INVARIANCE IMPROVED LEGENDRE MOMENTS AS CONTOUR-SHAPE DESCRIPTOR

V. P. Dinesh Kumar, J. Linesh, Roopesh Katan, Tessamma Thomas

Department of Electronics, Cochin University of Science and Technology, Kerala, Cochin - 682 022, INDIA, e-mails: *dineshvp@cusat.ac.in, tess@cusat.ac.in*

Abstract. A new shape descriptor based on improved Legendre moments, invariant to affine transformations and suitable as both contour and region-shape descriptor is proposed in this paper. The objective performance comparison of the improved Legendre descriptor with modified Zernike descriptor was carried out. The Average normalized modified retrieval rank (ANMRR) obtained on SQUID database for improved Legendre descriptor was 0.1673, and for modified Zernike descriptor was 0.3514, which depicts the better accuracy of the proposed technique. The subjective evaluation of the retrieval results of Legendre and Zernike descriptor also shows the better performance of the Legendre descriptor.

1. Introduction

Object shape features provide a powerful clue to object identity and functionality, and can be used for computer vision applications. A good descriptor captures characteristic shape features in a concise manner, which is invariant to scaling, rotation, translation and to various types of shape distortions. The multimedia content description interface, MPEG-7 [9], provides contour / region-based descriptors, for 2D shapes.

The region-based shape descriptor expresses pixel distribution within a 2-D object region, whereas, the contour-based shape descriptor contains the shape properties of the object outline (contour). Shape analysis techniques, based on moments have given rise to various region-based shape descriptors [3]. The contour-based shape descriptor is based on the curvature scale- space (CSS) [2] representation of the contour. But, these dedicated descriptors cannot describe the region-shape properties. It has been shown that Zernike moment [10] descriptor outperforms many other shape descriptors like geometric moments, grid descriptor etc in terms of representation effectiveness [5], [6], [11]. It is important to have unique visual feature descriptors that are capable of distinguishing different region-shape properties and contour-shape properties alike, for content based image retrieval applications in an image database. In pursue of this, the effectiveness of orthogonal moments as contour-shape descriptors is explored in this paper.

The Legendre moments [10], which are orthogonal like the Zernike moments, have better image representation capability with fewer coefficients compared to Zernike Moments [8]. But, proper investigation to their usefulness as shape descriptors is rarely done. This is due to their lack of inherent invariance to affine transformations. In this paper, methods for invariance improvement of Legendre moments are discussed and their usefulness as shape-descriptor is established by objective and subjective comparison with the Zernike moment descriptor.

2. Orthogonal moment descriptors

The orthogonal moments, which is the projection of an image function onto an orthogonal polynomial, provides compact image representation due to the lack of information redundancy.

2.1. Zernike moment descriptor

The kernel of Zernike moments is the set of orthogonal Zernike polynomials defined over the polar coordinates inside a unit circle. The Zernike basis function [10] of order n, and repetition l is

$$V_{n\,l}(x, y) = V_{n\,l}(\rho, \theta) = R_{n\,l}(\rho) \exp (j\, l\, \theta) \quad (1)$$

where n is a positive integer or zero, ρ is the length of vector from origin to (x, y) pixel, θ is the angle between vector ρ and x axis in counterclockwise direction and the repetition l is an integer subject to the constraints: $n - |l|$ is even and $|l| \le n$. $R_{n\,l}(\rho)$ is the radial polynomial given by,

$$R_{n\,l}(\rho) = \sum_{s=0}^{n-|l|/2} (-1)^s . \frac{(n-s)!}{s!(\frac{n+|l|}{2}-s)!(\frac{n-|l|}{2}-s)!}\rho^{n-2s}. \quad (2)$$

The Zernike moment for a discrete image function $F(x, y)$ (a digital image) of order n with repetition l that vanishes outside the unit circle is

$$Z_{nl} = \frac{n+1}{\pi} \sum_x \sum_y F(x, y) V^*_{nl}(\rho, \theta) \tag{3}$$

where $x^2 + y^2 \leq 1$ and $V^*_{n\,l}$ is a complex conjugate of $V_{n\,l}$.

