An Experimental Comparison of Seven Shape Descriptors in the General Shape Analysis Problem

Dariusz Frejlichowski

West Pomeranian University of Technology, Szczecin,
Faculty of Computer Science and Information Technology,
Zolnierska 49, 71-210, Szczecin, Poland
dfrejlichowski@wi.zut.edu.pl

Abstract. The general shape analysis is a problem similar to the recognition or retrieval of shapes. The most important difference is that the processed object does not have to belong to a base class, but usually is only similar to the template representing the class. The most general information about a shape is here concluded, i.e. how round, elliptical, triangular, etc. it is. Such a problem can occur in applications with few general base classes, e.g. in pre-classification or the assignment of stamps extracted from an image to few classes in order to find the fraudulent stamp images (mainly governmental, official ones). In the paper seven shape descriptors were explored using the template matching approach. In order to select the best approach their performance was compared with results provided by almost two hundred humans and collected using appropriate inquiry forms.

1 Introduction

The idea of image recognition can be realized in many various ways. One of them is based on identification of objects placed in a digital image. In that case one can use some features that are supposed to appropriately represent an object. Usually color, texture and shape are taken for that purpose. In many cases the last one is especially useful (however, lately the idea of combining completely different features becomes more tempting and popular, [1]). The recognition of shapes can be realized through the *template matching* approach. Roughly speaking, in this method an object under identification is matched with the base objects (*templates*). Obviously the matching of objects itself is insufficient, since in real situations they tend to be strongly deformed. In case of shapes not only the affine transformations (e.g. rotation, scaling and translation), but also noise and occlusion have to be considered ([2]). Therefore the so-called *shape descriptors* are used in order to represent a shape invariantly to particular deformations.

However, there is a class of applications, where the problem can not be considered as the traditional recognition. In that case the processed shape does not have to belong to a base class and usually it does not. The problem is depicted

A. Campilho and M. Kamel (Eds.): ICIAR 2010, Part I, LNCS 6111, pp. 294–305, 2010.

in Fig. 1. An object can be considered as similar to the three templates, yet it definitely is not one of them. It can be assigned to one of the very general shape classes, e.g. circle, triangle, star. It means that we are not interested in the exact identification of an object, but we are trying to find one or several general shapes, which are the most similar to the object being processed.

Fig. 1. Illustration of the problem — which general shape is the most similar?

The problem of general shape analysis can be utilized in various applications. Here, three examples will be briefly described. The first one is searching for probable false documents stored on a hard drive in a digital form ([3]). In this problem we can identify a general template — type of seal (e.g. official, public, business, institutional) instead of performing the exact process of recognition. This is based on the assumption that particular types of seal have an expected shape, e.g. official ones are round, whereas medical ones are rectangular (in Poland). The second example is the process of initial classification when working with large databases. In order to speed up the whole process the object is firstly matched with small number of general classes. Later it can be recognized more precisely within the preliminarily selected general class. In fact this process can be performed several times, with the classes becoming more detailed at each subsequent iteration. The third example is the possibility of using voice commands (e.g. 'find round red objects') for shape retrieval in large multimedia databases.

In the paper seven shape descriptors are used to indicate the general shapes. Their selection is not random. Firstly two so-called *simple shape descriptors* were taken, namely *Roundness* ([4]) and one of the *Feret measures* ([5]) — the *X/Y ratio*. Also, five more sophisticated shape descriptors were applied: *Moment Invariants* ([6]), *2D Fourier Descriptors* ([7]), *Point Distance Histogram* ([8]), *UNL* ([9]) and *UNL-Fourier* ([10]). Two of them (*PDH* and *UNL-F*) are invariant to rotation, the other three — are not. The *2D FD* is known for its ability of generalization ([7]). On the other hand, the *PDH*, thanks to the combination of polar coordinates with histogram, can emphasize the small differences between objects. Finally, the *UNL* and *UNL-F* have been successfully used in shape recognition, and are invariant (especially *UNL-F*) to many shape deformations.

The shape under analysis can be represented in two different ways. The first one is the outline of an object and the second is the whole region covered by it ([11]). Each of the algorithms works with one of the mentioned representations. In the paper the contour was explored using *PDH*, *UNL*, *UNL-F*, *Roundness* and the *X/Y Feret ratio*. The region was a subject for *MI* and *2D FD*.

