means of the affine-equivalent shape ratio invariance. We investigate two types of salient points on the base of binary morphology. First ones are vertices of convex hulls and second ones belong to skeletons. Finally, solution of overdetermined system of linear equations gives us affine transformation parameters. Results of the exterior points and skeleton normalization methods comparative analysis are discussed.

Keywords

Salient points, mathematical morphology, image normalization.

1. INTRODUCTION

Invariance to natural and artificial deformations or their elimination represent one of the traditional pattern recognition problems, which still does not have a final solution. With reference to image processing it is typical to deal with geometric distortions induced by change of a mutual location of object and videosensor. In most cases distortions of this kind are simulated by affine transformations. Despite of numerous researches on normalization [see, e.g. Rot96, She99, Pei95] there are some classes of images for which it is expedient to synthesize new algorithms based on the shape feature analysis.

A shape analysis is most convenient applied to binary images. Nevertheless the valid acquisition of binary image is sufficiently challenging problem for arbitrary gray-level videodata. Reasonable approach is to process a whole set of image binary cuts simultaneously. The reduction of gray-level image processing to binary cuts sets analysis provides high accuracy and reliability of normalization and is rather simple for unified hardware-software implementation including SIMD architectures first of all. Thus the solution can be based on the image affine-equivalent salient points, which steadily characterize shape properties.

Steadfast attention is devoted to the points of interest (salient points) search [see, e.g. Cha97, Dau01, Mik04, Zhu95, Gol98, Spr94, Sud97]. Furthermore, traditional use of interactively indicated reference points should be mentioned [Dau01, Zhu95]. However in some cases human interference in image processing system operations is difficult or generally impossible. At the same time, known automatic procedures of affine-equivalent points search do not have a sufficient noise stability with reference to normalization problems.

Thus, our objective is to present and experimentally investigate reliable procedures of searching the image salient points and detecting one-to-one affine corresponding point (feature points) among them. We describe techniques of salient points (and

Conference proceedings ISBN 80-86943-03-8
WSCG'2006, January 30-February 3, 2006
Plzen, Czech Republic.

indexed subsets) search and experimentally analyze their effectiveness.

Indices of processing speed, accuracy and noise stability are used as effectiveness characteristics of normalization. Offered normalization methods are based on binary morphology operations [Ser82, Hei94, Gia88, Mar02] and lie in the automatic search of binary images (or cuts) affine-equivalent points. Two techniques of feature points search are investigated. They depend on external points and on image skeletons. The statistical analysis of image indexed salient points subsets search methods has been carried out to define specification of their practical use.

2. METHODS OF IMAGE AFFINE-EQUIVALENT POINTS SETS SEARCH

2.1 Normalization problem statement

Suppose that in some time poins t', t'' two images $B'(x), B''(x)$ ($x \in R^2$) of the same object are fixed in videosensor field $D \subset R^2$ and they are connected by affine planar transformation $B'(x) = g \circ B''(x)$, i.e. $x \in D, g \in G = Aff(2, R)$. The group G action with identical transformation denoted by e has form

$$B'(x) = B''(Ax + b), \quad (1)$$

where $A \in Gl(2, R)$ is a nonsingular matrix of the second order, $b \in R^2$ is a translation vector. Note, that each image can be represented as a union of m binary cuts (generally in mixed scale of notation)

$$B = \bigcup_{i=1}^{m} B_i w_i \ .$$

Here $B_i = 1 \ \forall \gamma \in D: B(\gamma) \in [\alpha_i, \beta_i]$, α_i, β_i are given or found thresholds of gray levels partitions and $B_i = 0$ otherwise, w_i are weight numbers connected to the image representation i.e. the radix number steps of gray levels can be chosen as their discriminant homogeneous regions (see Figure 1). It is obvious, that $B'_i = g \circ B''_i$.

The problem consists in searching image geometric transformation parameters $A \in Gl(2, R)$, $b \in R^2$ on given sets $\{B'_i\}_{i=1}^{m}, \{B''_i\}_{i=1}^{m}$.

Let us underscore that transfer to binary cuts analysis reduces a normalization problem only to proceeding shape characteristics of object image and its components, in particular its individual points. Further, we shall omit index i and consider one cut as a binary image B, identifying it, if necessary, with a gray-level image.

