International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Various Shape Descriptors in Image Processing – A Review

Preetika D'Silva[1], P. Bhuvaneswari[2]

[1]PG Student, Department of Electronics and Communication, Rajarajeswari College of Engineering, Bangalore, India

[2]Associate Professor, Department of Electronics and Communication, Rajarajeswari College of Engineering, Bangalore, India

Abstract: ***Images have different features associated to them. Shape is an important visual feature in an image. It plays a vital role in Content Based Image Retrieval (CBIR). There are many shape descriptors that can be used in image processing. This paper presents a study on different shape descriptors. We have reviewed the contour based shape descriptors like Fourier Descriptors, Curvature Scale Space Descriptors and region based shape descriptors like Angular Radial Transform (ART), Image Moment Descriptors, Zernike Moments descriptors (ZMD), Geometric Moments Descriptors (GMD) and Grid Descriptors (GD). The advantages and disadvantages of these shape descriptors are also discussed.***

Keywords: Shape feature, Shape descriptors, Contour-based, Region-based.

1. Introduction

The different features related to images are texture, shape, color etc. Images when represented by shape compared to the other features, is much more effective in characterizing the content of an image [1], [10]. Though, the difficult task of shape descriptors is the precise extraction and representation of shape information. The construction of shape descriptors is even more complex when invariance, with respect to a number of possible transformations, like scaling, shifting and rotation, is required [1], [11]. The overall performance of shape descriptors can be divided into qualitative and quantitative performances. The qualitative characteristics involve their retrieval performance based on the captured shape details for representation. Their quantitative performance includes the amount of data required to be indexed in terms of number of descriptors, in order to meet certain qualitative standards [1], [12] as well as their retrieval computational cost. [1]

Contour-based shape descriptors such as Fourier descriptors [13], [14], shape signatures [16] and curvature scale space [15] use only boundary information [3], they cannot capture shape interior content. In addition these methods cannot deal with disjoint shapes where boundary may not be obtainable; hence they have limited applications. Now in region based techniques, all the pixels contained by the shape area are taken into account to gain the shape representation. Common region based methods use moment descriptors to describe shape [3], [4], [8]. These include geometric moments, Zernike moments Legendre moments, and pseudo Zernike moments. Recently, numerous researchers also use the grid method to describe shape [3], [11], [17], [18].

An additional type of multi-scale descriptors is defined directly on the original shape contours with no any preprocessing, including triangle area representation, shape tree and hierarchical procrustes matching. Triangle area representation (TAR) presents a measure of convexity/concavity of each contour point making use of the signed areas of triangles created by boundary points at different scales. The area value of every triangle is a measure for the curvature of corresponding contour point, and the sign of the area is zero, positive or negative when the contour point is on a straight line, convex or concave, respectively. This representation is effective in capturing both local and global characteristics of a shape [19], [22]

In Section 2, we describe the shape descriptors like Fourier descriptors, curvature scale space descriptors, Angular radial transform (ART), Image moment descriptors, Zernike moments descriptors (ZMD), geometric moments descriptors (GMD) and grid descriptors (GD).

2. Shape Descriptors

2.1 Fourier Descriptors

Fourier descriptors have characteristics, like simple derivation, simple normalization and its robustness to noise, which have made them very popular in a wide range of applications [1], [7]. Fourier descriptors are obtained by applying Fourier transform to a shape signature. The Fourier transform on a complex vector derived from the shape boundary coordinates gives us the Fourier descriptors. The shape boundary coordinates can be given as $(x_n\ y_n)$, . n=0,1,…….,N-1.[1]

The complex vector $\bar{U}$ is given by the difference of the boundary points from the centroid (x_c, y_c) of the shape

$$\bar{U} = \begin{pmatrix} x_0 - x_c + i(y_0 - y_c) \\ x_1 - x_c + i(y_1 - y_c) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_n - x_c + i(y_n - y_c) \end{pmatrix}, \text{ n=0, 1,…, N-1} \quad (1)$$

where

$$x_c = \frac{1}{N}\sum_{n=0}^{N-1} x(n) \quad , \quad y_c = \frac{1}{N}\sum_{n=0}^{N-1} y(n) \quad (2)$$

The location of the shape from boundary coordinates gives the centroid subtraction, which makes the representation invariant to translation. The Fourier transformed coefficients are then obtained by applying One-dimensional Fourirer transform on $\bar{U}$ [1]. These coefficients are given as

$$\overline{F_k}=\mathrm{FFT}[\bar{U}] \quad (3)$$

For scaling invariance, the magnitudes of the Fourier transformed coefficients $|F_k|$ are normalised by the magnitude of the first Fourier transformed coefficient $|F_0|$.

