Object matching by means of matching likelihood coefficients

Jan Flusser *

Institute of Information Theory and Automation, Czech Academy of Sciences, Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic

Abstract

The paper deals with matching of two sets of objects, which differ from each other by translation, rotation and scaling. A new two-stage algorithm for feature-based object matching has been developed. In order to reach accurate matching results and to reduce computing complexity, it uses local information (represented by a set of invariant features) as well as information about object-to-object distances in the image plane. To measure the reliability of determining correspondence between two objects, matching likelihood coefficients are introduced.

An improved version of shape matrix method for description of 2-D shapes is introduced in this paper. The shape matrices are used as the features in the first stage of the matching algorithm.

Keywords: Object matching; Shape description; Similarity measure; Matching likelihood coefficients; Shape matrix

1. Introduction

Analysis of two digital images of the same scene taken at different times, from different places or by different sensors often requires automatic matching of the images. By *image matching* we understand the process of determining mutual correspondence between two sets of objects.

The problem of image matching seems to be similar to the well-known feature-based classification (when the objects in the first image represent classes and the objects in the second image represent unknown patterns), but in image matching an important additional information is available:

- Each object of the first image cannot have more than one corresponding object in the second image, i.e. each class either contains only one element or is empty;

- The images differ from each other by some kind of geometric transformation which is supposed to be known (its coefficients, however, are unknown).

Object detection can be done by well-known segmentation techniques (see (Haralick and Shapiro, 1985) for a survey). Two binary images with extracted objects are obtained as the results of the segmentation.

There were described many matching methods in the literature. These methods are based on combinatorial approach (Goshtasby and Stockman, 1985), graph matching (Zahn, 1974), parameters clustering (Stockman et al., 1982), probabilistic relaxation (Ranade and Rosenfeld, 1980) and logical tree classification (Ventura et al., 1990). Generally, most of them have high computing complexity or low reliability. The algorithm described below combines local and global approaches in order to reduce computing complexity and to reach high robustness.

---

* Email: flusser@utia.cas.cz

2. Object matching algorithm

Let us consider matching of two images which differ from each other by a similarity transform (i.e. by translation, rotation and scaling). Let us denote N and M the numbers of objects in the first and second images, respectively.

We suppose that any object A is represented by its feature vector $\mathbf{A}$, which is invariant under similarity transform. At this moment, it is not essential what kind of features is used, because the proposed matching algorithm is feature-independent.

We suppose that a similarity relation $p(A, B)$ between the objects A and B is defined. If the feature space is a metric one with metric ϱ, $p(A, B)$ can be defined for instance by the following formula:

$$p(A, B) = \frac{1}{1 + \varrho(\mathbf{A}, \mathbf{B})}. \tag{1}$$

The presented algorithm for object matching consists of two stages. The first stage is performed in the feature space, the second one is performed in the image space.

In the first stage, the local information about the objects represented by their feature values is used to find two pairs of the most likely corresponding objects.

Let us denote $F_1, \dots F_N$ and $G_1, \dots G_M$ the objects detected in the first and second images, respectively. Define the $N \times M$ matrix $\mathbf{K}$ by the relation

$$\mathbf{K}_{ij} = p\left(F_i, G_j\right), \quad i = 1, \dots, N; \ j = 1, \dots, M.$$

For each i find $j_i \leqslant M$ such that $\mathbf{K}_{ij_i} = \max_j \mathbf{K}_{ij}$.

In order to minimize the possibility of the false match, we define the *matching likelihood coefficients*. For each i we find an index $\ell_i \leqslant M$ such that $\mathbf{K}_{i\ell_i} = \max_{j \neq j_i} \mathbf{K}_{ij}$.

The matching likelihood coefficient d_i is defined by the formula

$$d_i = \mathbf{K}_{ij_i} \cdot \left(\mathbf{K}_{ij_i} - \mathbf{K}_{i\ell_i}\right), \quad i = 1, \dots, N$$

and expresses the degree of reliability, with which objects F_i and G_j correspond with each other.

