SHAPE-BASED IMAGE MATCHING USING HEAT KERNELS AND DIFFUSION MAPS

Yu. V. Vizilter *, V.S. Gorbatsevich, A. Yu. Rubis, S.Yu. Zheltov

State Research Institute of Aviation Systems (GosNIIAS),
125319, 7, Viktorenko str., Moscow, Russia – (viz,gvs,arcelt,zhl)@gosniias.ru

Commission III, WG III/4

KEY WORDS: Image Matching, Diffusion Maps, Mathematical Morphology.

ABSTRACT:

2D image matching problem is often stated as an image-to-shape or shape-to-shape matching problem. Such shape-based matching techniques should provide the matching of scene image fragments registered in various lighting, weather and season conditions or in different spectral bands. Most popular shape-to-shape matching technique is based on mutual information approach. Another well-known approach is a morphological image-to-shape matching proposed by Pytiev. In this paper we propose the new image-to-shape matching technique based on heat kernels and diffusion maps. The corresponding Diffusion Morphology is proposed as a new generalization of Pytiev morphological scheme. The fast implementation of morphological diffusion filtering is described. Experimental comparison of new and aforementioned shape-based matching techniques is reported applying to the TV and IR image matching problem.

1. INTRODUCTION

This paper addresses the problem of shape-based image matching with no dependence on the concrete intensity or radiometric pixel values. For example, one can compare images of one scene captured in different seasons, times of day, weather and lighting conditions, spectral ranges and so on. In such cases the traditional image matching techniques like least square matching (LSM) or zero mean normalized cross correlation (ZNCC) are usually unstable due to the variability of intensity.

The most popular technique for image shape comparison utilizes the mutual information (MI) measure based on probabilistic reasoning and information theory (Maes et al, 1997). The other approach was proposed by Pytiev (so called "Pytiev morphology") based on geometrical and algebraic reasoning (Pyt'ev, 1993). This approach expresses the geometrical idea of image shape in the evident form and provides the image-to-shape matching technique based on morphological correlation coefficient (MCC). In our paper (Vizilter, Zheltov, 2012) this Pytiev morphological approach was applied for obtaining the shape-to-shape matching technique based on mean square effective morphological correlation coefficient (MSEMCC).

Unfortunately, the experiments with TV and IR image matching demonstrate that all these shape-based matching techniques are not robust enough relative to noise and high-frequency damage of images. We suppose that such insufficient robustness is a result of poor quality of image segmentation, because the traditional morphological 2D shape description as a set of flat zones of frame tessellation is too sensitive relative to noise, segmentation scheme and its parameters.

From the other hand, in the area of manifold learning and data shape matching some stable shape description and matching techniques are known, based on heat kernels and diffusion maps (Belkin, Niyogi, 2001), (Lafon, 2004), (Coifman et al, 2007), (Memoli, 2011). These techniques do not require preliminary data segmentation. So, there is a chance that modified morphological tools based on such "diffusion" ideas will be more robust.

In this paper we propose a new generalization of Pytiev morphological scheme based on heat kernels and diffusion maps. In the framework of this Diffusion Morphology the new image-to-shape matching technique is implemented and tested relative to the aforementioned matching techniques applying to the TV and IR images of the same scene.

2. RELATED WORKS

This section contains the brief overview of both shape-based 2D image comparison techniques and data description and matching techniques based on heat kernels and diffusion maps.

2.1 Mutual information

The approach for 2D shape matching based on mutual information (MI) notion from information theory was proposed in (Maes et al, 1997). *Mutual information $I(A,B)$* estimates the dependence of two random variables A and B by measuring the distance between the joint distribution $p_{AB}(a,b)$ and the distribution of complete independence $p_A(a)p_B(b)$:

$$I(A,B) = H(A) + H(B) - H(A,B), \qquad (1)$$
$$H(A) = -\sum_a p_A(a)\log p_A(a),\ H(B) = -\sum_b p_B(b)\log p_B(b),$$
$$H(A,B) = \sum_a \sum_b p_{AB}(a,b)\log p_{AB}(a,b),$$

where $H(A)$ is an entropy of A, $H(B)$ is an entropy of B, and $H(A,B)$ is their joint entropy. For two intensity values a and b of corresponding pixels in the two images, the required empirical estimations of joint and marginal distributions can be obtained by normalization of the joint (2D) and marginal (1D) histograms of compared image fragments. Maximal $I(A,B)$ value over a set of relative fragment positions corresponds to the best geometrical matching of image fragments. This approach does not describe the image shape in the evident form, but different successful shape matching applications were created based on this MI approach in recent years, for example (Goebel, 2005), (Ji, 2005).

---

* Corresponding author.