The obtained FD are translation, rotation and scale invariant. The high-frequency noise can be reduced to a large extent by limiting the number of coefficients k, leaving at the same time the main details of the patterns. However, this reduction also leads to loss of spatial information in terms of fine detail. [1]

The advantages of Fourier descriptors are its robustness, being able to capture some perceptual characteristics of the shape and they are easy to derive. The computation is efficient with fast Fourier transform (FFT). The disadvantage of Fourier descriptors is local features cannot be located, since in Fourier transform only the magnitudes of the frequencies are known and not the locations. [2]

2.2. Curvature Scale Space Descriptors

The Curvature Scale Space (CSS) descriptor treats shape boundary as a 1D signal, and analyses this signal in scale space [4], [6]. CSS is based on multi-scale representation and curvature to represent planar curves. Curvature is a local measure on how fast a planar contour is turning. The concavities/convexities of shape contour are found by examining zero crossings of curvature at different scales. These concavities/convexities are helpful for shape description since they represent the perceptual features of shape contour. The first step of the process is as same as that in computing FD [6]. The shape boundary coordinates are given as $(x_n\ y_n)$, n=0,1,…….,N-1.

Scale normalization is the second step which samples the whole shape boundary into a fixed number of points so that shapes with a different number of boundary points can be matched. The other two main steps in the process are the CSS contour map computation and CSS peaks extraction [6]. To calculate the CSS contour map, curvature is first derived from shape boundary points as follows.

$$k(t) = (\dot{x}(t)\ddot{y}(t) - \ddot{x}(t)\dot{y}(t))/(x^2(t) + \dot{y}(t)^2)^{3/2} \quad (4)$$

where $\dot{x}(t), \dot{y}(t)$ and $\ddot{x}(t), \ddot{y}(t)$ are the first and the second derivatives at location t, respectively.

Curvature zero-cross points are then positioned in the shape boundary. Gaussian smoothing is then applied which hence evolves the shape into the next scale. [6]

$$x'(t) = x(t) \otimes g(t,\sigma,c)\ ,\ y'(t) = y(t) \otimes g(t,\sigma,c) \quad (5)$$

where $\otimes$ means convolution, and $g(t,\sigma,c)$ is the Gaussian function as follows

$$g(t,\sigma,c) = e^{[-(t-c)^2]/2\sigma} \quad (6)$$

The evolving shape becomes smoother as σ increases. At each scale, new curvature zero-crossing points are located. This process continues until no curvature zero-crossing points are found. Then the acquired zero-crossing points are plotted onto the (t,σ) plane to construct the CSS contour map. The local maxima or peaks of the CSS contour map are subsequently extracted out and sorted in descending order of σ. These CSS peaks are used as CSS descriptors to index the shape after normalization. [1]

CSS is robust to noise and is reliable and fast. The main weakness of this technique is due to the problem of shallow concavities/convexities on a shape.

2.3 Angular Radial Transform

The Angular Radial Transform (ART) is a moment-based image description method .The ART is a complex orthogonal unitary transform defined on a unit disc based on complex orthogonal sinusoidal basis functions in polar coordinates. [1]

The ART coefficients, F_{mn} of order n and m, are defined by

$$F_{mn} = \iint V^*_{m,n}(x,y) f(x,y)\, dx dy \quad (7)$$

where $f(x,y)$ is an image function in polar co-ordinates and $V_{m,n}(x,y)$ is the ART basis function that is separable along the angular and radial directions [1]

$$V_{m,n}(r,\theta) = R_n(r) A_m(\theta) \quad (8)$$

with

$$A_m(\theta) = \frac{1}{2\pi} e^{jm\theta} \quad (9)$$

and

$$R_n(r) \begin{cases} 1\ , & (n=0) \\ 2\cos(\pi n r)\ , & (n>0) \end{cases} \quad (10)$$

The ART descriptor is defined as a set of normalised magnitudes of the set of ART coefficients. By using the magnitude of the ART coefficients rotational invariance is obtained [1]. This descriptor is robust to translations, scaling, multi-representation (remeshing, weak distortions) and noises. [9]