Now we can find indexes i_0 and i_1 which describe the most reliable matching:

$$d_{i_0} = \max_i d_i, \qquad d_{i_1} = \max_{i \neq i_0} d_i.$$

In this way, we have determined two pairs of objects corresponding with maximum likelihood:

$$F_{i_0} \approx G_{j_{i_0}} \quad \text{and} \quad F_{i_1} \approx G_{j_{i_1}}.$$

In the second stage, having two pairs of corresponding points (the centres of gravity of the objects $F_{i_0}, G_{j_{i_0}}, F_{i_1}$ and $G_{j_{i_1}}$), the parameters of the geometric transformation between the images can be simply estimated. The type of the transformation is supposed to be

$$u = c\,(x \cdot \cos\alpha - y \cdot \sin\alpha) + a$$
$$v = c\,(x \cdot \sin\alpha + y \cdot \cos\alpha) + b, \tag{2}$$

where (x, y) and (u, v) are the coordinates in the first and second images, respectively, a and b denote the translation in horizontal and vertical directions, c is the scaling factor and α is the angle of rotation.

Knowing the parameters, we can map the objects of the first image into the second image. Object-to-object correspondence is then established by the nearest neighbour rule. However, some rejection threshold should be applied to avoid false match.

More formally, the whole algorithm can be described as follows:

Algorithm MLC (*Matching by Matching Likelihood Coefficients*)

1. Denote $F_1, \dots, F_N$ the objects in the first image and $G_1, \dots, G_M$ the objects in the second one.
2. Construct two point sets

$$\mathcal{Q} = \{(x_i, y_i) \mid x_i, y_i \text{ are coordinates of the center of gravity of object } F_i\},$$

$$\mathcal{S} = \{(u_j, v_j) \mid u_j, v_j \text{ are coordinates of the center of gravity of object } G_j\}.$$

3. Compute the $N \times M$ matrix $\mathbf{K}$

$$\mathbf{K}_{ij} = p\left(F_i, G_j\right).$$

4. **for** $i = 1$ **to** N
   find indexes j_i and ℓ_i such that

$$\mathbf{K}_{ij_i} = \max_j \mathbf{K}_{ij}, \qquad \mathbf{K}_{i\ell_i} = \max_{j \neq j_i} \mathbf{K}_{ij}$$

   and compute the matching likelihood coefficient

$$d_i = \mathbf{K}_{ij_i} \cdot \left(\mathbf{K}_{ij_i} - \mathbf{K}_{i\ell_i}\right).$$

   endfor

5. Find indexes i_0 and i_1 such that

$$d_{i_0} = \max_i d_i, \qquad d_{i_1} = \max_{i \neq i_0} d_i.$$

6. Solve the following system of linear equations with variables $c \cdot \cos \alpha, c \cdot \sin \alpha, a$ and b:

$$u_{j_{i_0}} = c \left(x_{i_0} \cdot \cos \alpha - y_{i_0} \cdot \sin \alpha \right) + a$$

$$v_{j_{i_0}} = c \left(x_{i_0} \cdot \sin \alpha + y_{i_0} \cdot \cos \alpha \right) + b$$

$$u_{j_{i_1}} = c \left(x_{i_1} \cdot \cos \alpha - y_{i_1} \cdot \sin \alpha \right) + a$$

$$v_{j_{i_1}} = c \left(x_{i_1} \cdot \sin \alpha + y_{i_1} \cdot \cos \alpha \right) + b.$$

7. Define a distance threshold $r > 0$.
   for $i = 1$ **to** N compute

$$u = c \left(x_i \cdot \cos \alpha - y_i \cdot \sin \alpha \right) + a$$

$$v = c \left(x_i \cdot \sin \alpha + y_i \cdot \cos \alpha \right) + b.$$

Let $(u_j, v_j) \in \mathcal{S}$ be the closest point to (u, v). If

$$(u - u_j)^2 + (v - v_j)^2 < r^2,$$

then F_i and G_j are considered as a corresponding pair, otherwise the object F_i is marked to have no corresponding object in the second image.
endfor

8. Final check. Define a rejection threshold $t \in (0; 1)$. For each couple of corresponding objects F and G compare $p(F, G)$ with t. If $p(F, G) < t$, reject the objects F and G from the list of corresponding objects.

Now we shall discuss some basic properties of algorithm MLC.

- *Computing complexity.* Computing complexity of the algorithm is $O(M \cdot N)$ regardless of the type of features used (computing complexity of object feature evaluation does not depend on the total number of objects).