Figure 1. An image and binary cuts

Let us consider methods based on the salient points analysis. It is related to the fact that stable automatic selection of affine-equivalent points provides simple algorithmization of geometric transformation parameters search (i.e. normalization problem solution) as they are found straightly from a linear equations system

$$\begin{cases} x'_{1,k} = a_{11} x''_{1,k} + a_{12} x''_{2,k} + b_1, \\ x'_{2,k} = a_{21} x''_{1,k} + a_{22} x''_{2,k} + b_2, \end{cases} \quad k = \overline{1, K}, \ K \geq 3. \quad (2)$$

Especially we shall mark, that accurate and reliable affine-equivalent points search alongside with its computational simplicity, like other local methods provides a number of additional possibilities e.g. partial objects overlapping consideration, removing the object from videosensor field, i.e.

$$\exists x \in D : B(x) \neq 0, \quad Ax + b \notin D .$$

2.2 Image salient points sets

Let us introduce the basic definitions necessary for binary image affine-equivalent points search and validate effectiveness of mathematical morphology methods application as theoretical and practical toolkit. First we shall suppose that a map belonging to the power map $\mathcal{F} : B \to 2^B$ is given and it extracts the affine-equivalent points families. Then (after the necessary properties are established) we shall find such maps.

Definition 1. For arbitrary binary image B' elements of set $\mathcal{F}(B') = \{x\}$ we shall name salient points, if

$$\forall g \in G \ \forall x' \in \mathcal{F}(B') \ \ \exists!\, x'' \in \mathcal{F}(\mathfrak{B}'') : x' = g \circ x''. \ (3)$$

It is clear, that equation in (3) is also valid inversely

$$\forall g \in G \ \forall x'' \in \mathcal{F}(B'') \ \exists!\, x' \in \mathcal{F}(B') \ : \ x'' = g^{-1} \circ x'.$$

As an example of set $\mathcal{F}(B)$ it is possible to specify boundary ∂B of image B. At first sight uniqueness of the affine-equivalent points existence in (3) is rather obvious by virtue of the group transformations properties. However it is necessary to take into account, that images are actually considered not in $D \subset R^2$, but in $D \subset Z^2$, what may lead to violation of the group transformations properties, i.e. from the equality $\forall x \in Z^2,\ g \circ x = x$ does not necessarily follow $g = e$, and also there may exist a subregion of the videosensor field $D^* \subseteq D$ such, that

$$\exists g \in G : B(x) \neq 0, \forall x \in D^* \Rightarrow Ax + b = const$$

Sufficient illustration can be a scaling, when one boundary ∂B point can correspond to several points under same transformation parameters.

Definition 2. The indexed salient points $\Theta \subseteq \mathcal{F}(B)$, $\Theta = \{\theta_i\}_{i=1}^{I}$, $I \leq card\ \{\mathcal{F}(B)\}$ we shall name feature points, if $\forall B \in M$, $\forall g \in G$ there is a permutation $\pi_{\mathcal{F}} = \begin{pmatrix} 1 & 2 & \cdots & I \\ \alpha_1 & \alpha_2 & \ldots & \alpha_I \end{pmatrix}$ such, that

$$\theta_i = g \circ \theta_{\alpha_i},\ \theta_i \in \mathcal{F}(B),\ \theta_{\alpha_i} \in \mathcal{F}(g \circ B).$$

Directly from the definition it follows that condition $card\ \{\mathcal{F}(B)\} = card\ \{(g \circ B)\}$ or more precisely $card\ \Theta = card\ g \circ \Theta$ should be satisfied. In other words, it is necessary that map $\mathcal{F}$ would not depend on planar affine transformations.

When we select n feature points and form matrices

$$\Lambda' = \begin{pmatrix} \theta'_{x1} & \theta'_{x2} & \cdots & \theta'_{xn} \\ \theta'_{y1} & \theta'_{y2} & \cdots & \theta'_{yn} \end{pmatrix}$$

$$\Lambda'' = \begin{pmatrix} \theta''_{x1} & \theta''_{x2} & \ldots & \theta''_{xn} \\ \theta''_{y1} & \theta''_{y2} & \ldots & \theta''_{yn} \\ 1 & 1 & \ldots & 1 \end{pmatrix}$$

required affine parameters are

$$(A\ b) = \Lambda' \Lambda''^{T} (\Lambda'' \Lambda''^{T})^{-1}.$$

Let $B_\phi \subseteq B, \phi \in \Phi$, we shall impose on producing salient points map $\mathcal{F} : B \to 2^B$ conditions