The rest of the paper is organized as follows. The second section describes precisely the shape descriptors used in the problem of general shape analysis.

The third one provides experimental results achieved using them. The fourth section presents the results provided by humans and the comparison between them and the artificial algorithms. The last section concludes the paper and provides some ideas for future work with the problem.

2 The Description of Selected Algorithms

Seven various shape descriptors were taken for the experiments with the general shape analysis. The methods were selected deliberately in order to consider various ways of performing the analysis by humans. Some people take into account only the simplest features, e.g. curvature; others analyze an object on a higher level. However, the most important difference is the invariance to rotation. Some people during general analysis of a shape assume that it can not be rotated and that strongly influences the results.

In this section all the algorithms used will be described in detail. The order of their presentation is based on the ascending level of difficulty and sophistication.

2.1 Simple Shape Descriptors

Two simple shape descriptors were chosen. Those fast methods for measuring a shape are less popular nowadays thanks to the increasing computational power of computers. However, the generality of the simple methods can be an advantage in the presented problem. The first one, *X/Y Feret shape measure*, can be computed using the formula:

$$F_{xy} = \frac{x_{max} - x_{min}}{y_{max} - y_{min}}, \tag{1}$$

where:
x_{min}, x_{max} — minimal and maximal horizontal coordinates of a contour shape,
y_{min}, y_{max} — minimal and maximal vertical coordinates of a contour shape.

The *Roundness* (a measure of the sharpness of a shape) was a second simple descriptor used in the experiment. This measure is based on two other shape features: the area (A) and the perimeter (P), and can be formulated as ([4]):

$$R = \frac{4\pi A}{P^2}. \tag{2}$$

2.2 Moment Invariants

The *moments* are commonly used in image representation. Hu introduced the name of *Moment Invariants* in 1962 ([12]). For shapes this representation uses only two values of image function $f(x, y)$ — 1 for a pixel belonging to an object and 0 for background pixels — instead of 256 levels as for a grayscale image. The *general geometrical moments* are given by the formula ([6]):

$$M_{pq} = \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy, \tag{3}$$

where: $p, q = 0, 1, ..., \infty$.

For the discrete values in the image the above can be written as ([13]):

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y). \tag{4}$$

We calculate the *centre of gravity* of an object ([13]):

$$x_c = \frac{m_{10}}{m_{00}} \qquad y_c = \frac{m_{01}}{m_{00}}. \tag{5}$$

Then we calculate the *central moments* ([13]):

$$\mu_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q f(x, y). \tag{6}$$

The next step is the calculation of *normalized central moments* ([13]):

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q+2}{2}}}. \tag{7}$$

Finally we can derive the *Moment Invariants* (*MI*). In practice usually the first seven MI are used ([14]):

$$\begin{aligned}
\varphi_1 &= \eta_{20} + \eta_{02} \\
\varphi_2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \\
\varphi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\varphi_4 &= (\eta_{30} - \eta_{12})^2 + (\eta_{21} - \eta_{03})^2 \\
\varphi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{03} + \eta_{21})[3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2] \\
\varphi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{03} + \eta_{21}) \\
\varphi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{03} + \eta_{21})^2] \\
&\quad - (\eta_{30} - 3\eta_{12})(\eta_{03} + \eta_{21})[3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2].
\end{aligned} \tag{8}$$

The received shape representation is very compact — it is constituted by a vector of only seven values.

2.3 Fourier Descriptors

The *Fourier transform* is widely used in pattern recognition. In case of shapes usually its one-dimensional version is applied to the contour representation (e.g. [15]). However, in the literature another approach is also present. The so-called *2D Fourier Descriptors* are applied to a region (e.g. [16]).

In the experiments the *2D FD* were utilized. They can be derived using the *2D Fourier transform*, where only the absolute spectrum is used ([7]):

$$C(k, l) = \frac{1}{HW} \left| \sum_{h=1}^{H} \sum_{w=1}^{W} P(h, w) \cdot e^{(-i\frac{2\pi}{H}(k-1)(h-1))} \cdot e^{(-i\frac{2\pi}{W}(l-1)(w-1))} \right|, \tag{9}$$

where:
H, W — height and width of the image in pixels,
k — sampling rate in vertical direction ($k \geq 1$ and $k \leq H$),
l — sampling rate in horizontal direction ($l \geq 1$ and $l \leq W$),
$C(k,l)$ — value of the coefficient of *discrete Fourier transform* in the coefficient matrix in k row and l column,
$P(h,w)$ — value in the image plane with coordinates h, w.