- *Robustness.* The first stage of algorithm MLC (i.e. steps 1–5) could generate a false match in some particular cases. If all objects have very similar shapes, their feature vectors are close to each other and the objects are not well-separated in the feature space. The matching results could be then significantly affected even by small noise or by inaccurate segmentation of object boundaries. The second stage (i.e. steps 6 and 7) cannot fail, if the first stage produced a correct match.

The robustness can be controlled by appropriate threshold selection in steps 7 and 8. The less the threshold r is (and similarly the higher the threshold t is), the less likely the false match is but, however, the less number of corresponding pairs is found.

The matching results in every particular task are also affected by the features used for object description (especially by their stability with respect to shape distortions and random noise). Anyway, algorithm MLC gives much better results than minimum-distance shape matching in the feature space.

- *Adaptability.* The algorithm can be simply modified to deal with more general geometric transformations between the input images (affine or projective transform for instance). Let the transform be uniquely determined by P pairs of corresponding points ($P = 2$ for similarity transform, $P = 3$ for affine transform, $P = 4$ for projective transform, etc.). In step 5, we find P indexes $i_0, \ldots, i_{P-1}$ so that

$$d_{i_j} = \max_{i \neq i_0, \ldots, i_{j-1}} d_i, \quad j = 0, 1, \ldots, P - 1.$$

A system of $2P$ equations is then solved in step 6 to estimate the transform coefficients which are required in step 7.

3. The features for shape description

Although the algorithm MLC can work with any kind of features, the selection of appropriate shape descriptors plays an important role and can affect the matching results. Traditionally, the features used for 2-D shape description are categorized into five groups:

- Visual features (i.e. features directly describing a shape, such as shape vector (Peli, 1981) and shape matrix (Goshtasby, 1989; Taza and Suen, 1989; Flusser, 1992));
- Transform coefficient features (Fourier descriptors (Lin and Chellappa, 1987; Kiryati, 1988) or Hadamard coefficients (Kuncheva, 1988));
- Algebraic features (based on matrix decomposition of the image (Hong, 1991));
- Moment-based invariants (Hu, 1962; Flusser and Suk, 1993; Reiss, 1993);

• Differential invariants (used specially for curved objects (Weiss, 1988)).

In this paper, we present a new shape descriptor belonging to the first category which is called *shape matrix* (SM).

The idea of a shape matrix was originally introduced by Goshtasby (1989) and performs a generalization of Peli's shape vector (Peli, 1981). The main idea is to define polar coordinates with the origin in the center of gravity of the object. Then, starting from the maximum radius, the binary object B is sampled with steps θ and Δ, where $\theta = 2\pi/m$, $\Delta = L/n$, L is the length of the maximum radius and m, n are desired dimensions of the shape matrix. In this way, each sample uniquely corresponds to one element of binary shape matrix **B**. If the point with polar coordinates $(i \cdot \theta, j \cdot \Delta)$ lies inside the shape, then $\mathbf{B}_{ij} = 1$; otherwise $\mathbf{B}_{ij} = 0$.

The method described in (Goshtasby, 1989) has a major weakness. The density of sample points is higher near the center then near the outermost circle. Therefore, the elements of the shape matrix in different columns contain different type of information about the original shape. This is a significant phenomenon.

Several attempts to overcome this weakness have been recently published. Taza and Suen (1989) introduced the "weight" for each column of the shape matrix. The weight of a column is directly proportional to the order of this column. However, this technique improves original Goshtasby's method only slightly.

In this section, the shape matrix method is significantly improved and a new type of the shape matrix construction is introduced. Uniform sampling in a square grid is proposed instead of polar sampling. Therefore, the sample points are equidistantly distributed over the shape. In this way, the shape is encoded into a binary matrix with optional resolution. The presented method is able to describe every planar object. There is no limit to the scope of the shapes that the shape matrix can represent. The ability to describe even shapes with holes is a very important property of the shape matrix. In many instances, information about internal geometry of a shape considerably enhances the shape recognition. Naturally, the "new" shape matrix keeps all nice properties of the "old" one – it is simple to construct, shape preserving and invariant under translation, rotation and scaling.

For given planar object B we define its shape matrix of size $n \times n$ as follows.