2.2 Morphological image analysis and geometrical correlation

In the framework of Morphological Image Analysis (MIA) proposed by Pytiev (Pyt'ev, 1993) images are considered as piecewise-constant 2D functions

$$f(x,y) = \sum_{i=1}^{n} f_i \chi_{F_i}(x,y),$$

where n – number of non-intersecting connected regions of *tessellation* $\mathbf{F}$ of the frame Ω, $\mathbf{F}=\{F_1,\dots,F_n\}$; $\mathbf{f}=(f_1,\dots,f_n)$ – corresponding vector of real-valued region intensities; $\chi_{Fi}(x,y)\in\{0,1\}$ – characteristic (support) function of i-th region:

$$\chi_{F_i}(x,y) = \begin{cases} 1, & if\ (x,y)\in F_i; \\ 0, & otherwise. \end{cases} \quad (2)$$

This tessellation of image is supposed to be obtained by some image segmentation procedure. Set of images with the same tessellation $\mathbf{F}$ is a convex and close subspace $F\subseteq L^2(\Omega)$ called *shape-tessellation*, *mosaic shape* or simply *shape*:

$$\mathbf{F} = \left\{ f(x,y) = \sum_{i=1}^{n} f_i \chi_{F_i}(x,y),\ \mathbf{f}=\{f_1,\dots,f_n\}, \mathbf{f}\in R^n \right\} \quad (3)$$

For any image $g(x,y)\in L^2(\Omega)$ the *projection onto the shape F* is determined as

$$g_F(x,y) = P_F g(x,y) = \sum_{i=1}^{n} g_{F_i} \chi_{F_i}(x,y), \quad (4)$$

$$g_{F_i} = (\chi_{F_i},g)/\|\chi_{F_i}\|^2,\quad i=1,\dots,n.$$

Pytiev morphological comparison of images $f(x,y)$ and $g(x,y)$ is performed using the *normalized morphological correlation coefficients* (MCC) of the following form

$$K_M(g,F) = \frac{\|P_F g\|}{\|g\|},\quad K_M(f,G) = \frac{\|P_G f\|}{\|f\|}.$$

The first formula estimates the closeness of image g to the "shape" of image f. Second formula measures the closeness of image f to the "shape" of image g. For elimination of constant non-informative part of image brightness following image normalization is usually performed:

$$K_M(g,F) = \frac{\|P_F g - P_O g\|}{\|g - P_O g\|},\ K_M(f,G) = \frac{\|P_G f - P_O f\|}{\|f - P_O f\|}, \quad (5)$$

where $P_O f$ – projection of image f onto the "empty" shape O with one flat zone. This projection is a constant-valued image filled by mean value of projected image.

Later some advances and modifications of this approach were proposed in (Pyt'ev et al, 2006), (Falomkin, Pyt'ev, 2007), (Vizilter, Zheltov, 2009), but their basic ideas are still based on projection of one image to the shape of other image or image class.

In (Vizilter, Zheltov, 2012) the geometrical correlation approach was developed based on Pytiev's morphological image analysis. Let $f(x,y)$ from F is a piecewise-constant 2D function described above and image $g(x,y)$ from G is an analogous 2D function with m as a number of tessellation regions $\mathbf{G}=\{G_1,\dots,G_m\}$; $\mathbf{g}=(g_1,\dots,g_m)$ – vector of intensity values; $\chi_{Gj}(x,y)\in\{0,1\}$ – support function of j-th region. Let's introduce following additional set of "S-variables": S – area of the whole frame Ω; $S_i = \|\chi_{Fi}(x,y)\|^2$ – area of tessellation region F_i; $S_j = \|\chi_{Gj}(x,y)\|^2$ – area of tessellation region G_j; $S_{ij} = (\chi_{Fi}(x,y),\ \chi_{Gj}(x,y))$ – area of intersection $F_i \cap G_j$. *Mean square effective morphological correlation coefficient* (*MSEMCC*) for shapes F and G is determined as

$$K_M^2(F,G) = \sum_{j=1}^{m}\sum_{i=1}^{m} \frac{S_{ij}}{S}\frac{S_{ij}}{S_j} = \qquad, \quad (6)$$

$$= \sum_{j=1}^{m}\sum_{i=1}^{m} K_\Omega(F_i,G_j)\, K_M^2(G_j,F_i)$$

where $K_\Omega(F_i,G_j) = S_{ij}/S$ – normalized *influence coefficient* for pair of regions F_i and G_j; $K_M^2(G_j,F_i) = S_{ij}/S_j$ – square of *normalized morphological correlation* for pair of regions.

2.3 Manifold learning, diffusion maps and shape matching

The application of heat kernels and diffusion maps to shape analysis is initially inspired by the manifold learning technique developed for *nonlinear dimensionality reduction* (NLDR). Most interesting NLDR techniques are the following: Isomap (Tenenbaum et al, 2000), Locally Linear Embedding (LLE) (Roweis et al 2000), Kernel Principle Component Analysis (Scholkopf, 1999), Laplacian Eigenmaps (Belkin, Niyogi, 2001), Hessian LLE (Donoho, 2003), Manifold Sculpting (Gashler, 2008) and some other. The terms "heat kernel" and "heat dissipation" were introduced in (Belkin, Niyogi, 2001) in the context of Laplacian Eigenmap. In this concept they play the role of some manifold shape characteristics. Based on this, authors of (Lafon, 2004), (Coifman, Lafon 2006), (Coifman et al, 2007) introduced and developed the theory of diffusion maps (DM).