The magnitude of the Zernike moment coefficients gives rotation invariance. Translation invariance is achieved by moving the origin to the centre of the image by using the centralized moments. Zernike basis function radius is chosen as the maximum radius of the image shape. Since normalization of digital images generates errors of re-sampling and re-quantifying, zeroth-order geometric moment (Equation (4)) of the image is used to normalize the Zernike moments to provide scale invariance [1].

$$g_{00} = \sum_x \sum_y F(x, y) \tag{4}$$

The modified Zernike moment descriptor has 35 coefficients (order n = 1…10).

2.2. Legendre moment descriptor

The Legendre moments are based on Legendre polynomials [10] which form an orthogonal basis set in the interval [-1 1]. The p^{th} order Legendre polynomial is defined as follows.

$$P_p(x) = \frac{1}{2^p\, p!} \frac{d^p}{dx^p} (x^2 - 1)^p = \sum_{j=0}^{p} a_{pj} x^j \tag{5}$$

The values of $a_{p\,j}$ can be expressed as:

$$a_{pj} = \sum_{\substack{j=0 \\ j,p-odd \\ j,p-even}}^{p} (-1)^{(p-j)/2} \frac{(p+j)!}{2^p ((p-j)/2)!((p+j)/2)!\, j!} \tag{6}$$

The Legendre moments of order $(p + q)$ for a continuous image function $f(x, y)$ are defined as

$$\lambda_{pq} = \frac{(2p+1)(2q+1)}{4} . \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P_p(x) P_q(y) f(x, y) dx dy \tag{7}$$

where $p, q \geq 0$ and $p, q \in Z$.

In the case of a digital image $F(x, y)$, Equation (7) can be re-written as

$$\lambda_{pq} = \frac{(2p+1)(2q+1)}{4} . \sum_x \sum_y P_p(x) P_q(y) F(x, y) \tag{8}$$

and the pixel coordinates (x, y) are mapped to the interval [-1, 1].

2.3. Improvement of Legendre descriptor

In this section, the proposed improvements to Legendre moment descriptors for effective invariance to transformations like scale, translation and rotation, are discussed.

1. Rotate the image object so that major axis is horizontal.
2. The bounding rectangle (with the maximum width and height of the object as size) of the image object is mapped onto the basis set [-1 1] in each dimension.
3. Compute Legendre moments as per Equation (8).
4. Normalize the Legendre moments with the mass of the image (g_{00}), $\bar{\lambda}_{pq} = \lambda_{pq} \cdot (g_{00})^{-1}$.
5. Use the absolute value of improved Legendre moments, $\left|\bar{\lambda}_{pq}\right|$ as the features of shape.

The rotation invariance is ensured by aligning the major axis of the image with x-axis. Since, the bounding rectangle of the image object was mapped onto the basis set, even scale change such as zoom due to the camera operations does not affect the improved Legendre descriptor. The object location in the canvas is also immaterial due to translational invariance. Final normalization using the mass of the image brings in more scale invariance to the descriptors. The shape feature descriptors based on improved Legendre moments have 20 coefficients (order p = 1…5).

3. Experimental results and discussion

To evaluate the performance of the proposed improved Legendre moment descriptors, experiments were conducted. The tests were carried out using MPEG-7 approved contour-shape database SQUID [12], which consists of 1100 binary fish contours. All the distance measures between the feature vectors were taken in the L1 norm.