Algorithm SMC (*Shape Matrix Construction*)
1. Find the centre of gravity $T = (x_T, y_T)$ of object B.
2. Find point $R = (x_R, y_R)$ such that $R \in B$ and $e(R, T) = \max_{X \in B} e(X, T)$, where e is the Euclidean metric in the plane.
3. Construct the square with the centre at T and with the side length $2 \cdot e(R, T)$. Point R lies at the centre of a side of the square.
4. Divide the square into $n \times n$ subsquares S_{kj}; $k, j = 1, \ldots, n$.
5. Define the $n \times n$ binary shape matrix **B** as follows:

$$\mathbf{B}_{kj} = \begin{cases} 1 & \text{iff } \mu\left(S_{kj} \cap B\right) \geqslant \mu\left(S_{kj}\right)/2, \\ 0 & \text{otherwise,} \end{cases}$$

where $\mu(F)$ denotes the area of the planar region F.

It is easy to prove that the shape matrix is invariant to translation, rotation and scaling of the object. Moreover, the shape of the object can be reconstructed from the shape matrix. The accuracy of the reconstruction is given by the size of the subsquares S_{kj}, i.e. we can reach more accurate (and, of course, more computationally expensive) shape description by increasing n. An algorithm for determination of optimal shape matrix dimension can be found in (Flusser, 1992).

To determine the degree of similarity between two objects, their shape matrices **A** and **B** are compared. However, the dimensions of the matrices must be equal. The similarity relation $p(A, B)$ is defined by the following formula:

$$p(A, B) = 1 - \frac{1}{n^2} \sum_{j=1}^{n} \sum_{i=1}^{n} |\mathbf{A}_{ij} - \mathbf{B}_{ij}|. \tag{3}$$

Note that it always holds

$$0 \leqslant p(A, B) \leqslant 1. \tag{4}$$

However, there exist shapes with more than one maximum radius (i.e. more than one point R is found during the second step of algorithm SMC) which could produce different shape matrices depending on the maximum radii used. Let the object A be described by s different shape matrices $\mathbf{A}^1, \ldots, \mathbf{A}^s$ and the object B be described by q different shape matrices $\mathbf{B}^1, \ldots, \mathbf{B}^q$. We define the similarity $p(A, B)$ by the following relation:
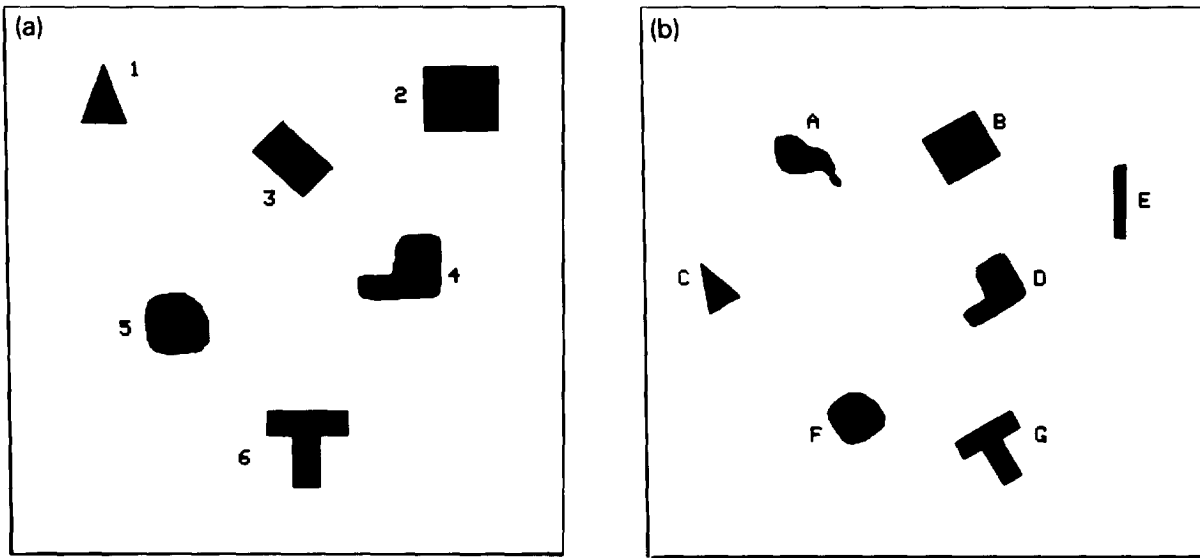
Fig. 1. Two sets of objects to be matched.

