Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP 2007)

A New Invariant Descriptor For Shape Representation And Recognition

Xiying Wang[1] Guozhong Dai[1]
[1]Institute of Software, Chinese Academy of Sciences, Beijing 100080, China
E-MAIL:wangxy0823@sohu.com or xiying04@ios.cn,

Abstract: *Traditional Fourier Descriptor suffers the drawback of being sensitive to translation, rotation and scale transform. This paper presents a new method of shape representation, Polygon Fourier Descriptor (PFD), which can achieve the invariance on translation, rotation and scale transform naturally and without need of a process of normalization. By applying polygonal approximation to the contour firstly, the contour can be represented more efficiently. Moreover, PFD is robust enough to represent the contour even if some transformations happened to it. Finally, we apply PFD to the recognition of static gestures and obtain satisfying results.*

1. Introduction

In the object recognition or image processing, contour shape often plays a very important role for recognizing the object efficiently. Human being can obtain much useful information of object by means of observing its contour. However, computer cannot understand the image unless its information has been represented in a quantitative way. Thus, how to represent the contour efficiently has been always one of the most important problems in the field of image processing and computer vision [1].

Till now, there are a number of ways to represent the contour, for example, Freeman Chain Code[2], moment invariants[3][4] and Fourier descriptor[5]]. From Zhang[1]'s theoretical analysis and Kauppinen[9]'s experimental result, it can be said safely that Fourier descriptor (FD), which represents the shape of the object in the frequency domain, is one of the most effective methods for the shape representation.

In the traditional FD method, each point P_k (x_k, y_k) *on* the contour can be expressed as a complex number, $s(k) = x_k + j y_k$, where j is the square root of -1. The Fourier transform will be applied to those complex numbers .The contour of object is represented by a series of Fourier descriptors. The calculation of traditional Fourier descriptor is often time-consuming and it always suffers the drawback of being sensitive to translation, rotation and scale transform. To obtain the invariance on translation, rotation and scale transform, some methods have been proposed, one of which is that the phase information of FD is removed and only the magnitude of FD is used to represent the object's contour. Although this method is simple, sometimes it will lead to recognition error duo to the loss of phase information that may includes some important information about the contour [10] for recognition.

In Keyes's paper[11], rotation normalization is achieved by finding the two coefficients with the largest magnitude and setting their phase angle equal to zero. In order to keep the orientation and the starting point of the contour defined uniquely, z(1) and z(2) is often assumed to be the two largest coefficients. However, it's generally not the case that z(2) is the second largest coefficient of magnitude[11]. Folkers and Samet[7] set the first Fourier descriptor z(0) to be equal to zero which moves the centroid of the contour onto the origin. As a result, the information of z(0) is discarded though it is very important for expressing the global character of the contour. Besides, rotation invariance is obtained by using the orientation of the basic ellipse, which leads to an ambiguity of π radians. Therefore, the rotation invariance is only rotation invariant modulo a rotation by π radians[7]. Beside, traditional Fourier descriptor is also sensitive to the starting point.

From the introduction, we can see that how to achieve invariance on translation, rotation, scale transform and even start point simply and efficiently is still an open problem. In this paper, we propose a new descriptor, Polygon Fourier Descriptor (PFD) that tries to solve the above problem.

2. Polygon Fourier Descriptor

Contour is composed of a series of pixels, in which there is usually lots of redundant information and noise that should be removed firstly. One of effective methods to remove noise is polygonal approximation. After the polygonal approximation, Fourier transform will be employed to calculate PFD.

The framework of the calculation of PFD is presented in Figure 1. In the following subsections, all steps will be explained in detail.

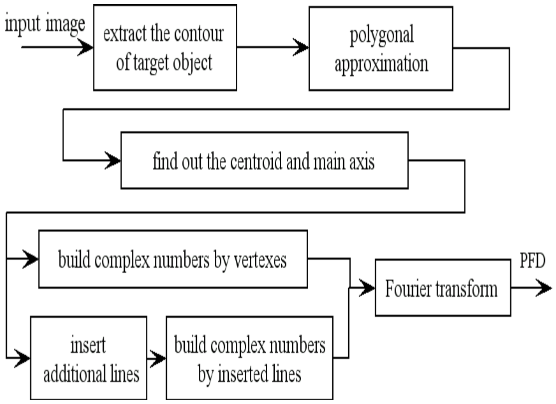


Figure 1. Calculation flowchart of PFD

2.1. Polygonal approximation of the contour

In this work, a Douglas-Peucker (D-P) approximation algorithm is adopted for polygonal approximation. The basic principle of D-P algorithm is that the chord of arc is used to approximate the contour with required accuracy. The approximation accuracy is determined by the largest distance between the point in the arc and the chord of the arc. Detailed description of D-P algorithm can be found in [12]:

The approximate polygon of the contour will be obtained with certain approximation accuracy ε . The original computational complexity of D-P algorithm is O(n^2). An improved version was proposed by Hershberger and Snoeyink[13], in which a worst-case running time is proportional to *nlogn*, and it has been adopted by this paper.