The performance of the improved Legendre descriptors for content based image retrieval (CBIR) application was evaluated. The retrieval accuracy was compared against that of modified Zernike descriptors

(Section 2.1). As the SQUID database is unclassified it was difficult to generate the ground truth set for all the 1100 images. So, 10 query images from the first 200 images were chosen and their ground truth was defined subjectively (features such as shape of the body, tail, shape and number of fins etc were used). The MPEG-7 accepted *average normalized modified retrieval rank* (ANMRR) is used as the numeric measure to weigh query results [7]. The rank of the k^{th} retrieved image for query q is defined as

$$\text{Rank}^*(k) = \begin{cases} \text{Rank}(k), & \text{if Rank}(k) \le K(q) \\ 1.25K, & \text{if Rank}(k) > K(q) \end{cases}$$

where K, the tolerance of the system, is defined with respect to the ground truth size NG(q) as follows:
$K = 2 * \text{NG}(q)$; if NG(q) >= 20; $K = 3 * \text{NG}(q)$; if NG(q) >= 10; $K = 4 * \text{NG}(q)$; otherwise.

The average rank (AVR) for query q, $AVR(q) = \frac{1}{NG(q)} \sum_{k=1}^{NG(q)} Rank^*(k)$ and modified retrieval rank (MRR) is $MRR(q) = AVR(q) - 0.5*[1+NG(q)]$. Normalizing with respect to NG(q) leads to *normalized modified retrieval rank* (NMRR)

No	Query image	Zernike descriptor (NMRR)	Legendre descriptor (NMRR)
1	kk3	0.1751	0.0276
2	kk12	0.2806	0.2889
3	kk20	0.2711	0.1761
4	kk35	0.4717	0.3208
5	kk41	0.4332	0.1935
6	kk53	0.4714	0.1429
7	kk76	0.3273	0.2545
8	kk100	0.1346	0.0699
9	kk104	0.4784	0.1991
10	kk157	0.4706	0
ANMRR		0.3514	0.1673

Table 1. *ANMRR results of moment based descriptors*

$$NMRR(q) = \frac{MRR(q)}{1.25*K - 0.5*[1+NG(q)]}. \tag{9}$$

NMRR(q) can take on values between 0 (indicating whole ground truth found) and 1 (indicating none from ground truth found), irrespective of ground truth size. The *average normalized modified retrieval rank* (ANMRR) indicates the retrieval quality over all queries and is defined as

$$ANMRR = \frac{1}{NQ} \sum_{q=1}^{NQ} NMRR(q) \tag{10}$$

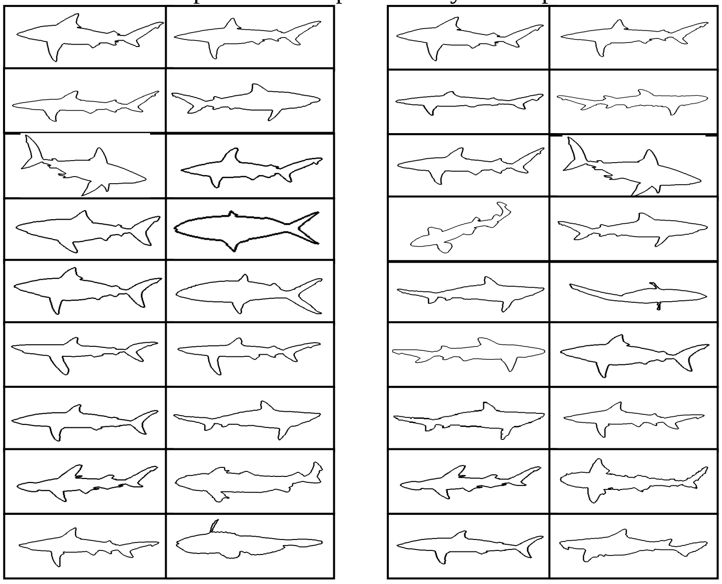
where NQ is the total number of queries. Table 1 lists the NMRR values obtained for 10 queries using modified Zernike moments and improved Legendre moments. The ANMRR value indicates the better performance of improved Legendre moments over the other.