Table 1

Matrix **K** and the matching likelihood coefficients (all values of **K** were multiplied by 100)

	1	2	3	4	5	6	d
A	88	61	68	83	45	76	440
B	62	85	81	69	80	65	340
C	96	65	68	84	49	78	1152
D	81	73	71	93	59	77	1116
E	77	53	56	72	37	65	385
F	49	82	73	59	88	58	528
G	79	66	65	76	60	98	1862

$$p(A, B) = \max_{k,\ell} p^{k\ell}(A, B),$$

$$k = 1, \ldots, s; \quad \ell = 1, \ldots, q \qquad (5)$$

where

$$p^{k\ell}(A, B) = 1 - \frac{1}{n^2} \sum_{j=1}^{n} \sum_{i=1}^{n} \left| \mathbf{A}_{ij}^k - \mathbf{B}_{ij}^\ell \right|. \qquad (6)$$

Some failures might be caused by random noise in the original image from which the objects were extracted and consequently by inaccurate boundary detection. It is well-known that the centre of gravity T is sufficiently stable under noise but the maximum radius point R could be detected in a false position. To overcome this, noise removal as well as boundary smoothing are required prior to the shape matrix construction.

Table 2

Distances (in pixels) between the centres of gravity of the objects calculated in Step 7 of Algorithm MLC

	A	B	C	D	E	F	G
1	146	259	0	253	376	171	288
2	143	0.1	259	134	153	267	276
3	108	131	142	118	234	149	208
4	211	134	253	0.1	140	174	145
5	243	267	171	173	312	0.1	128
6	315	277	289	145	254	128	0

Table 3

Correspondence between two test sets of objects determined by Algorithm MLC

(a)	1	2	4	5	6
(b)	C	B	D	F	G

4. Numerical experiments

To measure the performance of the proposed object matching algorithm, the following experiments were carried out.

In the first experiment, algorithm MLC was applied to match simple shapes. Two test binary images of the size 512×512 pixels were formed by a graphic editor (see Fig. 1). The images differ by a translation, rotation (30°) and scale $(1 : 0.8)$. There are six objects in the first image and seven objects in the second one.

First, the shape matrices of the size 16×16 of all objects were found. The similarities between the objects were calculated for each possible pair and the
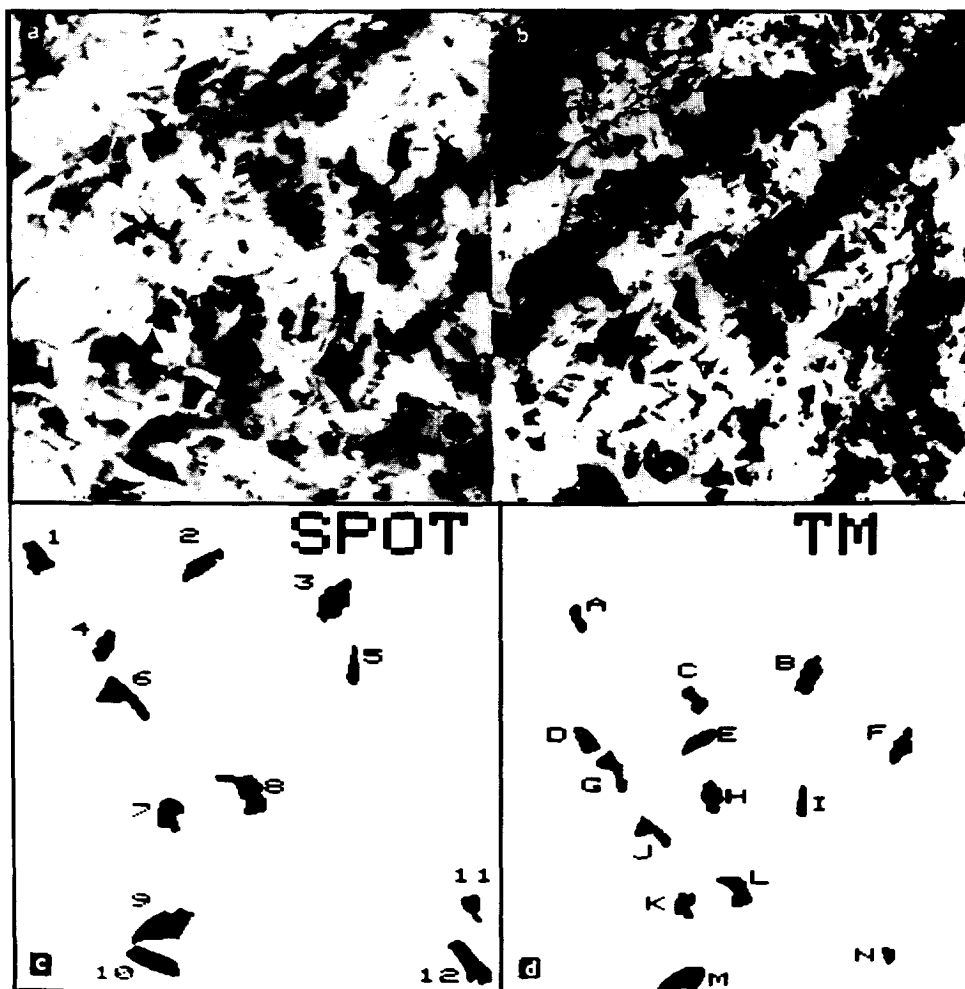
Fig. 2. (a) Reference image (SPOT subscene, 512 × 512 pixels, band 3). (b) Sensed image (Landsat TM subscene, 512 × 512 pixels, band 5). (c) Closed-boundary regions of the reference image. (d) Closed-boundary regions of the sensed image.