$$\mathcal{F}(\varnothing) = \varnothing,\ \forall B_\phi \subseteq B : B_\phi \neq \varnothing \Rightarrow \mathcal{F}(B) \neq \varnothing, \quad (4)$$

$$\mathcal{F}(\textstyle\bigcup_{\phi \in \Phi} B_\phi) = \bigcup_{\phi \in \Phi} \mathcal{F}(B_\phi), \quad (5)$$

$$\forall B', B'' \subset B : B' \subset B'' \Rightarrow \mathcal{F}(B') \subset \mathcal{F}(B''), \quad (6)$$

$$\mathcal{F}(\textstyle\bigcap_{\phi \in \Phi} B_\phi) \subset \bigcap_{\phi \in \Phi} \mathcal{F}(B_\phi). \quad (7)$$

Condition (4) specifies existence of the required map. The condition of additivity (5) together with a condition of inheriting (6) guarantees a possibility of fragment processing, i.e. creates premises for realization on SIMD-architecture. Condition (7) ensures processing of several images binary cuts, including outcomes of their interim processing. If $\mathcal{F}$ is a biunivocal mapping, inclusion in (7) grades into equality.

Let us search maps $\mathcal{F}$ by constructing set-theoretic operations superpositions and Minkowski algebra operations. It follows from the clear fact that the class of bit cuts is closed concerning Minkowski algebra operations and Minkowski addition and subtraction

$$X \oplus Y = \textstyle\bigcup_{y \in Y} X + y,\ X \ominus Y = \bigcap_{y \in Y} X + y,$$

(X, Y are arbitrary sets) satisfy conditions (4)–(7).

We shall emphasize that if fixing one of the sets, namely, its spatial form and structure in operations (8), (9), we can receive subsets with the given properties in a range of definition. Constructive and traditional way of Minkowski algebra operations realization with reference to image processing binary morphology operations [Ser82, Hei94].

Let still B be a binary image defined in videosensor field $D \subset Z^2$. Basic operations are: dilation, coinciding with Minkowski addition

$$\delta_H(B) = \textstyle\bigcup_{h \in H} (B + h) = \{x \in B : [(H + x) \bigcap B] \subset B\} \quad (8)$$

and erosion

$$\varepsilon_H(B) = \textstyle\bigcap_{h \in H} (B + h) = \{x \in B : (H^r + x) \subset B\} \quad (9)$$

where $H^r = \{-x \in B : x \in H\}$, $H \subseteq D$ is a fixed set named as a structural element. We used structural elements H^* which are shown in Figure 2.

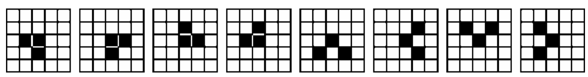

Figure 2. Set of structural elements

Sets of image affine-equivalent points can be of two types: interior and exterior. Interior points are related to images convex hulls, and exterior – to the centers of gravity, centroids, etc. For the considered two algorithms of interior and exterior feature points we used notations $N_8, N_4 \in H$:

$$N_4(\alpha) = \{\beta \in \mathbb{Z}^2 : |\alpha_x - \beta_x| + |\alpha_y - \beta_y| = 1\},$$

$$N_8(\alpha) = \{\beta \in \mathbb{Z}^2 : max\ \{|\alpha_x - \beta_x|, |\alpha_y - \beta_y|\} = 1\},$$

Then we can find boundaries

$$\partial_4(B) = B / \varepsilon_{N_8}(B),\ \partial_8(B) = B / \varepsilon_{N_4}(B)$$

and using combination of morphological operations (8) and (9) affine-equivalent skeleton was constructed for the interior feature points set search

$$\Phi_i(B,H) = (\varepsilon_{H^*}(B_i / \varepsilon_{N_4}(B_i)) \cup (B_{i-1} / \delta_{N_8}(B_i)), \quad (10)$$

where $B_0 = B$, $\Phi_0(B, H) = \varepsilon_{H^*}(\partial_8(B_0))$, $B_i = B_{i-1} / / \partial_4(B_{i-1})$, $i = \overline{1, K}$, K is such that $B_K = \varnothing$, then the aggregate of points obtained with the help of (10) forms the map $\mathcal{F} = \bigcup_{i=0}^{K} \Phi_i(B, H)$ which produces the salient points forming affine-equivalent skeleton $Sk(B)$. Obtained in this way skeleton became a basis of the interior points binary image normalization method.