Let the manifold be described by some set of points $\mathbf{X}=\{\mathbf{x}_i\}$ in a high-dimensional space. The solution of NLDR problem in DM approach has a following form:

- Generate a *neighborhood graph G*.
- Form a *heat kernel* (matrix of pairwise similarity weights) $\mathbf{H} = \|h_{ij}\|$ using the rule. If i-th and j-th points are connected in G, then

$$h_{ij} = \exp\left(-\frac{\|x_i - x_j\|}{2\sigma^2}\right)$$

else $h_{ij}=0$.
- Normalize the heat kernel, and obtain the *diffusion kernel* $\mathbf{P} = \mathbf{M}^{-1}\mathbf{H}$, where $\mathbf{M} = \|m_{ij}\|$ is a diagonal matrix of column sums for $\mathbf{H}$: $m_{ii} = \sum_i h_{ij}$.
- Select the scale parameter t, and form the t-degree diffusion matrix $\mathbf{P}^t$, $t\geq 1$.
- Compute the spectral decomposition of $\mathbf{P}^t$ with eigenvalues

$$1 = \lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \dots$$

and corresponding eigenfunctions $\{\psi_i\}$
- Map the data to the low-dimensional vector space via selection of l eigenvalues and forming new coordinates based on corresponding eigenfunctions:

$$x_i = [\lambda_1\psi_{i1}\quad \lambda_2\psi_{i2}\quad \dots\quad \lambda_l\psi_{il}]^T$$

The normalization of heat kernel provides here both the linear smoothing *diffusion operator* P^t, and its interpretation in terms of *Markov chain* with *transition matrix* $\mathbf{P}^t$ for t steps of random walking. Thus, the Euclidean distance in this new space can be interpreted as a probability of point-to-point transition in t steps of this random walking. Such distance is called a *diffusion distance*, and it is very popular now in the area of high dimensional data analysis and machine learning.

Being inspired by NLDR task, the DM approach was later successfully applied for other types of data analysis problems. In particular, the following image restoration technique was outlined in (Coifman et al, 2007). Let $I(p)$ be a 2D image with $p=(x_p, y_p)$. And let pixels be described by some feature vector $v(p)$. Then for given $\varepsilon > 0$ the *diffusion kernel* is defined as

$$A_{p,q} = \frac{\exp - \dfrac{\|v(p) - v(q)\|^2}{\varepsilon}}{\sum\limits_{q} \exp \dfrac{-\|v(p) - v(q)\|^2}{\varepsilon}}$$

The filtering of image I by this *diffusion filter* has a usual form
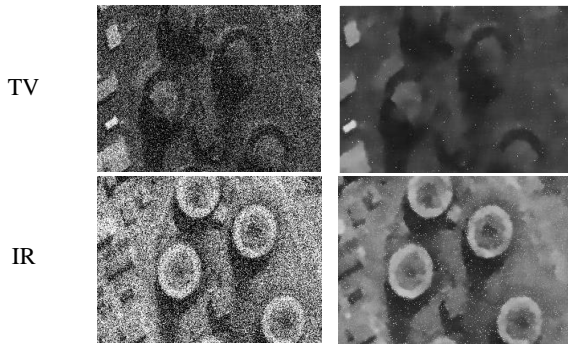
$$I_A(p) = \sum_q A_{p,q} I(q).$$

Figure 1. Examples of diffusion filtering for denoising of TV and IR images

Fig. 1 demonstrates the examples of TV and IR images diffusion filtering. In this case, like in paper (Coifman et al, 2007), $v(p)$ is just a 5×5 vector of grayscale image values in a 5×5 pixel neighborhood. Both for TV and for IR image the shape was preserved and the noise was essentially removed. The reason of such success in noise suppression is an adaptive smoothing with high kernel weights for similar neighbors and low or zero weights for dissimilar.

Let's note that idea of linear smoothing with adaptive kernel is utilized in different image restoration techniques, for example in (Takeda et al, 2007), (Milanfar, 2013). But the DM approach provides a unified way for shape description, shape comparison and shape-based image restoration.

Let's look at the data shape matching techniques based on heat kernels, diffusion maps and their spectral features. A number of such approaches including heat kernel signature (HKS), heat kernel spectrum (set of eigenvalues), heat kernel signature distribution (HKSD) and heat trace (HT) are proposed and developed in (Sun et al, 2009), (de Goes et al, 2008), (Lieu, Saito, 2008), (Reuter et al, 2006) and other papers. In the brilliant paper (Memoli, 2011) the overview, classification and unified mathematical description of these techniques are given. Moreover, the special Gromov-Wasserstein distances are

proposed for shape matching based on such spectral characteristics, and the stability (robustness) of such distances and matching procedures are theoretically proved (Memoli, 2011). These theoretical results are supported by impressive experiments with 3D models and real data collections. These results allow supposing, that the combination of DM and MIA approaches could create some effective 2D image and shape matching tools.

3. DIFFUSION MORPHOLOGY AND IMAGE MATCHING

In the first part of this section we propose a new generalized morphological framework based on diffusion shape models. In the second part the corresponding new morphological image and shape matching technique is described.

3.1 Generalized diffusion morphology

Let the image be a 2D function

$$f(x, y): \Omega \to R, \ \Omega \subset R^2$$

where R – set of real numbers, R^2 – image plane, Ω – rectangular frame region of image plane. Images are elements of Hilbert space $L^2(\Omega)$ with *scalar product* (f,g) and *norm* $\|f\| = (f,f)^{1/2}$.

Let's introduce the generalized definitions of basic MIA notions by substitution of "mosaic shape" to "diffusion shape" in order to obtain the relaxed version of MIA.