matrix **K** was constructed. Then the matching likelihood coefficients were computed (see Table 1). It is visible from Table 1 that two pairs of objects which correspond with each other with maximum reliability are $6 \approx G$ and $1 \approx C$.

The following values of the parametres of the geometric transformation between the images were obtained in step 6 of algorithm MLC: $a = -32.4$, $b = 237.1$, $c = 0.8$ and $\alpha = \pi/6$. The correspondence between the other objects was established in step 7 (see Table 2 and Table 3). Threshold values $r = 1$ and $t = 0.85$ were used in steps 7 and 8 of the algorithm.

It can be seen clearly by visual comparison of both

images that the final correspondence determined by algorithm MLC is correct everywhere.

An interesting result was obtained in the case of two multitemporal images of the same part of Czech territory, taken from satellites SPOT and Landsat. The first image was taken by SPOT HRV sensor in September 1987 (see Fig. 2a), the second one was taken by Landsat TM in August 1988 (see Fig. 2b). Subscenes of the size 512 × 512 pixels from the original images were used. Since the both sensors produce multispectral images, only one spectral band was selected for this experiment (band 3 from SPOT HRV and band 5 from Landsat TM images).

Table 4

Matrix **K** and the matching likelihood coefficients of the regions from satellite images (all values of **K** were multiplied by 100)

TM	SPOT												d
	1	2	3	4	5	6	7	8	9	10	11	12	
A	81	84	83	87	81	74	73	77	83	86	79	83	87
B	86	86	90	89	84	76	75	82	89	89	85	86	90
C	82	81	79	82	79	67	77	78	82	78	77	84	168
D	92	81	91	91	81	74	81	81	86	84	88	84	92
E	78	97	81	79	93	82	71	86	92	88	79	93	388
F	85	85	89	90	84	72	79	82	88	82	84	86	90
G	71	77	73	72	79	85	62	67	75	83	77	78	170
H	87	74	89	90	73	66	87	73	77	74	83	76	90
I	78	92	82	79	89	84	68	82	88	93	82	89	93
J	70	83	71	69	84	88	63	71	78	83	76	81	352
K	82	69	78	78	70	61	92	73	76	64	78	73	920
L	85	87	81	82	88	74	78	92	91	79	85	88	92
M	85	89	84	84	91	79	78	89	92	86	85	94	188
N	85	74	86	87	74	68	86	74	78	75	85	75	87

The geometric transformation between the images is affine, but we consider it as approximately similarity transform. In order to extract several closed-boundary regions, the following technique was used in the both images. (Note that our aim was not the complete image segmentation, but only detection and extraction of several homogeneous regions with high local contrast.)

First, each image was filtered by eight 3 × 3 Sobel masks to detect edges in various directions. The edge image was created as the maximum of those eight oriented-edge images. The edge image was binarized by low threshold ($h = 40$). After that, most of pixels were signed as "edge". In the binary image, closed boundary regions were found. Only the regions having perimeter between 10 and 100 pixels were taken into account.

Finally, the region boundaries were refined by contour-tracing technique in original images (not in the thresholded ones).

