

Affine Invariant Gradient Based Shape Descriptor

Abdulkerim Çapar, Binnur Kurt, and Muhittin Gökmen

Istanbul Technical University, Computer Engineering Department
34469 Ayazağa, Istanbul, Turkey
{capar, kurtbin, gokmen}@itu.edu.tr

Abstract. This paper presents an affine invariant shape descriptor which could be applied to both binary and gray-level images. The proposed algorithm uses gradient based features which are extracted along the object boundaries. We use two-dimensional steerable G-Filters [1] to obtain gradient information at different orientations. We aggregate the gradients into a shape signature. The signatures derived from rotated objects are shifted versions of the signatures derived from the original object. The shape descriptor is defined as the Fourier transform of the signature. We also provide a distance definition for the proposed descriptor taking shifted property of the signature into account. The performance of the proposed descriptor is evaluated over a database containing license plate characters. The experiments show that the devised method outperforms other well-known Fourier-based shape descriptors such as centroid distance and boundary curvature.

1 Introduction

Shape representation and description plays an important role in many areas of computer vision and pattern recognition. Neuromorphometry, character recognition, contour matching for medical imaging 3-D reconstruction, industrial inspection and many other visual tasks can be achieved by shape recognition [2].

There are two recent tutorials on the shape description and matching techniques. Veltkamp and Hagedoorn [3] investigated the shape matching methods in four parts: global image transformations, global object methods, voting schemes and computational geometry. They also worked on shape dissimilarity measures. Another review on shape representation methods is accomplished by Zhang and Lu [4]. They classified the problem into two class of methods; contour-based methods and region-based methods.

In this work, we proposed a contour-based shape description scheme using some rotated filter responses along the object boundary. Although, we extract the descriptors by tracing object boundary, we utilize local image gradient information also. The rotated G-filter kernels, which are a kind of steerable filters, are employed to obtain the local image gradient data. Steerable filters are rotated match filters to detect some local features in images [5]. Local descriptors are increasingly used for task of image recognition because of their perceived robustness with respect to occlusions and to global geometrical deformations [6,7,8,9].

In this study we interested in the filter responses only along object boundaries. These responses are treated as a one-dimensional feature signature. Fourier Descriptors of this feature signature are computed to provide starting point invariance and to have compact descriptors. Moreover, Fourier Descriptor (FD) is one of the most widely used shape descriptors due to its simple computation, clarity and coarse to fine description capability. Zhang and Lu [10] compared the image retrieval performance of FD with curvature scale space descriptor (CSSD) which is an accepted MPEG-7 boundary-based shape descriptor. Their experimental results show that FD outperforms CSSD in terms of robustness, low computation, hierarchical representation, retrieval performance, and suitability for efficient indexing. One-dimensional image signatures are employed to calculate FDs. Different image signatures are reported in the literature based on “color”, “texture” or “shape”. When we researched on the “shape” signatures, we face with several methods ([11,12,13,14]) based on topology of the object boundary contours. Fortunately, a general evaluation and comparison on these FD methods had not been accomplished until recent study of Zhang and Lu [15]. Zhang and Lu studied on different shape signatures and Fourier transform methods in means of image retrieval. They have the following conclusions, which are important for us; on retrieval performance, centroid distance and area function signatures are most suitable methods, 10 FDs are sufficient for a generic shape retrieval system.

In section 2 the directional gradient extraction is introduced. Section 3 is for the proposed gradient based shape descriptor. The experimental results and the conclusion are presented in section 4 and 5 respectively.

2 Directional Gradient Extraction Using Steerable Filters

In this study we deal with boundary based shape descriptors and assume that objects always have closed boundaries. Many shape descriptors exist in the literature and most of these descriptors are not able to address different type of shape variations in nature such as rotation, scale, skew, stretch, and noise. We propose an affine invariant shape descriptor in this study, which handles rotation, scale and skew transformations.