2.4 UNL and UNL-Fourier Shape Descriptors

The *UNL* (*Universidade Nova de Lisboa*) descriptor is based on the transform of the same name ([9]). It uses complex representation of Cartesian coordinates for points and parametric curves in discrete manner ([9]):

$$z(t) = (x_1 + t(x_2 - x_1)) + j(y_1 + t(y_2 - y_1)), \qquad t \in (0,1), \tag{10}$$

where $z_1 = x_1 + jy_1$ and $z_2 = x_2 + jy_2$ (complex numbers) and z_i denotes point with coordinates x_i, y_i. The *centroid* O is now calculated ([9]):

$$O = (O_x, O_y) = (\frac{1}{n}\sum_{i=1}^{n} x_i, \frac{1}{n}\sum_{i=1}^{n} y_i), \tag{11}$$

and the maximal Euclidean distance between points and centroid is found ([10]):

$$M = \max_i \{\|z_i(t) - O\|\}, \qquad \forall i = 1...n, \qquad t \in (0,1). \tag{12}$$

Now coordinates are transformed ([9]):

$$U(z(t)) = R(t) + j \times \theta(t) = \frac{\|z(t) - O\|}{M} + j \times atan(\frac{y(t) - O_y}{x(t) - O_x}). \tag{13}$$

The discrete version is formulated as follows ([9]):

$$\begin{aligned} U(z(t)) = & \frac{\|(x_1 + t(x_2 - x_1) - O_x) + j(y_1 + t(y_2 - y_1) - O_y)\|}{M} \\ & + j \times atan(\frac{y_1 + t(y_2 - y_1) - O_y}{x_1 + t(x_2 - x_1) - O_x}). \end{aligned} \tag{14}$$

The parameter i is discretized in the interval [0,1] with significantly small steps ([10]). Derived coordinates are put into a matrix, in which the row corresponds to the distance from centroid, and the column — to the angle. The obtained matrix is 128×128 pixels size.

Since after the *UNL-transform* we obtain 2-dimensional binary image again, the author of the approach proposed using the *2D Fourier transform* as the next step ([9]). That gave one of the best descriptors in shape recognition called shortly the *UNL-F* and achieved using the *UNL-Fourier transform*.

2.5 Point Distance Histogram

The *Point Distance Histogram* ([8]) is an algorithm combining the advantages of histogram with the transformation of contour points into polar coordinates. Firstly the mentioned coordinates are derived (with $O = (O_x, O_y)$ as the origin of the transform) and put into two vectors Θ^i for angles and P^i for radii ([8]):

$$\rho_i = \sqrt{(x_i - O_x)^2 + (y_i - O_y)^2}, \qquad \theta_i = atan\left(\frac{y_i - O_y}{x_i - O_x}\right). \tag{15}$$

The resultant values are converted into nearest integers ([8]):

$$\theta_i = \begin{cases} \lfloor \theta_i \rfloor, & if\ \theta_i - \lfloor \theta_i \rfloor < 0.5 \\ \lceil \theta_i \rceil, & if\ \theta_i - \lfloor \theta_i \rfloor \geq 0.5 \end{cases}. \tag{16}$$

The next step is the rearrangement of the elements in Θ^i and P^i according to increasing values in Θ^i. This way we achieve the vectors Θ^j, P^j. For equal elements in Θ^j only the one with the highest corresponding value P^j is selected. That gives a vector with at most 360 elements, one for each integer angle. For further work only the vector of radii is taken — P^k, where $k = 1, 2, ..., m$ and m is the number of elements in P^k ($m \leq 360$). Now, the normalization of elements in vector P^k is performed ([8]):

$$M = \max_k \{\rho_k\}, \qquad \rho_k = \frac{\rho_k}{M}, \tag{17}$$

The elements in P^k are assigned to r bins in histogram (ρ_k to l_k,[8]):

$$l_k = \begin{cases} r, & if\ \rho_k = 1 \\ \lfloor r\rho_k \rfloor, & if\ \rho_k \neq 1 \end{cases}. \tag{18}$$

3 Conditions and Results of the Experiments

The problem was explored using objects from [17]. The database included 10 templates (the general shapes) and 40 tested objects (see Fig. 2).