The *relational model* of *diffusion shape F* for image f is a heat kernel $h_F(x,y,u,v)$: $\Omega \times \Omega \to [0,1]$, such that

$$h_F(x, y, x, y) \geq h_F(x, y, u, v); \ h_F(x, y, u, v) \geq h_F(u, v, x, y),$$

and the unique *basic similarity measurement function* η exists providing

$$\eta(f(x, y), f(u, v)) = h_F(x, y, u, v).$$

The *operator model* of *diffusion shape F* is a diffusion operator P_F with normalized diffusion kernel $p_F(x,y,u,v)$, such that

$$P_F g(x, y) = \iint_{\Omega} p_F(x, y, u, v) \, g(u, v) \, du \, dv$$

$$p_F(x, y, u, v) \geq 0; \ p_F(x, y, x, y) \geq p_F(x, y, u, v);$$

$$\iint_{\Omega} p_F(x, y, u, v) \, du \, dv = 1;$$

$$\iiint_{\Omega \times \Omega} \int p_F^2(x, y, u, v) \, dx \, dy \, du \, dv < \infty.$$

So, any relational model h_F corresponds to the unique operator model with normalized kernel p_F:

$$p_F(x, y, u, v) = h_F(x, y, u, v) \Big/ \iint_{\Omega} h_F(x, y, a, b) \, da \, db$$

The *spatial model* of *diffusion shape F* for image $f(x,y)$ with precision n is an eigenspace of diffusion operator P_F

$$F = \text{span}\{\lambda_1 \varphi_1(x, y), \ldots, \lambda_n \varphi_n(x, y)\}$$

where $\{\lambda_1, \ldots, \lambda_n\}$ – n first eigenvalues, a $\{\varphi_1(x,y), \ldots, \varphi_n(x,y)\}$ – n first eigenfunctions of *morphological diffusion operator P_F*:

$$P_F \varphi_i(x, y) = \lambda_i \varphi_i(x, y), i = 1, \ldots n.$$

It is easy to see, that in particular case of heat kernel of the form

$$h_F^*(x, y, u, v) = \begin{cases} 1, & \text{if } f(x, y) = f(u, v); \\ 0, & \text{otherwise} \end{cases}$$

it expresses the binary relation "points of equal values", and the diffusion-based morphological definitions stated above degrade to the MIA definitions given in section 2.2. Really, in this case the similarity relation η becomes an equivalent relation on the image frame points splitting them to a set of non-overlapping regions $\mathbf{F} = \{F_1, \ldots, F_n\}$, where n is a number of regions of frame tessellation $\mathbf{F}$. Hence, the morphological diffusion operator becomes a morphological projector $P_F = P_F P_F$ with n 1-valued eigenvalues and characteristic (support) functions $\varphi_i(x, y) = \chi_{Fi}(x, y)$ (2) of n regions as eigenfunctions. So, the diffusion shape in this particular case becomes a mosaic shape F of the form (3). And for any $g(x, y) \in L^2(\Omega)$ the diffusion filtering P_F becomes a morphological projection onto the mosaic shape F of the classic form (4).

Thus, the Pytiev MIA approach is a particular case of generalized *diffusion morphology* described in this section. So, all shape analysis schemes and tools of MIA can be recovered on this wider basis just using the diffusion operator instead of Pytiev morphological projector.

Let's note that the formal definition of diffusion morphology can start directly from a heat kernel of the classic form

$$h_{ij} = \exp(-d_{ij}/\varepsilon),$$

where d_{ij} – *basic distance* between i-th and j-th points of discrete digital image. In this formulation of diffusion morphology the MIA case corresponds to a special selection of basic distance as a *discrete distance by pixel value*:

$$d_{ij} = \begin{cases} 1, & \text{if } f(x_i, y_i) = f(x_j, y_j); \\ \infty, & \text{otherwise}. \end{cases}$$

Then the heat kernel takes a form $h_{ij} = h_F^*(x_i, y_i, u_j, v_j)$ of binary relation "points of equal values" that forces the transformation of diffusion operator to the mosaic projector.

The final note here is that different choice of descriptors and metrics for pixel or/and region comparison provides the design of different diffusion morphologies with different semantic properties. Additionally, one can use the diffusion filter P^t with scale parameter t for morphological scale space analysis.

Thus, this generalized morphological approach provides more information about the image shape than just the information about shape of frame tessellation exploited by MI or original MIA. And if this new information is robust relative to noise, it will support the higher quality of matching.

3.2 Image and shape matching based on diffusion morphology

In Pytiev morphology the comparison of image $g(x, y)$ and shape of image $f(x, y)$ is performed using the normalized morphological correlation coefficient of the following form

$$K_M(g, F) = \frac{\| P_F g \|}{\| g \|},$$

and this coefficient satisfies the property $K_M(f, F) = 1$ due to the fact that $f = P_F f$. The diffusion morphological operator is not a projector, but it is a smoothing filter with $\| P_F f \| \le \| f \|$. Nevertheless, it is natural to suppose that the smoothing power of P_F will be essentially less for images with similar shapes than for images with different shapes. So, the *morphological diffusion correlation coefficient* (MDCC) is defined as a ratio of Pytiev morphological coefficients

$$K_{MD}(g, F) = \frac{K_M(g, F)}{K_M(f, F)} = \frac{\| P_F g \| \, \| f \|}{\| P_F f \| \, \| g \|},$$

where $K_M(f, F)$ describes the power of self-smoothing of f by F. Let's note that MDCC is a correct generalization of MCC, because in case of projective morphology $\| f \| = \| P_F f \|$ and $K_{MD}(g, F) = K_M(g, F)$.