2.2. Complex number series for Fourier transform

In traditional Fourier descriptor or most modified versions of it, the coordinates of the contour points are used to calculate the Fourier descriptor. Unlike traditional method, we construct the complex number by using the connecting line between the vertex and the centroid of the polygon. Firstly, we define v_kC as the connecting line between vertex v_k and centroid C, ω_k as the angle formed by v_kC and the main axis, ranged from –180 degree to 180 degree, and l_k as the length of v_kC. And then, the corresponding complex number $f(k)$ can be formed by $f(k) = \omega_k + l_k j$.

Obviously, given the values of the angle ω_k and the length of connecting line l_k, a unique vertex can be determined. The position of centroid can be computed by Eq.2.1, where m_{ij} represents $(i + j)$ order moment of the shape.

$$x_c = \frac{m_{10}}{m_{00}}, \quad y_c = \frac{m_{01}}{m_{00}} \qquad (2.1)$$

The main axis of the approximate polygon is determined through the method described as follows. Firstly, the length of connecting lines between each vertex and centroid should be calculated, and then the direction of the connecting line with the largest length is defined as the direction of main axis. If there is more than one line with the largest length, we choose the line with the smallest angle to X axis's positive direction in the counterclockwise direction to be the main axis.

Furthermore, in order to keep invariant, the length component l_k is set to be the relative length to the main axis's, i.e. $l_k = |v_k c| / |\text{length}_{\text{main}}|$. The value of l_k ranges from 0 to 1. Considering the different value scope of the angle component and the length component, a normalization process is applied to ω_k by defining $\omega_k = \omega_k / 360$.

Consequently, no matter how the position of contour changes, the value of PFD will not be influenced. Similarly, the invariance to rotation and scale transform can also be achieved naturally, since the values of ω_k and l_k would not change even if the image is rotated、enlarged or reduced.

2.3. Robustness of PFD

It should be noted that, in the process of polygonal approximation, the approximate polygon of the contour usually has different number of vertices due to the different approximation accuracy. PFD should be robust enough so that those contours, which have similar geometrical characteristics but different polygon vertices after polygonal approximation, can be represented identically.

To enhance the robustness of PFD, a number of additional lines are inserted between the adjoining connecting lines of vertex and centroid by a certain interval angle. As shown in Fig.2, the points A and B are two neighboring vertices on the approximate polygon, and the centroid of the polygon is O. The angle between the connecting lines *OA* and *OB* is labeled by α , and the interval angle used to insert additional lines is β .The inserted lines intersect the line *AB* at points *C, D*, …and *G* respectively, the number of which is $\left\lfloor \frac{\alpha}{\beta} \right\rfloor$, i.e., the largest integer that not larger than α / β .

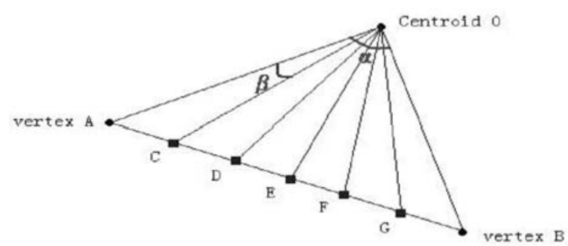


Figure 2. Insert lines between adjoining connecting lines

As for the length of inserted line, e.g. |*OC*|, it can be deduced according to the formula of the area of triangle. The length of other inserted lines can also be worked out like |*OC*|. Eq.2.2 is the general formula for the calculation of the length of inserted lines.

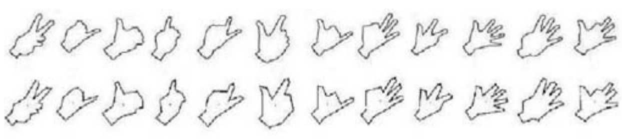
$$|OI| = \frac{|OA| * |OB| * \sin\alpha}{|OA| * \sin(i * \beta) + |OB| * \sin(\alpha - i * \beta)} \qquad (2.2)$$

where $|OI|$ is the length of the i-th inserted line and i is set to $1,2,\ldots\ldots\left\lfloor \frac{\alpha}{\beta} \right\rfloor$. Then, we can build new complex numbers by $f(x_i) = (i * \beta) + |OI|\, j,\ i = 1,2,\ldots\ldots\left\lfloor \frac{\alpha}{\beta} \right\rfloor$.

Finally, these new complex numbers, together with those complex numbers previously formed by connecting vertexes and centroid, are applied by Fourier transform to generate the Polygon Fourier Descriptor series.

3. Experiments

In order to testify the availability of PFD, we apply it to the recognition of static hand gestures in our experiments. Firstly, twelve different static hand gestures are designed as templates in Fig.3 in which the above row contains twelve gesture contours, and the second row includes the corresponding approximate polygons of the contours above.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k) (l)

Figure 3. The twelve hand gesture contours and their approximate polygons

In our first experiment, the invariance on translation, rotation and scale transform of PFD was testified. Firstly, one contour was selected from the twelve templates randomly and some combined transformations were applied to it, and then PFD of the transformed contour was calculated. We can test the invariance of PFD by matching the transformed contour with those templates.