For each explored algorithm the idea of a test was simple. A test object was represented using a shape descriptor, and so were all the general shapes (templates). Basing on the typical template matching approach, the description of a test shape was matched using *Euclidean distance* with all the descriptions of the templates. The three smallest dissimilarity values indicated the general shapes closest to a test shape according to the algorithm explored. Pictorial representations of the results will be provided in consecutive figures.

The results for *Roundness* presented in Fig. 3 can not be considered ideal. However, in some cases they seem to be correct. For example, the rhombus (object no. 2) is firstly connected to the square and secondly to the rectangle, which is a very good result. However, the third indicated general shape — the disc — completely does not fit to the previous ones. Similarly, the triangle was

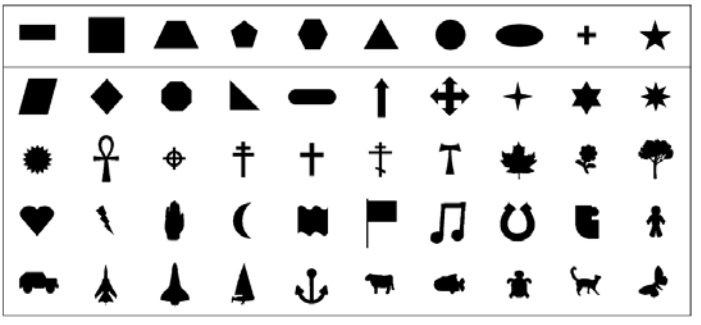

Fig. 2. The division of shapes into 10 templates and 40 test objects

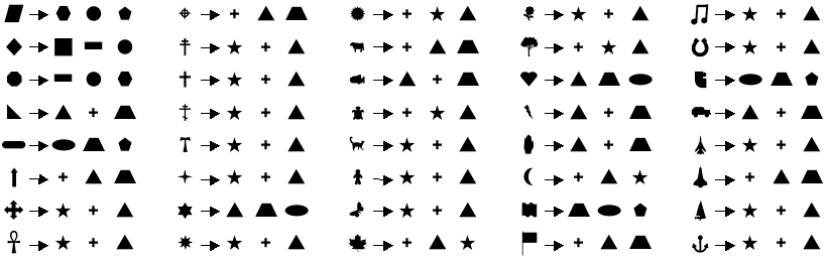

Fig. 3. Results of the experiment on general shape analysis using Roundness

properly matched with a general triangular shape and a rectangle with rounded corners was connected with the ellipse. In other cases the star was selected very often. That is clearly seen for various versions of crosses. The results for more complicated objects are less convincing. Usually the star, triangle or cross were indicated.

As one might expect, the results achieved using such a simple approach (X/Y *Feret*) are not satisfactory (see Fig. 4). Only in a few cases the results can be considered correct. The rounded rectangle is the first example. Definitely the first two selected general shapes (the square and the ellipse) are similar to the test object. Similarly, the results achieved for the car are correct in all three cases — the ellipse, the trapezoid and the rectangle. In few other cases the first result is also acceptable. That concerns for example the heart that is similar to the triangle and the flag close to the square.

The results provided using *Moment Invariants* (see Fig. 5) are usually very promising for the first indicated general shape. This time the rectangle 'became more popular'. In fact, e.g. for the crosses, the human, the cat, the car, etc. one can agree that when rotated they are very similar to the rectangle.

The results for *Fourier Descriptors* (see Fig. 6) are very interesting. For the first five simple test shapes they are very proper. The results achieved for group

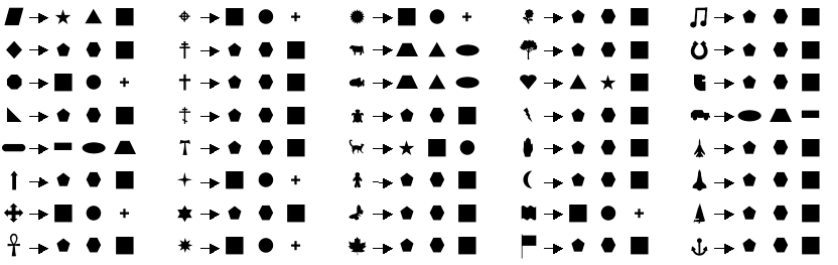

Fig. 4. Results of the experiment on general shape analysis using X/Y Feret ratio