In this way, twelve regions from the SPOT image and fourteen regions from the Landsat TM image were extracted (see Figs. 2c and 2d). The extracted regions represent mostly fields and water bodies (lakes and ponds). However, there were some regions that appeared in one of the images only.

The matrix **K** and the matching likelihood coefficients are shown in Table 4. It is visible from Table 4 that two pairs of the regions which correspond with each other with maximum reliability are $7 \approx K$ and $2 \approx E$. In steps 7 and 8 of algorithm MLC, the following values of parametres were used: $r = 5$ and

Table 5

Correspondence between two sets of closed-boundary regions from satellite images determined by Algorithm MLC

SPOT	1	2	5	6	7	8	9	11
Landsat TM	D	E	I	J	K	L	M	N

$t = 0.85$. Table 5 summarizes the matching results. All regions were matched correctly which is really a hopeful result.

5. Summary and conclusion

In this paper, the problem of object matching has been addressed and a new matching algorithm has been presented. This algorithm uses the object description by a set of invariant features and combines it with relations in the image space. The performance of the proposed image matching algorithm has been demonstrated by experiments on the artificial test images and remotely sensed satellite images. All objects in the both experiments have been matched correctly.

Moreover, a new space- and scale-invariant description of planar shapes by coding into binary shape matrix has been developed and presented in this paper. The shape matrix has been successfully used as a shape descriptor in MLC algorithm.

The presented technique of image matching may have numerous practical applications in the field of pattern recognition, computer vision and remote sensing.

Acknowledgement

References

Flusser, J. (1992). Invariant shape description and measure of object similarity. *Proc. 4th Internat. Conf. Image Processing*, Maastricht, The Netherlands, 139–142.

Flusser, J. and T. Suk (1993). Pattern recognition by affine moment invariants. *Pattern Recognition* 26, 167–174.

Goshtasby, A. (1989). Description and discrimination of planar shapes using shape matrices. *IEEE Trans. Pattern Anal. Mach. Intell.* 7, 738–743.

Goshtasby, A. and G.C. Stockman (1985). Point pattern matching using convex hull edges. *IEEE Trans. Syst. Man. Cybernet.* 15, 631–637.

Haralick, R.M. and L.G. Shapiro (1985). Image segmentation techniques. *Computer Vision, Graphics, and Image Processing* 29, 100–132.

Hong, Z.Q. (1991). Algebraic feature extraction of image for recognition. *Pattern Recognition* 24, 211–219.

Hu, M.K. (1962). Visual pattern recognition by moment invariants. *IRE Trans. Inform. Theory* 8, 179–187.

Kiryati, N. (1988). Calculating geometric properties of objects represented by Fourier coefficients. *Proc. CVPR '88*, Ann Arbor, MI, 641–646.

Kuncheva, R.A. (1988). Pattern recognition method using two-dimensional Hadamard transform. *Proc. 7th Internat. Conf. on Robot Vision Sensory Controls*, Zurich, Switzerland, 231–236.

Lin, C.C. and R. Chellappa (1987). Classification of partial 2-D shapes using Fourier descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.* 9, 686–690.

Peli, T. (1981). An algorithm for recognition and localization of rotated and scaled objects. *Proc. IEEE* 69, 483–485.

Ranade, S. and A. Rosenfeld (1980). Point pattern matching by relaxation. *Pattern Recognition* 12, 269–275.

Reiss, T.H. (1993). *Recognizing planar objects using invariant image features*. Lecture Notes in Computer Science 676. Springer, Berlin.

Stockman, G.C., S. Kopstein and S. Benett (1982). Matching images to models for registration and object detection via clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* 4, 229–241.

Taza, A. and C.Y. Suen (1989). Discrimination of planar shapes using shape matrices. *IEEE Trans. Syst. Man Cybernet.* 19, 1281–1289.

Ventura, A.D., A. Rampini and R. Schettini (1990). Image registration by recognition of corresponding structures. *IEEE Trans. Geosci. Rem. Sensing* 28, 305–314.

Weiss, I. (1988). Projective invariants of shapes. *Proc. DARPA Image Understanding Workshop*, Cambridge, MA, 1125–1134.

Zahn, C.T. (1974). An algorithm for noisy template matching. *Proc. IFIP Congress*, Stockholm, Sweden, 698–701.