The second offered algorithm is based on the external points search. It can be proved that the following morphological operations combination allows to find external salient points set

$$\mathcal{F} = \varepsilon_{H^*}(\partial_8(B)) \quad (11).$$

As algorithm, described by (11), searches for the salient points, feature points set needed to be obtained by building an image convex hull, which peaks formed the required set. For affine-equivalent skeletons (10) node points (point which have more than two neighbors according to 8-connectivity paradigm) will be the feature points. To find correspondence between feature points sets of template and distorted images it is necessary to use property of areas ratio invariance under affine transformations. For the algorithm based on external point search we find permutation

$$\pi_{\mathcal{F}} = \begin{pmatrix} p & mod_n(p+1) & \dots & mod_n(p+n-1) \\ \alpha^* & mod_n(\alpha^*+1) & \dots & mod_n(\alpha^*+n-1) \end{pmatrix},$$

where p is a fixed feature point on the convex hull of image B'_i and α^* is the corresponding point of the convex hull of B''_i, n is the number of feature points. For internal feature points we proceed by analogy.

The point set obtained by means of such map has laid the foundation for the feature points binary images normalization methods.

However the question arises as to these methods accuracy, reliability and stability. With the purpose of these characteristics clarification, and also to determine their practical applications, we shall carry out the comparative analysis of these morphological normalization methods.

3. COMPARATIVE ANALYSIS OF MORPHOLOGICAL NORMALIZATION METHODS

Let us introduce the affine transformation group operation $Aff(2, R)$ (1) in homogeneous coordinates

$$B'(x) = g \circ B''(x) \sim B'(x) = B''(Hx),$$

where $x = (x_1, x_2, 1)^T$, $H = \begin{pmatrix} A & & b \\ 0 & 0 & 1 \end{pmatrix}$. For realization of authentic statistical researches the knowledge of affine transformation true parameters and the values found by means of feature points is necessary. In other words, we shall obtain affine-equivalent images from templates $B_0^k(x)$, $k = 1, 2, \dots$ using affine transformation, defined by 3×3 matrix H_t, i.e. $B'(x) = B_0(H_t x)$. We shall denote obtained values as H_v, then the normalization error in a feature space will be

$$\varepsilon_M = \sqrt{tr\,[(H_v - H_t)(H_v - H_t)^T]}, \quad (12)$$

where $tr\,[\circ]$ is a spur of a matrix. We shall mark that given estimation can be used only if the normalization accuracy is high enough, as affine transformations are non-isometric.

Error in images space we shall define as measure of intersection of template $B_0(x)$ and normalized $B_v(x) = B'(H_v^{-1}x)$ images in the videosensor field

$$\varepsilon_I = 1 - \frac{card\,(B_0 \cup B_v) - card\,(B_0 \cap B_v)}{card\,(B_0 \cup B_v)}. \quad (13)$$

In recognition problems the estimation (12) is more preferable; however at the experimental researches (when fixing some registration conditions and image processing) expression (13) gives a clearer understanding.

Under research of feature points normalization methods characteristics with the purpose of the binary cuts deriving algorithms influence elimination we shall use binary images only. 88 images used in a simulation modeling are shown in Figure 3.

Examples of normalization results using salient points and skeletons are shown in Figure 4 and Figure 5 accordingly.

On first column of Figure 4 one can see template images, on second one – distorted images. Feature points sets and convex hulls for template and transformed images are shown in columns 3 and 4 respectively. Last column contains normalization results obtained by method using (11).

In a similar manner in first two columns of Figure 5 template and distorted images are shown. Columns 3 and 4 contain skeletons for these images respectively and in the last column normalization results obtained by method using (10) are presented.

To determine whether the type of entering images has any influence on the normalization accuracy, i.e.