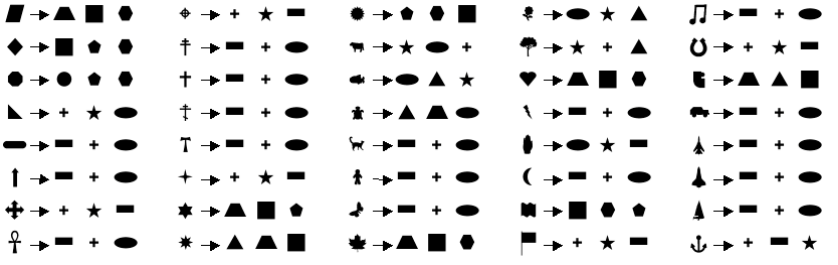

Fig. 5. Results of the experiment on general shape analysis using Moment Invariants

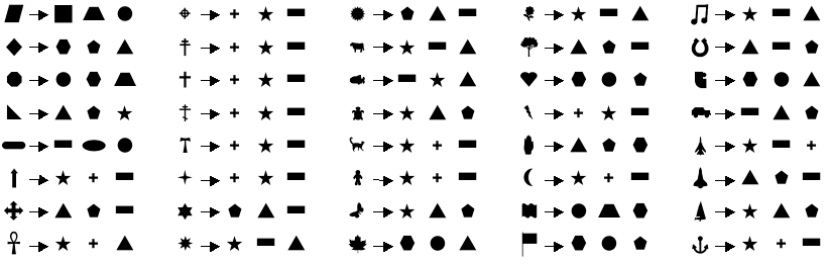

Fig. 6. Results of the experiment on general shape analysis using Fourier Descriptors

of crosses is also correct. For other objects the indication of the star as the most similar general shape is clearly visible yet not always correct.

The *UNL* shape descriptor (Fig. 7) completely failed in our problem. Only few results can be considered proper. Usually, the square and the disc were taken at the first place. Sometimes the pentagon and the hexagon were selected. The lack of other templates in results, e.g. the star and the cross, is noticeable.

The results achieved using the *UNL-Fourier* descriptor (Fig. 8) are not convincing. This method is very effective in the traditional shape recognition problem. However in the general shape analysis it often fails. For example, the results achieved for the parallelogram and crosses are irrational. On the other hand, the

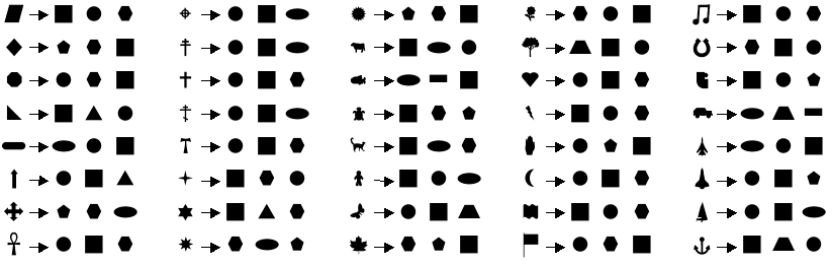

Fig. 7. Results of the experiment on general shape analysis using UNL descriptor

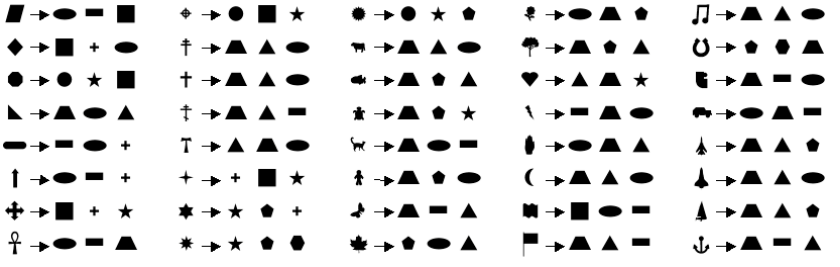

Fig. 8. Results of the experiment on general shape analysis using UNL-F descriptor

performance of the descriptor for the rounded rectangle, stars and some more complicated test objects is definitely acceptable.

In some cases the results of the *PDH* approach (Fig. 9) are acceptable. For example, the first indicated general shape is correct for the first four test objects as well as for the human, the hand and the car. In general this approach gave slightly worse results than expected. However, in comparison with other explored descriptors it seems to be appropriate for the problem under consideration.