Basically the proposed descriptor uses gradient information at the boundaries rather than the boundary locations. We use 2D Generalized Edge Detector [1] to obtain object boundary. We then trace the detected boundary pixels along the clock-wise direction to attain the locations of the neighboring boundary pixels denoted as (x_i, y_i) . So the object boundary forms a matrix of size n -by-2

$$\Gamma = \begin{bmatrix} x & y \end{bmatrix}, x = [x_1, \dots, x_n]^T, y = [y_1, \dots, y_n]^T \quad (2.1)$$

where n is the length of the contour such that $n = |\Gamma|$. We are interested in the directed gradients at these boundary locations. We utilize steerable G-Filters to obtain the gradient at certain directions and scales as

$$D_{(\lambda, \tau)}^\theta(I, x_i, y_i) = (I * G_{(\lambda, \tau)}^\theta)(x_i, y_i) \quad (2.2)$$

where I is the image intensity. Steerable $G_{(\lambda, \tau)}^\theta$ filter is defined in term of $G_{(\lambda, \tau)}^{\theta=0}$ as

$$G_{(\lambda,\tau)}^\theta(x', y') = G_{(\lambda,\tau)}^{\theta=0}(x, y) \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2.3)$$

Detailed analysis of these filters ($G_{(\lambda,\tau)}^{\theta=0}$) is given in [1]. Let us denote the response matrix as $F(\Gamma) = [f_{k,m}]$ where $f_{k,m}$ is equal to $D_{(\lambda,\tau)}^{\theta_m}(I, x_k, y_k) = I * G_{(\lambda,\tau)}^{\theta_m}(x_k, y_k)$. Let us assume that we use M number of steerable filters whose directions are multiple of $\frac{\pi}{M}$ such that $\theta_m = m \frac{\pi}{M}$. In this case the size of F is $M|\Gamma|$. When the object is rotated α degrees about the center of gravity, the columns of the matrix F is circular shifted to left or right. The relationship between the rotation angle and amount of shifting can be stated as follows assuming that the rotation angle is multiple of $\frac{\pi}{M}$

$$F(\Gamma_\alpha) = [f'_{k,m}], \quad \alpha = s \frac{\pi}{M} \quad (2.4)$$

$$f'_{k,(m+s) \bmod M} = f_{k,m}$$

where $\Gamma_\alpha = R(\alpha)\Gamma$ and $R(\alpha)$ is the rotation matrix.

In order to analyze the filter response along different directions, we first apply the filter at each pixel on the sample object boundary (Fig. 1.2(a)) where we change the angle from 0^0 to 180^0 with 1^0 increments and obtain the response matrix F . The response plot is given in Fig. 1.2(b). The experimental results also verify the circular shifting property. Fig. 1.2(c) shows how the steerable filter response changes with the rotation angle.

We also analyze how the shifting property is affected when the rotation angle is not multiple of $\frac{\pi}{M}$. Let us assume that rotation angle is $s \frac{\pi}{M} + \phi$ where $0 \leq \phi < \frac{\pi}{M}$ and we have M steerable filters $\{G^{\theta_k}; k = 0, 1, \dots, M-1\}$. Any steerable filter response at one direction can be expressed as a linear combination of M different G^{θ_k} filter responses as

$$I * G^\phi = \sum_{k=0}^{M-1} c_k(\phi) (I * G^{\theta_k}) \quad (2.5)$$

The difference between the signatures obtained for the same object, one is rotated $s \frac{\pi}{M}$ degrees and the other is rotated $s \frac{\pi}{M} + \phi$ degrees can be found as

$$\begin{aligned}
f_{k,m} &= R_\alpha(I) * G^{\theta_m} \\
\hat{f}_{k,m} &= R_{\alpha+\phi}(I) * G^{\theta_m} = R_\alpha(I) * G^{\theta_m+\phi} \\
|f_{k,m} - \hat{f}_{k,m}| &= R_\alpha(I) * (G^{\theta_m} - G^{\theta_m+\phi})
\end{aligned} \tag{2.6}$$

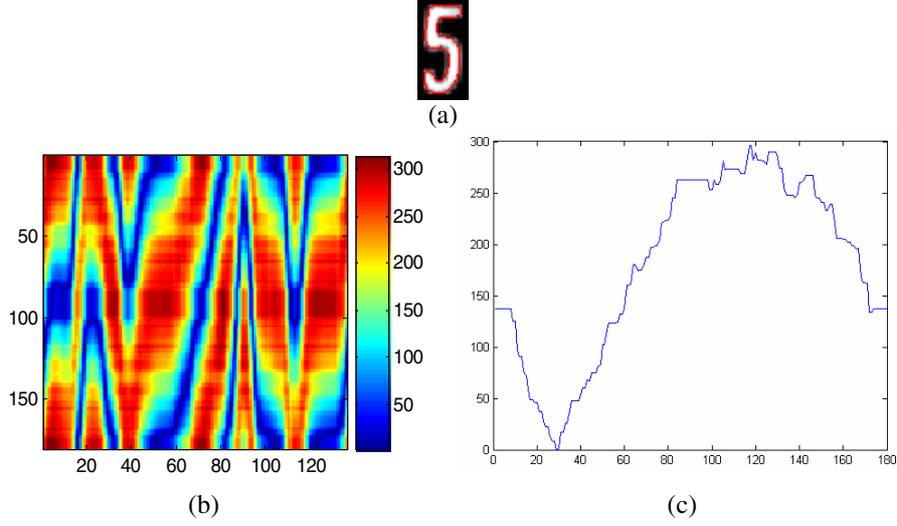


Fig. 1. (a) Sample digit, (b) Filter responses at each direction and boundary pixels, (c) Steerable filter response with respect to rotation angle