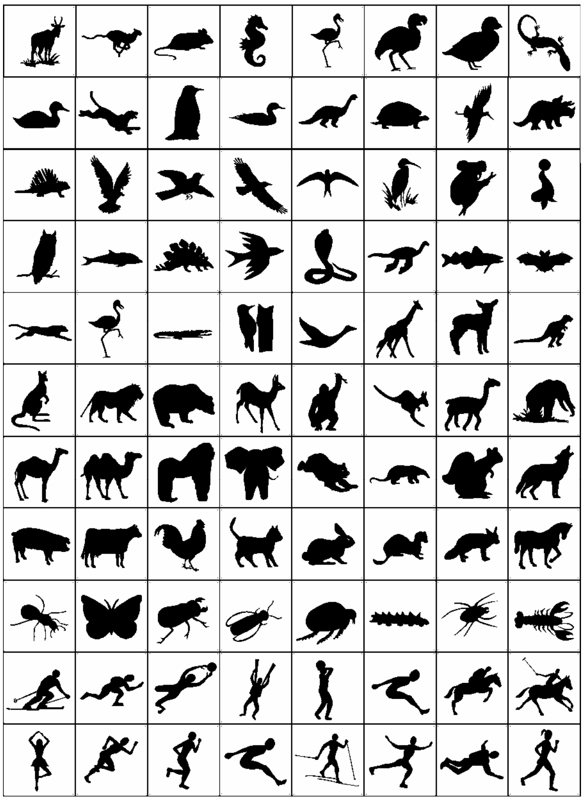

Figure 3. Binary images used in experiments

whether normalization averages are comparable, the variance analysis tools were used.

Let us note that independence of each sample from the others is provided by independency of processes image ensemble choice.

All images were introduced on the discrete raster having 512×512 pixels. To provide the equivalence of all images presentation in digital type (eliminations of quantization and digitization errors) they were obtained by converting various vectorized fonts versions of TTF format into BMP format files.

For variance analysis application it is necessary, that normalization errors would be distributed under the normal law and their variances are uniformal. To check the hypothesis of distribution normality, the computations results were broken into 10 intervals.

For each of them number of results m_i and probability of p_i hitting this interval under the normal probability distribution law was determined.

Estimation of criterion value χ^2 at $l = 10$

$$\chi^2 = \sum_{i=1}^{l} \frac{(m_i - np_i)^2}{np_i} = 1.05,$$

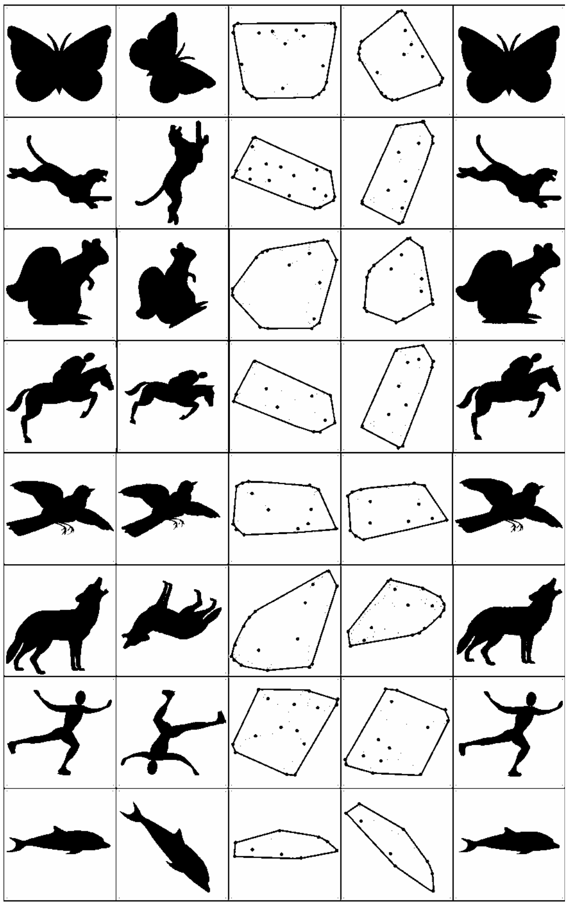

Figure 4. Normalization using exterior points

and critical criterion value under $k = 7$ degrees of freedom and $p = 0.99$ confidence probability is equal $\chi^2 = 1.24$, i.e. there is no sufficient foundation to reject a hypothesis about normal distribution.

Checking line uniformity hypothesis has provided the following results: choosing a confidence probability $p = 0.99$, we discover the upper critical limit χ^2, under number of degrees of freedom equal 87, $\chi^2 = 59.279$, i.e. we have accordingly

$$Q_1 = 21.74 < \chi^2 = 59.28, \; Q_2 = 4.02 < \chi^2 = 59.28,$$

what means that the hypothesis about variances line uniformity does not contradict the experiment results.