4 The Results Provided by Humans

The results provided and briefly described in the previous section can not be adequately judged, since we can only guess if a method is working well or not. There is no independent measure for the problem of general shape analysis. Therefore, in order to estimate the behavior of explored algorithms an inquiry form was filled by 187 persons (124 men, 63 women, aged from 9 to 62) in. This inquiry was conducted in order to investigate the manner in which humans perform the task of general shape analysis. This can serve as a benchmark, an ideal result. Now, we only have to investigate which of the explored artificial algorithms is the most similar to it and in what degree. The results of general shape analysis performed by humans are depicted in Fig. 10. The analysis of the inquiry forms could be performed in various ways. Here, the most popular result at the particular place was selected. That gave the most common general shapes indicated by humans for particular test objects.

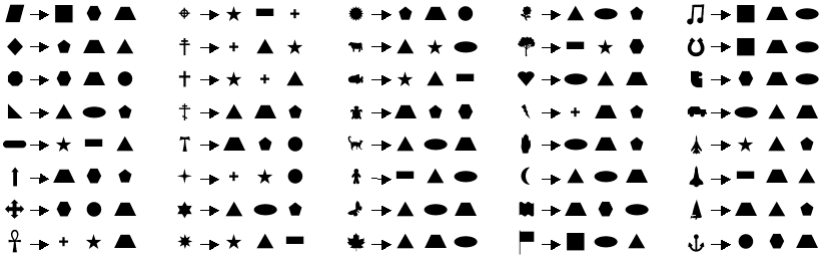

Fig. 9. Results of the experiment on general shape analysis using PDH descriptor

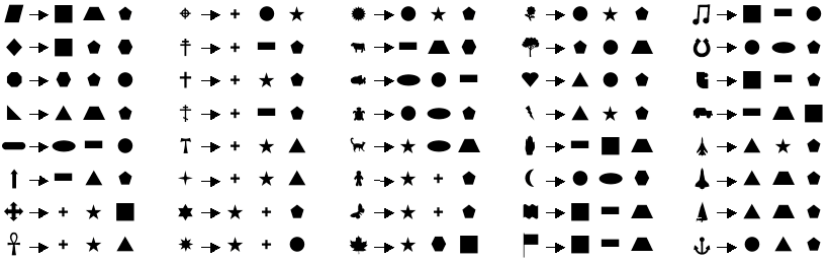

Fig. 10. Results of the general shape analysis test performed by humans, a benchmark for the artificial methods

Table 1. The comparison of the general shape analysis performed by humans and artificial algorithms - the percentage of convergence between a shape descriptor and benchmark human results

Shape descriptor	1st indication	2nd indication	3rd indication
1. FD	35%	22.5%	17.5%
2. PDH	25%	15%	27.5%
3. Roundness	25%	12.5%	17.5%
4. MI	20%	17.5%	5%
5. UNL-F	17.5%	10%	12.5%
6. UNL	15%	5%	12.5%
7. X/Y Feret	10%	7.5%	2.5%

The degree of similarity between shape descriptors and results provided by humans is presented in Table 1. Each time the percentage of proper indications compared to benchmark human results is presented, separately for the three firstly selected templates. As it can bee seen, the *Fourier Descriptors* work most similarly to the benchmark human statistics. The result is much higher than for other six explored methods. That concerns the first (35%) and the second (22.5%) selected template. Only for the third selection another descriptor works better, namely — *PDH*. It is accordant with the human results in 27.5 %.

5 Conclusions and Future Plans

The paper described experimental results on usage of shape descriptors in the general shape analysis. The problem is similar to recognition or retrieval, but here we assume that the processed shape does not have to belong to any of the template classes, which include the few most general shapes — triangle, square, disc, etc. Therefore, it can be considered as the determination how triangular, square, round, etc. is a tested shape. The general shape analysis, as presented in the paper, can be useful in many applications. The first one is the preliminary classification of shapes, when we firstly assign an object to a major class and subsequently we increase the level of details in the identification. Another example is the shape retrieval based on the similarity of an object to few the most general shapes. It can be combined for example with the usage of voice commands. The third example is the analysis of seals when searching for probable false documents stored on a hard drive in a digital form.