As in MCC, for elimination of non-informative part of image brightness images should be normalized before comparison:

$$K_{MD}(g, F) = \frac{\| P_F g - P_O g \| \, \| f - P_O f \|}{\| P_F f - P_O f \| \, \| g - P_O g \|}, \quad (7)$$

where $P_O f$ – morphological filtering of image f by the "empty" diffusion shape O. But in diffusion morphology such normalization is not the trivial subtraction of global mean value (like in Pytiev MIA). It is a subtraction of mean value in a sliding window determined by support neighborhood (effective size) of heat kernel. This subtraction preserves the local informative features only (in the corresponding scale of analysis). These informative elements of image g will be passed (if the shape G is similar to shape F) or extremely smoothed (if the shape G is essentially different) by the diffusion filter P_F. This trick is called *morphological image normalization* (Fig.2).
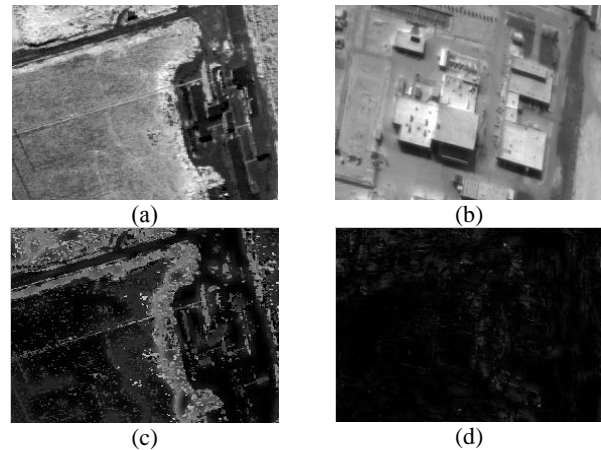

(a)   (b)

(c)   (d)

Figure 2. Example of morphological image normalization: a) image f; b) image g; c) image f normalized by self-shape F ($P_F f - P_O f$); d) image g normalized by shape F ($P_F g - P_O g$).

If the effective size of heat kernel is small, such diffusion image-to-shape matching technique uses the local features only like the points-based and contour-based matching techniques. Fortunately, as we stated above, the theory of diffusion maps has a natural instrument for multi-scale data analysis – parameter t (number of Markov random walking steps). The description of the image shape by the set of different scale diffusion operators $\{P^t\}$ allows performing the *morphological scale-space analysis*.

Finally, as well as for Pytiev correlation coefficient, we prefer to use the square of MDCC instead of MDCC, because both $K_M^2(g,F)$ and $K_{MD}^2(g,F)$ can be interpreted as statistical *coefficient of determination* between the model (shape F) and observed data (image g). So, the generalized morphological technique for image-to-shape matching is formed.

4. IMPLEMENTATION AND EXPERIMENTAL RESULTS

In the first part of this section we describe the original fast implementation of diffusion filters and morphological diffusion operators. This implementation provides a possibility for real-time image processing. It was utilized in our comparative experiments with different shape matching techniques for TV and IR image matching task. These experiments and corresponding results are outlined in the second part of this section.

4.1 Fast implementation of morphological diffusion operators

The computation of diffusion filtering with heat kernel of the form

$$h(p,q) = \exp\left(-\frac{\|v(p) - v(q)\|^2}{\varepsilon}\right),$$

where $v(p)$ is a neighborhood of image point p, is an extremely time-consuming procedure even for reasonable size of $v(p)$. In this paper we propose to substitute such computationally unpleasant descriptors by the new type of point feature descriptors – *iLBP* (intensity + LBP):

$$iLBP(p) = (m(p),\ LBP(p)),$$

where $m(p)$ – mean value of $v(p)$; $LBP(p)$ – threshold LBP (Ahonen, 2004) for $v(p)$. The local binary pattern (LBP) is calculated here as an 8-bit vector for each pixel p based on a comparison of its value and values of its 8 nearest neighbors. If the value of neighbor pixel is less than the value of central pixel and the difference between them is greater than threshold, then the corresponding bit is set to 1, otherwise – to 0. Correspondingly we substitute the original neighborhood matching metrics by the special iLBP matching metrics:

$$\|v(p) - v(q)\|^2 \sim \beta|m(p) - m(q)|^2 + d_{ham}(LBP(p), LBP(q))$$

where d_{ham} – Hamming distance, β – importance weight for intensity part of iLBP. Local binary patterns are stored as bit fields, and the computation of Hamming distance is performed via bitwise XOR operation. The exponent is calculated using table values. Mean value in a sliding window is computed by a fast algorithm with sliding sum recalculation. Due to this, the usage of iLBP allows both increasing the computational speed and obtaining heat kernels very similar to original. The software implementation of this idea provides the processing time about 300 ms for image size 640×480 and PC with Core i7 – 860 /8Gb/ GeForce GTX 680 in one thread.

The application of CUDA technology (NVIDIA, 2014) additionally improves the performance of diffusion filtering. Our current CUDA-based implementation provides the processing time about 40 ms for image size 640×480 and the same PC configuration. Thus, it allows performing the diffusion filtering and correspondingly the proposed shape-based matching procedure in real-time.

4.2 TV and IR matching experiments and results

The scheme of experimental comparison of matching techniques over a set of real images is following. Some clear fragments of TV image ("etalon" fragments) are compared with all equal-sized fragments of corresponding IR image without segmentation. TV fragments are segmented with n =4 levels using least square optimal segmentation procedure.

For segmented image fragments f (TV) and non-segmented g (IR) at each fragment g position (x,y) following similarity measures are calculated: mutual information $MI(F,G)$ (1); square of centered Pytiev morphological coefficient $K_M^2(f,G)$ (5); square of MSEMCC $K_M^2(F,G)$ (6); square of MDCC $K_{MD}^2(f,G)$ (7).