Further we shall calculate an unbiased estimator of deviations inside each series and get for the indicated data at degree of freedoms $k - 1 = 87$ and $N - k = 528$ accordingly

$$F_1 = s_{a,1}^2 / s_{r,1}^2 = 1.63 > F_{0.01} = 1.43$$

and

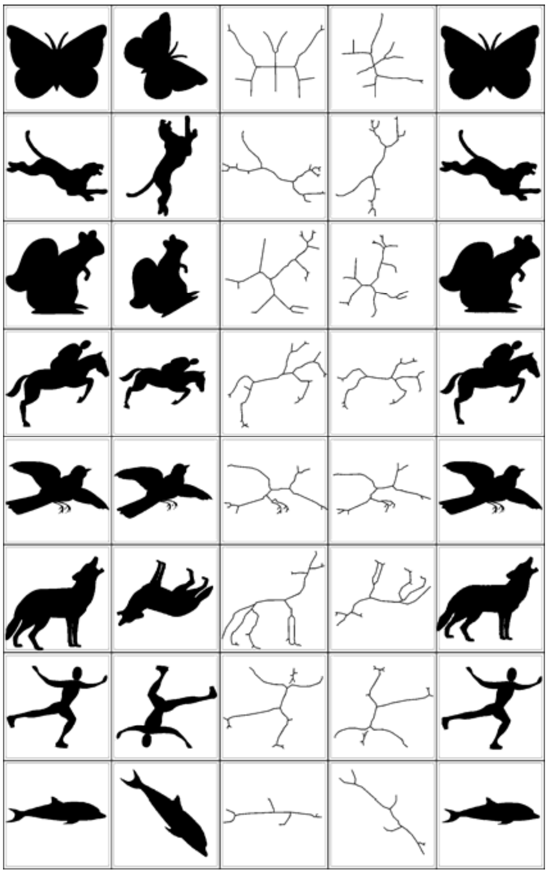

Figure 5. Normalization using interior points

$$F_2 = s_{a,2}^2 / s_{r,2}^2 = 0.161 < F_{0.01} = 1.43 .$$

Thus, it is possible to state that based on the carried out experiments there are no foundation to consider that skeleton based normalization accuracy depends on the binary images type. At the same time, there is no reason to believe that extreme points morphological normalization does not depend on input images. The obtained result has a rather transparent explanation: external points extraction is not sufficiently steady at the small areas image fragments analysis, what brings the erroneous data into the simple equations systems. Elimination of such errors can be carried out by similar feature points analysis exclusion, and also by increasing the number of images binary cuts. On the other hand, skeletons construction is based on a local symmetry of binary cuts, i.e. actually in an implicit aspect some integrated estimation of images is used, what ensures higher stability of the method. In other words, it is necessary to choose either faster or more reliable method with reference to applied specification of solving tasks, or use their combinations.

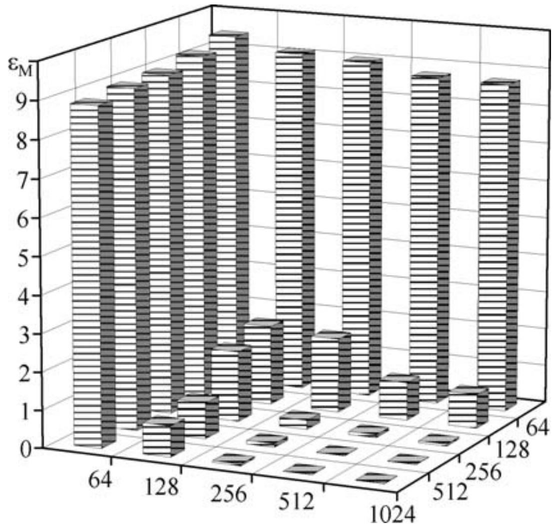

Figure 6. Averaged normalization error under different discrete raster

Let us underline, that discrete raster observation number also influences the end results. In Figure 6 averaged on 10 images and 5 affinities (50 experiments) skeleton normalization errors dependency on traditional for digital processing of videosensor field characteristics is shown. It is visible, that accuracy raises dramatically remains rather stable under observation/indication number exceeding 256.