The Eq. 2.5 can be used to compute the difference between ϕ degrees rotated filter kernels such that

$$|G^{\theta_m} - G^{\theta_m+\phi}| = \left(\sum_{k=0}^{M-1} |\cos(\theta_m - \theta_k) - \cos(\theta_m + \phi - \theta_k)| \right) G^{\theta_m} \tag{2.7}$$

This formulation helps us to decide how many directions we should use, hence the size of the descriptor. The equation suggests that small ϕ rotations are tolerated for small M values and even large rotations are tolerated for large M values. We explored this phenomenon in our experiments. The experimental results are reported and discussed in Section 4.

Next section will discuss how to use Fourier Descriptor applied to the gradient signature discussed above and the distance metric between two descriptors.

3 Affine Invariant Shape Descriptor

Fourier Descriptors (FDs) are mostly employed for boundary shape description. Zhang and Lu [15] compared shape retrieval using FDs derived from different shape signatures in terms of computation complexity, robustness, convergence speed, and

retrieval performance. They reported that the centroid distance shape signature outperforms other signature methods in terms of above criteria. Selecting the shape signature is the most critical step for FDs. Various signature models were proposed in the literature such as complex boundary coordinates, centroid distance and boundary curvatures [15].

We apply Fourier transform to the gradient based shape signature described in the previous section in order to obtain a compact descriptor. We denote the descriptor as

$\tilde{F} = \left[\left\| \tilde{f}_{k,m} \right\| \right]$ where the coefficients are computed as follows

$$\tilde{f}_{k,m} = \frac{1}{N} \sum_{t=0}^{N-1} f_{t,m} \exp\left(\frac{-j2\pi kt}{N}\right) \quad (3.1)$$

Note that we take the magnitude of the Fourier coefficients to keep the descriptor invariant to where you start tracing on the boundary. Note also that \tilde{F} also satisfies circular shifting property when we rotate the object:

$$\tilde{F}(\Gamma_\alpha) = \left[\tilde{f}'_{k,m} \right], \alpha = s \frac{\pi}{M}, \tilde{f}'_{k,(m+s) \bmod M} = \tilde{f}_{k,m} \quad (3.2)$$

Scale invariance is achieved by dividing the magnitudes by the DC component, i.e., $\tilde{f}_{0,m}$.

Finally we present a distance metric to compare two descriptors. Assume that we are given two descriptors such as $\tilde{f}_{k,m}$ and $\tilde{g}_{k,m}$ then we define the distance as

$$SD(\tilde{f}, \tilde{g}) = \min_{0 < r < M-1} \left\| \tilde{f}_{k,(m+r) \bmod M} - \tilde{g}_{k,m} \right\|_2 \quad (3.3)$$

The computational complexity of Eq. 3.3 is $M(M \times L)$ where M is the number of filter kernels used and L is the number of Fourier coefficients used.

4 Experimental Results

We have run our algorithm on a license plate characters. Fig. 2 shows typical segmented license plate characters selected from the database. The database contains 8321 gray level digit characters segmented from the real traffic license plates. About the half of the database characters, namely 4121 digits, is used in training, and the remaining 4200 digits are used in testing. In practice, although there are 10 digits appear on the license plate images, we only labeled 9 classes because one can not distinguish between 6 and 9 if rotation is considered and assuming that no prior information is available.

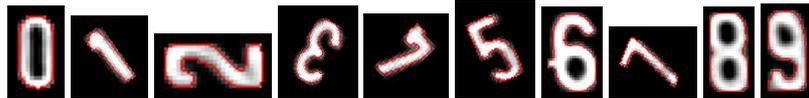


Fig. 2. Typical segmented license plate characters selected from the database

Table 1. Recognition rate with respect to Descriptor Size (M,L)