The quality of these measures is estimated by following statistics of 2D correlation function $C(x,y)$ – *signal-to-noise ratio* (SNR) and *exceeding* of first maximum to second one (E):

$$SNR = \frac{|C_1 - \mu|}{\sigma};\quad E = \frac{|C_1 - \mu|}{|C_2 - \mu|},$$

where C_1 – global maximum of correlation value; C_2 – second global maximum of correlation value out of some small neighborhood of first global maximum; μ – mean value of correlation function; σ – dispersion of correlation function.

Figures 3, 5, 7, 9 demonstrate four matching examples (correspondingly Examples 1, 2, 3, 4) of experimental clear TV data and IR data corrupted by Gaussian noise. Figures 4, 6, 8, 10 demonstrate corresponding 2D-graphs of correlation fields for different matching scores. Corresponding numerical data are listed in tables 1-4.

According to these experiments, proposed MDCC provides essentially higher and sharper correlation peak relative to other similarity measures including state-of-art mutual information score.

Measure	Max value	SNR	E
$MI(F,G)$	0.10812	4.4657	3.1907
$K_M^2(f,G)$;	0.1006	4.5859	2.7913
$K_M^2(F,G)$	0.29344	4.5531	3.3613
$K_{MD}^2(f,G)$	0,00339	9.9809	4.6698

Table 1 Numeric data for TV-IR matching (Example 1).

Measure	Max value	SNR	E
$MI(F,G)$	0.20079	8.2561	1.3156
$K_M^2(f,G)$;	*(mismatch)*	-	-
$K_M^2(F,G)$	0.49002	9.0658	1.3958
$K_{MD}^2(f,G)$	0.00807	15.2569	1.5128

Table 2 Numeric data for TV-IR matching (Example 2).

Measure	Max value	SNR	E
$MI(F,G)$	0.048	3.0601	1.0293
$K_M^2(f,G)$;	0.0391	2.8682	1.2919
$K_M^2(F,G)$	0.39671	3.5967	1.1626
$K_{MD}^2(f,G)$	0,00103	10.2286	3.8530

Table 3 Numeric data for TV-IR matching (Example 3).

a)

b)

c)

Figure 3. Example 1: a) etalon TV fragment; b) test IR image; c) segmented TV fragment.
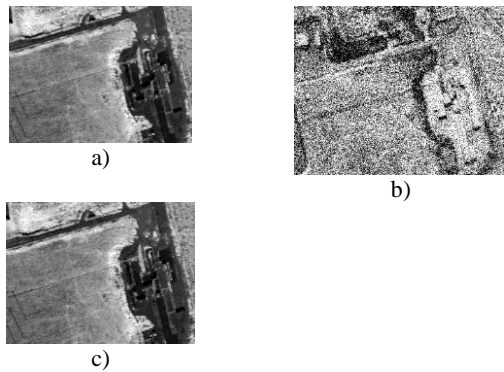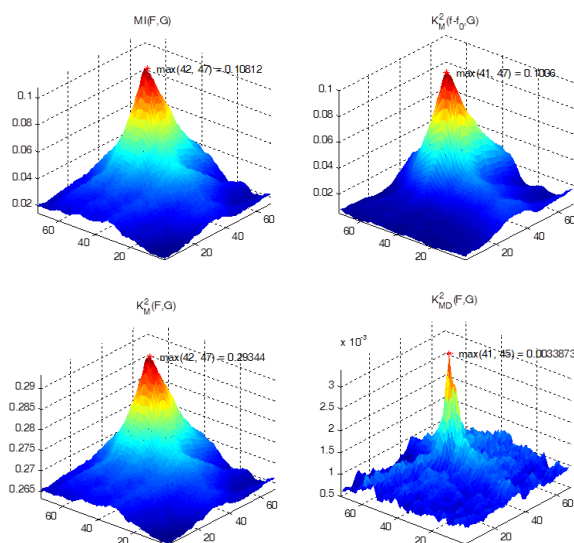

a)

b)

c)

Figure 5. Example 2: a) etalon TV fragment; b) test IR image; c) segmented TV fragment.

Figure 4. Correlation fields for TV-IR matching (Example 1).
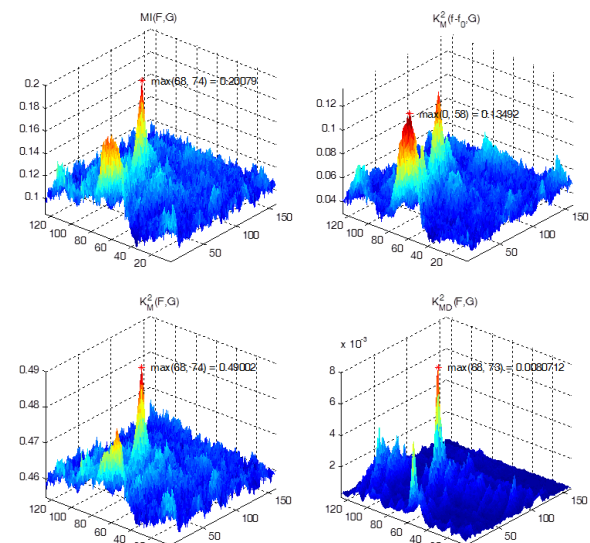
Figure 6. Correlation fields for TV-IR matching (Example 2).