Also experimental check of morphological normalization methods stability under conditions of noise presence was carried out. At first, the normalization error dependency from noise level and density in spatial area is considered. Gaussian and uniform noises were used as distortions. Following results were obtained: despite the density of noise, at the level up to 40 %, the normalization error in space does not exceed 2 % from the shape area, what is an acceptable result. Further, depending on a density it begins to increase considerably. The normalization error dependency from noise type was also considered. Results showed that normalization error in space changes more for uniform noise. Dependency from the noise level is much stronger than from their density.

Further the experimental analysis concerning geometric transformation parameters improvement on binary cuts sets was carried out. It is possible to draw a conclusion that under increasing of processed informative binary cuts number for the given class of images, the error of parameters determination decreases. However, if at the initial stage addition of new informative cut considerably reduces the average error for a class of images, starting with 6 cuts the error remains practically constant. From this it is possible to draw a conclusion, that the optimal amount of processed informative cuts for considered classes of images is a number in a range from 3 to 6. It is necessary to note, that one should not misapply the amount of cuts since multiple increasing of their number will lead to the case when informativity of each single cut will decrease, what can lead to a significant error of image normalization parameters determination.

4. CONCLUSIONS

The basic scientific significance of this work is the new methods which describes image feature points sets on a discrete raster. Steady automatic search of those is one of necessary parts of image normalization problems solution, as used for geometrical transformation parameters search. We analyzed stability and reliability problems of searching the indexed salient points sets which are the solution of external points and skeleton morphology normalization tasks based on mathematical morphology operations combinations. For statistical analysis of the offered methods errors of parameters definition were introduced and experiments for detecting normalization accuracy dependency from input images type, noise, images physical sizes and quantity of analyzed cuts were carried out. As a practical significance of results it is necessary to specify, that the offered methods have sufficient effectiveness and can be used in technical vision systems, however their application is limited to specific classes of images.

5. REFERENCES

[Rot96] Rothe, I., Suesse, H., Voss, K. The method of normalization to determine invariants. IEEE Trans. on Pattern Analisys and Machine Intelegence 18, No4. pp. 366–377. 1996.

[She99] Shen D., Wong W.-H., Ip H.H.S. Affine-invariant image retrieval by correspondence matching of shapes. Image and Vision Computing 17, pp. 489-499. 1999.

[Pei95] Pei, S.C., Lin, C.N. Image normalization for pattern recognition. Image and Vision Computing 13, No8. pp. 711–723. 1995.

[Cha97] Chang S.H., Cheng F.H., Hsu W.H., Wu G.Z. Fast Algorithm for Point Pattern Matching: Invariant to Translations, Rotations and Scale Changes. Pattern Recognition 30, No2, pp. 311-320. 1997.

[Dau01] Daurat, A. Salient points of Q–convex sets. International Journal of Pattern Recognition and Artificial Intelligence 15, No7. pp. 1023–1030. 2001.

[Mik04] Mikolajczyk K., Schmid C. Scale & affine invariant interest point detector. International Journal of Computer Vision 60, No1. pp. 63–86. 2004.

[Zhu95] Zhuang, X., Zhao, D. A morphological algorithm for detecting dominant points on digital curves. SPIE Proc. «Nonlinear Image Processing VI» 24. pp.372–383. 1995.

[Gol98] Gold, S., Rangarajan, A., Lu, C.P., Pappu, S., Mjolsness, E. New algorithms for 2D and 3D point matching: Pose estimation and correspondence. Pattern Recognition 31, No8. pp. 1019–1031. 1998.

[Spr94] Sprinzak, J., Werman, M. Affine point matching. Pattern Recognition Letters 15, No4. pp. 337–339. 1994.

[Sud97] Sudhir G., Banerjee S., Zisserman A. Finding point correspondence in motion sequences preserving affine structure. Computer Vision and Image Understanding 68, No2, pp. 237-246, 1997.

[Ser82] Serra, J. Image analysis and mathematical morphology. London: Academic Press, 1982.

[Hei94] Heijmans, H.J.A.M. Morphological image operators. Boston: Academic Press, 1994.

[Gia88] Giardina, C.R., Daugherty, E.R. Morphological methods in image and signal processing. New York: Prentice–Hall, Englewood Clifs, 1988.

[Mar02] Maragos P., Schafer, R.W., Butt, M.A., (eds.) Mathematical morphology and its applications to image and signal processing. Computational Imaging and Vision 5. Dordrecht-Boston-London: Kluwer Academic Publishers, 2002.