2.4 Image Moments

Image Moments (IM) have been proved applicable in a variety of recognition tasks. The chosen image moments are not invariant only under translation, rotation and scaling of the object but also under general affine transformation [1], [8]. The affine moment invariants are derived by means of the theory of algebraic invariants and more specifically by means of decomposition of affine transformation into six one parameter transformations. The six affine invariants used are defined below [1]

$$I_1 = \frac{1}{\mu_{00}^4}(\mu_{20}\mu_{11}^2 - \mu_{11}^2)$$

$$I_2 = \frac{1}{\mu^{10}}(\mu_{30}^2\mu_{03}^2 - \mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3 + 4\mu_{03}\mu_{21}^3 - 3\mu_{21}^2\mu_{12}^2)$$

$$I_3 = \frac{1}{\mu_{00}^{7}}\left(\mu_{20}(\mu_{20}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2)\right)$$

$$I_4 = \frac{1}{\mu_{00}^{11}}\left(\mu_{20}^3\mu_{03}^2 - 6\mu_{20}^2\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}^2\mu_{21}\mu_{02}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2 + 12\mu_{20}\mu_{11}^2\mu_{03}\mu_{21} + 6\mu_{20}\mu_{11}\mu_{02}\mu_{30}\mu_{03} - 18\mu_{20}\mu_{11}\mu_{02}\mu_{21}\mu_{12} - 8\mu_{11}^3\mu_{03}\mu_{30} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} + 9\mu_{20}\mu_{02}^2\mu_{21}^2 + 12\mu_{11}^2\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}^2\mu_{02}^2\mu_{30}\mu_{21} + \mu_{02}^3\mu_{30}^2\right)$$

$$I_5 = \frac{1}{\mu_{00}^{6}}(\mu_{40}\mu_{04} - 4\mu_{31}\mu_{13} + 3\mu_{22}^2)$$

$$I_6 = \frac{1}{\mu_{00}^{9}}(\mu_{40}\mu_{04}\mu_{22} + 2\mu_{31}\mu_{22}\mu_{13} - \mu_{40}\mu_{13}^2 - \mu_{04}\mu_{31}^2 - \mu_{22}^3) \quad (11)$$

The proposers of the moment-based method have recommended making use of either all six or only the first four invariant moments for the description of objects. [1]. IM have the advantage of being computationally simple. Furthermore, they are invariant to rotation, scaling and translation. However, they have quite a few drawbacks like these moments suffer from a high degree of information redundancy since the basis is not orthogonal, higher-order moments are very sensitive to noise, and the moments computed have large variation in the dynamic range of values for different orders. This may cause numerical instability when the image size is huge. [5], [9].

2.5 Zernike Moment Descriptors (ZMD)

The block diagram of the whole process of computing ZMD is showing Figure 1. [3]

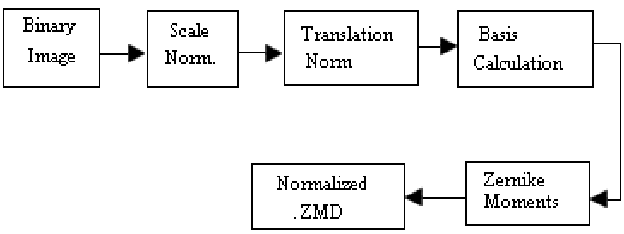

Figure 1: Block diagram of computing ZMD

Zernike moments allow independent moment invariants to be constructed to an arbitrarily high order. Zernike polynomials are used to get the complex Zernike moments. The Zernike polynomials are given as [23]:

$$V_{nm}(x,y) = V_{nm}(\rho\cos\theta, \rho\sin\theta) = R_{nm}\, e^{(jm\theta)} \quad (12)$$

where

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!}\rho^{n-2s} \quad (13)$$

where ρ is the radius from (*x*, *y*) to the shape centroid, θ is the angle between ρ and *x* axis, *n* and *m* are integers and subject to *n*-|*m*| = *even*, |*m*|≤*n*. Zernike polynomials are a total set of complex-valued function orthogonal over the unit disk, i.e., $x^2 + y^2 = 1$. Then the complex Zernike moments of order *n* with repetition *m* are defined as

$$A_{nm} = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)V_{nm}^*(x,y), \qquad x^2+y^2 \le 1 \quad (14)$$

where * means complex conjugate.

Since Zernike basis functions take the unit disk as their domain, this disk must be specified before moments can be computed. Only using magnitudes of the moments gives us the rotational invariance. The magnitudes are then normalized into [0, 1] by dividing them by the mass of the shape [23].The theory of Zernike moments is same to that of Fourier transform, which is to expand a signal into series of orthogonal basis. Though, the computation of Zernike moments descriptors does not need to know boundary information, making it suitable for more complex shape representation. [3]

Zernike moments have the following advantages like the magnitudes of Zernike moments are invariant to rotation, they are robust to noise and any minor variation in shape and as the basis is orthogonal, they have minimum information redundancy. However, the computation of ZM (in general, continuous orthogonal moments) have many problems like the image coordinate space must be transformed to the domain where the orthogonal polynomial is defined, the continuous integrals must be approximated by discrete summations. This approximation not only leads to numerical errors in the computed moments, but also strictly affects the analytical properties such as rotational invariance and orthogonality. Also as the order becomes large, the computational complexity of the radial Zernike polynomial increases.[9],[20].