For subjective evaluation of the performance of the proposed descriptor, first 18 retrieved images for the same query image using modified Zernike descriptor and Legendre descriptor is presented in Fig. (1). The top, left image is the query image. The query was done on the complete database. The visual features considered in the case of the fish in the query image are asymmetric tail, number and location of the fins etc. It can be observed from the figure, that the number of mismatches is more in the case of modified Zernike moment descriptor.

Thus, the proposed Legendre moment based shape-descriptor has good contour-shape description property. The effectiveness of orthogonal moments as region-shape descriptors is already established [4], [5], [6]. Another important criterion in the MPEG-7 specification is the compactness of the descriptor. Improved Legendre moments descriptor requires only 20 coefficients, whereas Zernike moment descriptor has 35 coefficients. Besides, Legendre moments are computationally less complex than Zernike moments [8].

4. Conclusion

In this paper, we have described an improved Legendre moment based contour-shape descriptor suitable for content based image retrieval applications. The comparative performance analysis between the proposed descriptor and the modified Zernike descriptor has shown the merit of the improved Legendre moment based approach. The improved Legendre descriptor is a suitable candidate as shape-descriptor for CBIR applications in an image database. As per the MPEG-7 standard, the proposed descriptor not only has better retrieval accuracy but also is more compact and computationally less expensive.



(a) (b)

Fig. (1). Retrieval results of shape descriptors for Query-by-Example (a) Modified Zernike descriptor (b) Improved Legendre descriptor.

Acknowledgements

This research work is funded by Council for Scientific and Industrial Research (CSIR), India. The authors would like to thank Prof. Mokhtarian, Centre for Vision, Speech, and Signal Processing Laboratory, University of Surrey, U.K., for providing the SQUID database for our experiments.

References

[1]. Ye Bin, Peng Jia-Xiong, "Invariance analysis of improved Zernike moments", *Jour. Opt. A: Pure App Optics*, **4**, (2002), 606-614.

[2]. Miroslaw Bober, "Shape descriptor based on curvature scale space", *MPEG-7 proposal P320*, Lancaster, U.K., 1998.

[3]. Miroslaw Bober, "MPEG-7 Visual Shape Descriptors", *IEEE Tran. Cir. Sys. Vid. Technology*, **11**, 6, (2001), 716-719.

[4]. Jiann-Jone Chen, Cheng-Yi Liu, "A Universal Query Mechanism for Similarity Retrieval Based on Shape Information in Image Databases", *Proc. IEEE Int. Con. on Aco. Spe. Sig. Processing*, 4, (2002), 676-3679.

[5]. Alireza Khotanzad, Yaw Hua Hong, "Invariant image recognition by Zernike moments", *IEEE Tran. Pat. Ana. Mac. Intelligence*, **12**, 5, (1990), 489-497.

[6]. H. Kim, J. Kim, "Region-based shape descriptor invariant to rotation, scale and translation", *Sig. Pro.: Ima. Communication*, **16**, (2000), 87-93.

[7]. B. S. Manjunath, Jens-Rainer Ohm, Vinod V. Vasudevan, Akio Yamada, "Color and texture Descriptors", *IEEE Tran. Cir. Sys. Vid. Technology*, **11**, 6, (2001), 703-715.

[8]. R. Mukundan, S. H. Ong, P. A. Lee, "Image Analysis by Tchebichef Moments", *IEEE Tran. Ima. Processing*, **10**, 9, (2001), 1357-1364.

[9]. Thomas Sikora, "The MPEG-7 Visual Standard for Content Description – An Overview", *IEEE Tran Cir. Sys. Vid. Technology*, **11**, 6, (2001), 696-702.

[10]. M. R. Teague, "Image analysis via the general theory of moments", *Jour. Opt. Soc. America*, **70**, 8, (1990), 920-930.

[11]. D. S. Zhang, G. Lu, "Content-based shape retrieval using different shape descriptors: A Comparative study", *Proc. IEEE Int. Con. Mul. Expo*, Tokyo, Japan, (2001), 317-320.

[12]. http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html.