Measure	Max value	SNR	E
$MI(F,G)$	0.17752	5.9937	1.2923
$K_M^2(f,G);$	0.12125	7.9213	1.5949
$K_M^2(F,G)$	0.41667	6.4704	1.3463
$K_{MD}^2(f,G)$	0,00216	15.4247	1.7667

Table 4 Numeric data for TV-IR matching (Example 4).

In some experiments MDCC provides SNR in 2-3 times better. Examples 2 and 3 demonstrate the higher stability of proposed MDCC relative to noise in comparison with mutual information (1), Pytiev MCC (4) and MSEMCC (5).

5. CONCLUSION

This paper presents the new image-to-shape matching technique based on heat kernels and diffusion maps.
The theoretical contribution of the paper is a formulation of Diffusion Morphology. It is designed as a generalization of Pytiev morphological image analysis (MIA) and contains MIA as a particular case.
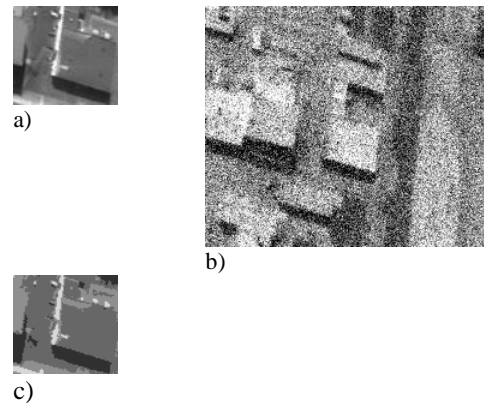
All shape analysis tools of MIA can be applied in diffusion morphology too using the morphological shape-based diffusion operator instead of morphological shape-based projector.

Practical contributions of the paper are the new tool for image shape-based matching called morphological diffusion correlation coefficient (MDCC) and the fast implementation of morphological diffusion filtering using the new type of point feature descriptors – iLBP (intensity + LBP). This implementation allows performing the diffusion filtering and correspondingly the proposed image-to-shape matching procedure in real-time.

Experimental comparison of shape-based matching techniques is performed applying to the TV and IR image matching problem. The square of MDCC is tested relative to the mutual information, square of centered Pytiev morphological correlation coefficient (MCC) and square of mean square effective morphological correlation coefficient (MSEMCC). The quality of matching is estimated by two statistics of correlation field: signal-to-noise ratio (SNR) and exceeding of first maximum to second one (E).
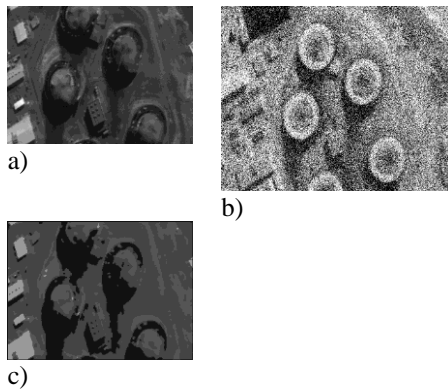
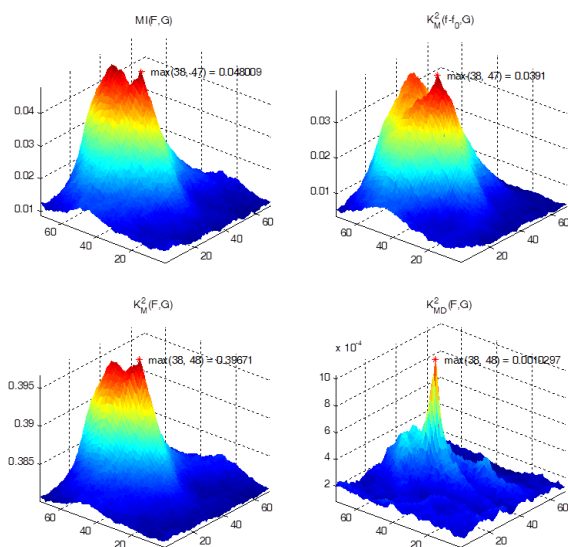Figure 7. Example 3: a) etalon TV fragment; b) test IR image;
c) segmented TV fragment.

Figure 9. Example 4: a) etalon TV fragment; b) test IR image;
c) segmented TV fragment.

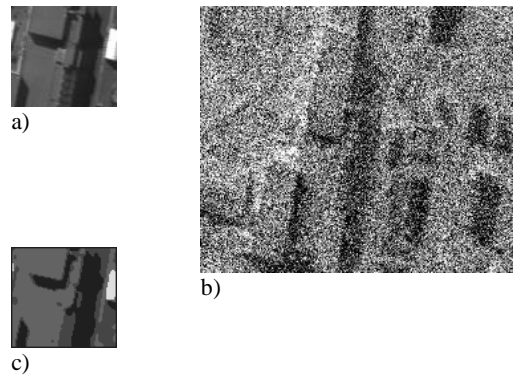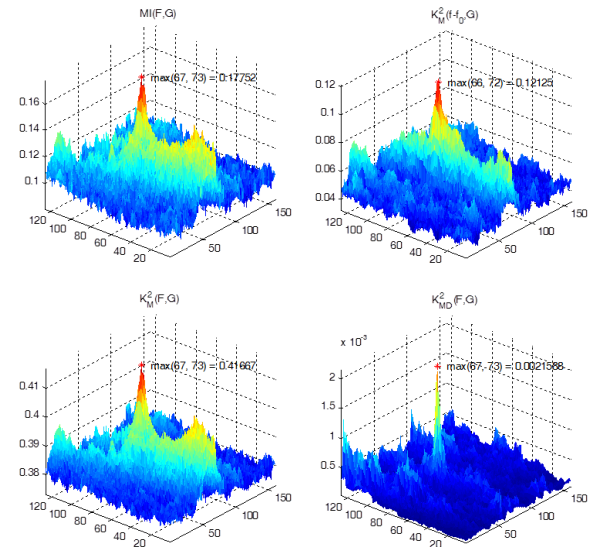Figure 8. Correlation fields for TV-IR matching (Example 3)

Figure 10. Correlation fields for TV-IR matching (Example 4)

In our experiments with airborne TV and IR imagery the proposed MDCC totally outperforms all compared matching scores by both statistics, especially, in SNR.