Method	Number of Fourier Coefficients (L)					
	15	10	7	5	3	2
Centroid Distance	90,21	90,14	84,54	76,11	59,9	52,95
$M=2$ angles	57,4	54,04	47,59	44,97	32,54	24,02
$M=3$ angles	82,69	81,4	76,83	73,4	59,26	42,61
$M=4$ angles	91,71	89,52	86,11	82,19	67,71	54,38
$M=5$ angles	96,14	95,64	93,57	90,35	81,81	71,45
$M=6$ angles	97,66	97,69	96,02	95,21	88,88	79,42
$M=8$ angles	99,28	99,47	99,14	98,73	95,73	88
$M=12$ angles	99,04	99,09	99,19	99,23	98	94,97
$M=16$ angles	99,16	99,42	99,33	99,61	98	96
Curvature	89,5	88,88	88,45	85,45	67,54	55,28
Complex Coordinates	93,66	89,28	91,54	73,19	50,42	35,9

In the first group of experiments we explored how the recognition rate changes with the descriptor size. Descriptor has two dimensions: number of kernel filters denoted as M and the number Fourier coefficients denoted as L . The results are summarized in Table.1. The proposed descriptor receives 99% recognition rate for $M>6$ and $L>3$. When compared to centroid distance, its performance is at most 90% for $L=15$. Another observation is that the performance of centroid distance drops dramatically when L gets smaller. For example its recognition performance is 60% for $L=3$ where as the proposed descriptor performance much better, 95%, for $M=8$ and $L=3$. In Fig. 3 we plot the recognition rates and how they change with L for four methods (i.e. the proposed descriptor ($M=8$), centroid distance, boundary curvature, complex coordinates).

In the second group of experiments we explored how the recognition rate changes with the scale parameter, λ , of the filter kernel. We change the filter size, hence the scale from 3×3 to 15×15 . The results are summarized in Table.2.

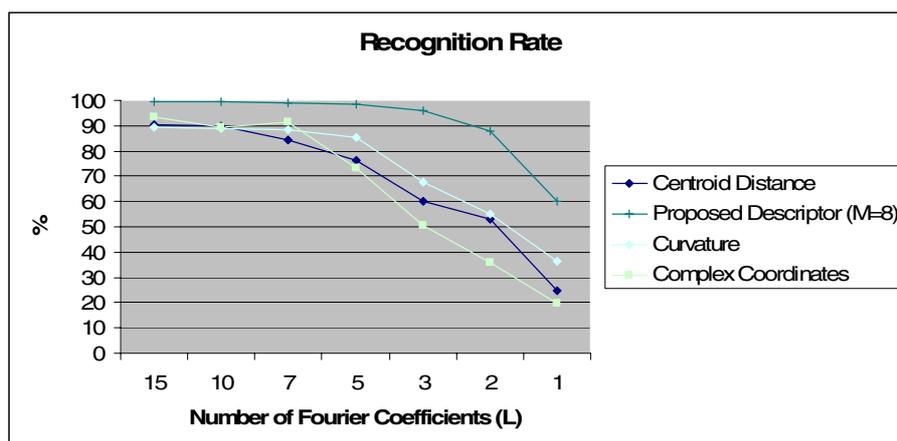
**Fig. 3.** Recognition rates for four methods

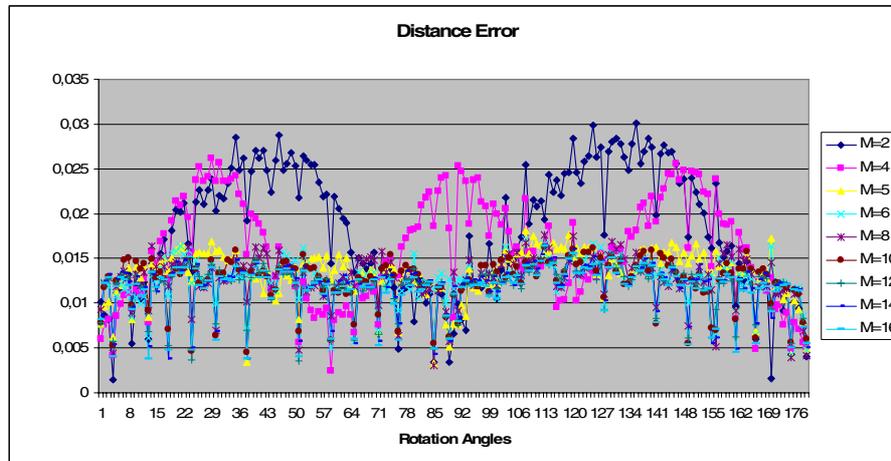
Table 2. Recognition rate with respect to filter scale

	Filter Size			
	7×7	9×9	11×11	13×13
The Proposed Descriptor ($M=8,L=10$)	97,19	98,73	99,59	99,54

Table 3. Average distances between the object and its 1° rotated versions (Binary object results are given in the first row, and Gray-level object results are given in the second row.)