2.6 Grid Descriptors (GD*)*

The Grid Descriptors were initially applied for contour based shape description [6] but in this paper we have studied it for region based shape description as well. In grid shape representation, a shape is projected onto a grid of set size, for example 16×16 grid cells [23]. The value 1 is assigned to the grid cells if they are covered by or inside the shape (or covered beyond a threshold) and 0 if they are outside the shape. The grid is then scanned from left to right and then from top to bottom and a binary sequence is created. Then this binary sequence is used as shape descriptors to index the shape. This binary sequence forms the shape number.

The complete process to construct the GD starts by finding out the major axis, i.e., the line connecting the two furthest points on the boundary. We can achieve Rotation normalization by turning the shape so that the major axis is parallel with the *x*-axis. To avoid multi normalization results for mirrored shape and flipped shape, the centroid of the rotated shape may be limited to the lower-left part, or a mirror and a flip operation on the shape number are applied in the matching step. [3]

We can implement Scale normalization by resizing the shape so that the length of the major axis is equivalent to the fixed grid width, and by shifting the shape to the upper-left of the grid, the representation becomes translation invariant. The next step is scanning the grid cells with the intention that a binary value is calculated for each cell based on the coverage of the cell by the shape boundary. Lastly, a binary sequence is generated as shape descriptors. The number of elements having different values gives the distance between two set of grid descriptors. This algorithm described is for contour based shape [3]

The block diagram of computing grid descriptors (GD) for a contour-based shape is given in Figure 2. [23]

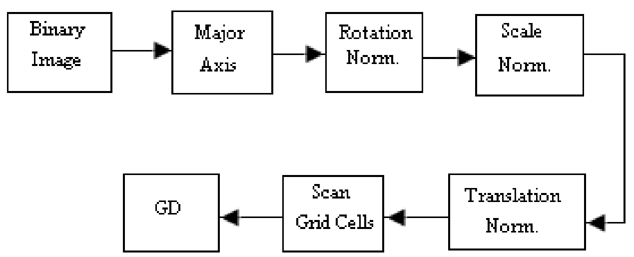

Figure 2: Block diagram of computing GD

For region-based shape, the process of GD generation is more complex. In the major axis computation step, to find the major axis of a region shape using point by point computation is not possible, the computation would be prohibitive. Hence, an algorithm of searching approximated major axis is used. The major axis for a region shape is found by finding the outer border point pairs on the shape boundary in numeral directions such as, 180 directions. The major axis is defined by the pair with the furthest distance. An interpolation process is followed the rotation normalization, since after rotation, the region points are scattered. GD is not robust, since a slight shape distortion, for example shear affine transform, skew and stretching can cause huge difference in the similarity measurement [3].

2.7. Geometric Moments Descriptors (GMD)

The Geometric Moment Descriptors is a technique based on moment invariants for shape representation and similarity measure and is widely used in shape recognition. Moment invariants are resultant from moments of shapes and are invariant to 2D geometric transformations of shapes.[6] The central moments of order *p*+*q* of a two dimensional shape represented by function *f*(*x*, *y*) are given by[23]

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y) \quad p,q = 0,1,2,\ldots \tag{15}$$

where $\bar{x} = \mu_{10}/m$, $\bar{y} = \mu_{01}/m$ and *m* is the mass of the shape region. μ_{pq} are invariant to translation. The first 7 normalized geometric moments that are invariant under translation, rotation and scaling are given as [3], [21]:

$$\phi_1 = \eta_{20} + \eta_{02}$$
$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4(\eta_{11})^2$$
$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$
$$\phi_4 = (\eta_{30} - \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$
$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$
$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$\phi_7 = (3\eta_{21} - \eta_{30})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$
(16)

A feature vector consists of the seven moment invariants $f = (\phi_1, \phi_2, \ldots \ldots, \phi_7)$is used to index each one shape in the database [3]. The values of the calculated moment invariants are generally small, values of higher order moment invariants are close to zero in few cases, thus, all the invariants are further normalized into [0, 1] by the limit values of each dimension. The benefit of using GMD is that it is a very compact shape representation and the computation is low, however, it is difficult to obtain higher order moment invariants [3].