During the experiments on the problem seven shape descriptors were explored. In order to measure their performance, a special inquiry form was developed. It was similar to the performed tests and it was filled in by almost two hundred persons. The selection of the best method was based on a very simple criterion. The artificial method with the results most similar to the ones provided by humans was treated as the best. As it turned out the *FD* were the best among the tested approaches. The second place went to the *PDH*, which was also rather successful. However, the achieved numerical results can not be treated as satisfactory enough. On the other hand, the results provided by humans are in many cases ambiguous as well. Nevertheless, the results achieved by the best shape descriptor among tested are worse than expected (35%), therefore there is still necessity of exploring some other algorithms in the problem. This is the first conclusion related to the future work. The second important issue is the a different way of constructing the benchmark. In the paper the simplest approach was utilized. Plainly, the most popular result in the inquiry forms was treated as the proper one. However, in some cases the differences between the most popular indication and the second one were very small. This can be taken into consideration in the future improved method of comparing the artificial results with the human benchmark. Finally, the experiments on some practical examples will be performed to illustrate the capabilities of the best approaches. The first problem to explore is the identification of document seals by means of the general shape analysis.

References

1. Forczmanski, P., Frejlichowski, D.: Strategies of Shape and Color Fusions for Content Based Image Retrieval. Advances in Soft Computing 45, 3–10 (2007)
2. Frejlichowski, D.: An Algorithm for Binary Contour Objects Representation and Recognition. In: Campilho, A.C., Kamel, M.S. (eds.) ICIAR 2008. LNCS, vol. 5112, pp. 537–546. Springer, Heidelberg (2008)

3. Forczmanski, P.: Stamp Detection in Scanned Documents. In: 8th Int. Conf. on Comp. Science - Research and Applications (IBIZA 2008), Poland, Kazimierz Dolny (2009)
4. Nafe, R., Schlote, W.: Methods for Shape Analysis of two-dimensional closed Contours — A biologically important, but widely neglected Field in Histopathology. Electronic Journal of Pathology and Histology 8(2) (2002)
5. Whang, S.S., Kim, K., Hess, W.M.: Variation of silica bodies in leaf epidermal long cells within and among seventeen species of Oryza (Poaceae). American Journal of Botany 85, 461–466 (1998)
6. Rothe, I., Süsse, H., Voss, K.: The method of normalization to determine invariants. IEEE Trans. on Pattern Anal. and Mach. Int. 18, 366–375 (1996)
7. Kukharev, G.: Digital Image Processing and Analysis. SUT Press (1998) (in Polish)
8. Frejlichowski, D.: The Point Distance Histogram for Analysis of Erythrocyte Shapes. Polish Journal of Environmental Studies 16(5b), 261–264 (2007)
9. Rauber, T.W., Steiger-Garcao, A.S.: 2-D form descriptors based on a normalized parametric polar transform (UNL transform). In: Proc. of MVA 1992 IAPR Workshop on Machine Vision Applications (1992)
10. Rauber, T.W.: Two-dimensional shape description. Technical Report: GR UNINOVA-RT-10-94, Universidade Nova de Lisboa (1994)
11. Zhang, D., Lu, G.: Review of shape representation and description techniques. Pattern Recognition 37(1), 1–19 (2004)
12. Flusser, J., Suk, T.: Affine moment invariants: A new tool for character recognition. Pattern Recognition Letters 15(4), 433–436 (1994)
13. Hupkens, T.M., de Clippeleir, J.: Noise and intensity invariant moments. Pattern Recognition Letters 16(4), 371–376 (1995)
14. Liu, C.-B., Ahuja, N.: Vision Based Fire Detection. In: Proc. of 17th Int. Conf. on Pattern Recognition (ICPR 2004), UK, Cambridge (2004)
15. Osowski, S., Nghia, D.: Fourier and wavelet descriptors for shape recognition using neural network — a comparative study. Pattern Recognition 35(9), 1949–1957 (2002)
16. Yadav, R.B., Nishchal, N.K., Gupta, A.K., Rastogi, V.K.: Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: A comparative study. Optics and Lasers in Engineering 45, 695–708 (2007)
17. Frejlichowski, D.: General Shape Analysis Using Fourier Shape Descriptors. In: Swiatek, J., et al. (eds.) Information Systems Architecture and Technology – System Analysis in Decision Aided Problems, pp. 143–154 (2009)