The shape of correlation field is essentially better (sharper) than for all concurrent coefficients. In particular, relative to the state-of-art mutual information approach this new morphological technique provides SNR in 2-3 times better. Thus, we can conclude that proposed MDCC is a best correlation score for TV and IR matching problem.

REFERENCES

Ahonen, T., Hadid, A., Pietikainen, M., 2004. Face recognition with local binary patterns. ECCV, pp 469-481

Belkin M., Niyogi P., 2001. Laplacian eigenmaps and spectral techniques for embedding and clustering. In Advances in Neural Information Processing Systems, 14, pages 585-591.

Coifman, R., Lafon, S., 2006. Diffusion maps, Appl. Comp. Harm. Anal.

Coifman, R., Lafon, S., Maggioni, M., Keller, Y., Szlam, A., Warner, F., Zucker, S., 2007. Geometries of sensor outputs, inference and information processing, NIPS.

de Goes, F., Goldenstein, S., Velho, L., 2008. A hierarchical segmentation of articulated bodies, Comput. Graph. Forum 27 (5), pp 1349–1356.

Donoho, D., Grimes, C., 2003. Hessian eigenmaps: locally linear embedding techniques for high dimensional data. Proc. of National Academy of Sciences, 100(10), pp 5591–5596.

Falomkin, I. I., Pyt'ev, Yu. P., 2007. Algorithm of Adaptive Morphological Filtering of Images, Pattern Recognition and Image Analysis, 2007, Vol. 17, No. 3, pp. 408–420.

Gashler, M., Ventura, D., Martinez, T., 2008. Iterative Non-linear Dimensionality Reduction by Manifold Sculpting., Advances in Neural Information Processing Systems 20, pp 513–520, MIT Press, Cambridge, MA.

Goebel, B. et al, 2005, An Approximation to the Distribution of Finite Sample Size Mutual Information Estimates,

Communications, 2005. ICC 2005. 2005 IEEE International Conference on (Volume:2), pp 1102-1106.

Ji, Y. et al, 2005, Direct and Recursive Prediction of Time Series Using Mutual Information Selection, Computational Intelligence and Bioinspired Systems, Lecture Notes in Computer Science Volume 3512, , pp 1010-1017

Lafon, S., 2004. Diffusion maps and geometric harmonics. PhD thesis, Yale University, Dept of Mathematics & Applied Mathematics.

Lieu, L., Saito, N., 2008. Pattern recognition by diffusion maps and earth mover's distance, Tech. rep., http://www.math.ucdavis.edu/~llieu/PRbyDMEMD.pdf.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. and Suetens, P., 1997. Multimodality Image Registration by Maximization of Mutual Information. IEEE Transactions on Medical Imaging, Vol. 16, No. 2, April, pp.187-198.

Memoli, F., 2011. A spectral notion of Gromov–Wasserstein distance and related methods, Appl. Comput. Harmon. Anal. 30, pp 363–401.

Milanfar P., 2013. A Tour of Modern Image Filtering, IEEE Signal Processing Magazine vol. 30, no. 1, January, pp. 106-128

NVIDIA, 2014. NVIDIA CUDA Compute Unified Device Architecture, http://www.nvidia.ru/object/cuda-parallel-computing-ru.html

Pyt'ev, Yu., 1993. Morphological Image Analysis. Pattern Recognition and Image Analysis. V.3. No 1, pp. 19-28.

Reuter, M., Wolter, F.-E., Peinecke, N., 2006. Laplace–Beltrami spectra as "Shape-DNA" of surfaces and solids, Comput. Aided Design 38 (4), pp 342–366.

Roweis, S. T., Saul, L. K, 2000. Nonlinear dimensionality reduction by locally linear embedding. Science, 290, pp 2323–2326.

Scholkopf, B., Smola, A. J., Muller, K.-R., 1999. Kernel principal component analysis. Advances in kernel methods: support vector learning, pages 327–352.

Sun, J., Ovsjanikov, M., Guibas, L., 2009. A concise and provably informative multi-scale signature based on heat diffusion, in: SGP.

Takeda H. et al., 2007. Kernel Regression for Image Processing and Reconstruction, IEEE Trans. Image Processing.

Tenenbaum, J. B., de Silva, V., Langford, J. C., 2000. A global geometric framework for nonlinear dimensionality reduction. Science, 290:2319–2323.

Vizilter, Yu., Zheltov, S., 2009. The Use of Projective Morphologies for Object Detection and Identification in Images // Journal of Computer and Systems Sciences International, Vol. 48, No. 2, pp. 282–294.

Vizilter, Yu., Zheltov, S., 2012. Geometrical Correlation and Matching of 2D Image Shapes, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., I-3, 191-196, doi:10.5194/isprsannals-I-3-191-2012.