3. Conclusion

In this paper, a few of the existing shape descriptors have been reviewed. There are two approaches in shape description namely contour-based and region-based. We have reviewed few techniques in each of these approaches. The contour-based approach utilizes only the outer boundary of the shape whereas the region-based approach utilizes the area covered inside the shape boundary. The contour-based approaches are simple whereas region-based approaches are more robust. Many shape descriptors other than the few mentioned in this paper can also be studied and implemented.

References

[1] Amanatiadi V.G. Kaburlasos A. Gasteratos S.E. Papadakis, "Evaluation of shape descriptors for shape-based image retrieval",IET Image Process., 2011, Vol. 5, Iss. 5, pp. 493–499

[2] Dengsheng Zhang, Guojun Lu, "Content-Based Shape Retrieval Using Different Shape Descriptors: A Comparative Study", Gippsland School of Computing and Information Technology

[3] Dengsheng Zhang, Guojun Lu, " A Comparative Study of Three Region Shape Descriptors", DICTA2002: Digital Image Computing Techniques and Applications, 21—22 January 2002

[4] Jalba, A., Wilkinson, M., Roerdink, J. , "Shape representation and recognition through morphological curvature scale spaces", IEEE Trans. Image Process., 2006, 15, (2), pp. 331–341

[5] M. E. Celebi and Y. A. Aslandogan, "A comparative study of three moment-based shape descriptors", in Proc. Of the International Conference of Information Technology: Coding and Computing, 2005, pp. 788-793
[6] CS Pillai, "A Survey of Shape Descriptors for Digital Image Processing", IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555,Vol. 3, No.1, February 2013
[7] Zhang, D., Lu, G., " Shape-based image retrieval using generic Fourier descriptor", Signal Process., ImageCommun., 2002, 17, (10), pp. 825–848
[8] Flusser, J., Suk, T., "Pattern recognition by affine moment invariants", Pattern Recognit., 1993, 26, (1), pp. 167–174
[9] Yang Mingqiang,Kpalma Kidiyo, Ronsin Joseph, " A survey of shape feature extraction techniques", HAL archives July 15, 2008
[10] Persoon, E., Fu, K., "Shape discrimination using Fourier descriptors", IEEE Trans. Syst. Man Cybern., 1997, 7, (3), pp. 170–179
[11] Belongie S., Malik, J., Puzicha, J., "Shape matching and object recognition using shape contexts", IEEE Trans. Pattern Anal. Mach. Intell., 2002, 24, (4), pp. 509–522
[12] Berretti, S., Del Bimbo, A., Pala, P., "Retrieval by shape similarity with perceptual distance and effective indexing", IEEE Trans. Multimedia, 2000, 2, (4), pp. 225–239
[13] G. Granlund., "Fourier Preprocessing for hand print character recognition", IEEE Trans. Computers, Vol. 21, 1972, pp.195-201
[14] Eric Persoon and King-sun Fu., "Shape Discrimination Using Fourier Descriptors", IEEE Trans. On Systems, Man and Cybernetics, Vol.SMC-7(3):170-179, 1977
[15] A. Sajjanhar, "A Technique for Similarity Retrieval of Shapes", Master thesis, Monash University, Australia, 1997
[16] E. R. Davies, "Machine Vision: Theory, Algorithms, Practicalities", Academic Press, 1997, pp.171-190
[17] G. J. Lu and A. Sajjanhar., "Region-based shape representation and similarity measure suitable for content-based image retrieval", Multimedia System 7:165-174, 1999
[18] M. Safar, C. Shahabi and X. Sun, "Image Retrieval by Shape: A Comparative Study", IEEE Int. Conf. On Multimedia and Expo, New York, USA, 2000, pp
[19] J. Wang, X. Bai, X. You, W. Liu, and L. J. Latecki, "Shape matching and classification using height functions",Pattern Recognit. Lett., Elsevier vol. 33, no. 2, pp. 134–143, 2012
[20] R. Mukundan, S. Ong, and P. Lee, "Image analysis by tchebichef moments", IEEE Trans. Image Processing, vol.10(9), pp. 1357_1364, 2001
[21] Ming-Kuei Hu., " Visual pattern Recognition by Moment Invariants",IRE Transactions on Information Theory, IT-8:179-187, 1962
[22] Alajlan, N., Rube, I.E., Kamel, M.S., Freeman, G.," Shape retrieval using triangle area representation and dynamic space warping",Pattern Recognition 40, 1911–1920, 2007.
[23] Dengsheng Zhang, Guojun Lu, "Evaluation of MPEG-7 shape descriptors against other shape descriptors", Multimedia Systems, Springer-Verlag, 2003