M=2	M=4	M=5	M=6	M=8	M=10	M=12	M=14	M=16
0,0179	0,0133	0,0126	0,0127	0,0125	0,0119	0,0119	0,0116	0,0116
0,0121	0,0074	0,0073	0,0063	0,0058	0,0052	0,0046	0,0040	0,0036

In the last group of experiments we explore how M and ϕ effect the shifting property of the descriptor. We rotate the same object by 1° angles and compute the distance defined in the previous section (Eq. 3.3) between the original object and its rotated versions. Due to lack of space only the average results are given in Table.3. We compute the distances for both gray level and binary objects. The first row of the Table.3 holds the average distances for 9 different kernel sets applied to binary objects. The results verify the Eq. 2.6. The shape is much better described as M increases. Gray-level object results are given in the second row. The results are collectively plotted in Fig. 4 for $M=2,4,5,6,8,10,12,14,16$.

**Fig. 4.** Distance error between the object and its rotated versions for $M=2,4,5,6,8,10,12,14,16$

5 Conclusion

In this study, we present a new affine invariant object shape descriptor, employing steerable filters and Fourier Descriptors. The proposed system utilizes not only the

boundary point coordinates of the objects, but also the filter responses along the boundaries. We compare the recognition performance of new shape descriptor with well-known boundary based shape descriptors on a database including rotated grey-level license plate characters. The experimental results show that, the proposed system dramatically outperforms the other shape descriptors.

Using such gradient based shape descriptor is very effective especially used with active contour segmentation techniques which employ shape priors. The main reason is, a gradient based shape recognizer does not describe the object during the active contour is not near the real object boundary; even if the contour gets the prior shape.

We will evaluate the performance of the proposed method on shape retrieval image databases as a future work.

References

- [1] Kurt B., Gökmen M., "Two Dimensional Generalized Edge Detector," 10th International Conference on Image Analysis and Processing (ICIAP'99), pp.148-151, Venice Italy, 1999.
- [2] Costa L. F. and Cesar Jr. R. M., "Shape Analysis And Classification: Theory And Practice, CRC Press New York, 2001
- [3] Veltkamp R. and Hagedoorn M., "State-of-the-art in Shape Matching", Technical Report UU-CS-1999.
- [4] Zhang D. and Lu G, "Review of shape representation and description techniques", Pattern Recognition 37 pp. 1 – 19, 2004
- [5] Freeman W.T. and Adelson E.H., "The Design and Use of Steerable Filters", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 13, no. 9, pp. 891-906, 1991.
- [6] Yokono J.J. and Poggio T., "Oriented Filters for Object Recognition: an Empirical Study", Automatic Face and Gesture Recognition, pp. 755-760, 2004
- [7] Balard, D.H.; Wixson, L.E., "Object recognition using steerable filters at multiple scales"; Qualitative Vision, Proceedings of IEEE Workshop pp. 2 – 10, June 1993
- [8] Talleux, S., Tavşanoğlu, V., and Tufan, E., "Handwritten Character Recognition using Steerable Filters and Neural Networks," ISCAS-98, pp. 341-344, 1998.
- [9] Li, S. and Shawe-Taylor, J. , "Comparison and fusion of multiresolution features for texture classification", Pattern Recognition Letters Vol:26, Issue:5, pp. 633-638, 2005
- [10] Zhang D. and Lu G, "A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval", Journal of VCIR, Vol. 14, pp. 39-57, 2003
- [11] Rafiei D. and Mendelzon A. O., "Efficient retrieval of similar shapes", The VLDB Journal 11, pp. 17-27, 2002
- [12] Antani S., Leeb D.J., Longa L.R., Thoma G.R., "Evaluation of shape similarity measurement methods for spine X-ray images", Journal of VCIR, Vol.15, pp. 285-302, 2004
- [13] Phokharatkul P., Kimpan C. "Handwritten Thai Character Recognition Using Fourier Descriptors and Genetic Neural Network", CI 18(3), pp. 270-293, 2002
- [14] Kunttu, I., Lepisto L., Rauhamaa J., Visa A., "Multiscale Fourier descriptors for defect image retrieval", Pattern Recognition Letters 27, pp. 123–132, 2006
- [15] Zhang D. and Lu G, "Study and evaluation of different Fourier methods for image retrieval", Image and Vision Computing, Vol. 23, No. 1, pp. 33-49, 2005