In order to do matching by PFD, the Euclidean distance was adopted to calculate the distance between transformed contour and each template, as shown in Eq.2.3,

$$dis(A,B) = \sqrt{\sum_{i=0}^{N-1} (MOD(A_i) - MOD(B_i))^2} \tag{2.3}$$

where A and B are characteristic vectors with the same dimension N.

In the experiment, template f was selected randomly and applied various combined transformations (see Table 1). After the combined transformation of rotation and scale, random translation was applied too. Subsequently, these transformed contours were approximated by D-P algorithm with same accuracy. The approximate polygons of transformed contours, labeled by f_{ij}, are shown in Table 1. Each approximate polygon f_{ij} was matched with the twelve templates by way of PFD, that is, first 20 Polygon Fourier Descriptors were adopted to calculate the distance according to the distance formula Eq.2.3.

Figure 4 illustrates the matching result that achieves high correct ratio, and all transformed templates are matched to the target correctly.

Table 1. Approximate polygons of the transformed templates

scaling	angle of rotation			
	40°	110°	240°	330°
50%	f_{11}	f_{12}	f_{13}	f_{14}
150%	f_{21}	f_{22}	f_{23}	f_{24}
400%	f_{31}	f_{32}	f_{33}	f_{34}

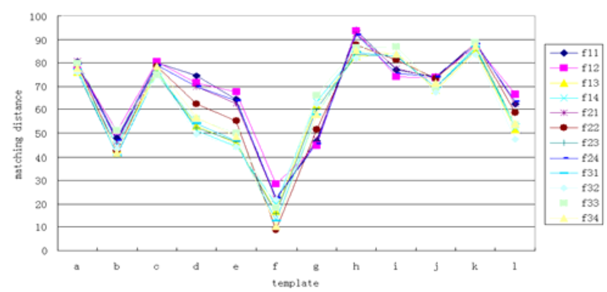


Figure 4. The graph of matching result

In order to compare the representation ability of PFD with traditional FD, in our second experiment, all twelve templates were transformed by those combined transformations described in Table 1, and calculated their traditional FD and PFD respectively. The normalization of traditional FD was achieved through the method used by Laura and Adam [10]. Then, we selected the first 20 traditional FDs and PFDs respectively as the input data to do matching according to Eq.2.3. Matching results of traditional Fourier descriptors and PFD are shown in Figure 5.

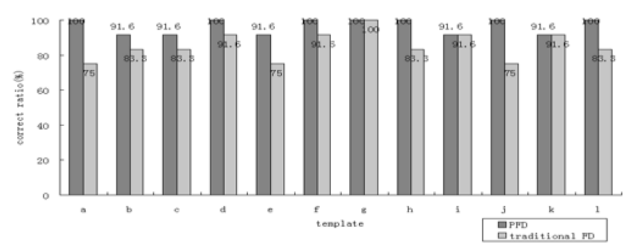


Figure 5. Comparison graph between PFD and traditional FD in correct matching ratio

Fig.5 shows that PFD achieves a higher correct ratio than traditional Fourier descriptor under the similar conditions, and thus it can be concluded that PFD has a more powerful ability of representation.

As for the computational cost of PFD, we can see that the total number of vertices and inserted lines is much less than the pixel number of the contour. Besides, there is no need of additional normalization process in PFD to achieve the invariance on rotation, scale and translation.

4. Conclusion

In this paper, we present a new approach, Polygon Fourier Descriptor, to represent object's contour. At the beginning, polygonal approximation is applied to the contour in order to remove the redundancy and noise from it. Because only vertices and some inserted lines are used to calculate PFD and additional normalization process is avoided, the computational cost of Fourier transform is reduced remarkably. Moreover, unlike most modified versions of traditional FD, PFD uses the relative value of angle and length to form complex number, which makes the invariance on translation, rotation and scale transform achieved in a natural way without losing any information.

According to the experiments and their results, it can be seen clearly that Polygon Fourier Descriptor has such characteristics as follows:

1). Invariant to translation, rotation and scale transform, and independent of starting point. Especially, the invariance is an inherent characteristic of PFD, that is, additional normalization process can be omitted.

2). Can be calculated more simply and quickly than traditional Fourier descriptor, and proved effective in gesture recognition.

3). Robust enough to represent the contour even if the contour is approximated by different approximation accuracy or some transforms are applied to it.

Acknowledgement

This work was supported by National Basic Research Program (973 Program) of China under Grant No. 2002CB312103 and 863 Program under Grant No. 2006AA01Z328.

References

[1] Dengsheng Zhang, Guojun Lu, "Review of shape representation and description techniques", Pattern Recogntion, 2004, 37, pp.1~19.

[2] H. Freeman, "On the encoding of arbitrary geometric configurations". IRE Trans. on Electronic Computer, 1961,10, pp. 260~268.

[3] M.K. Hu, "Visual pattern recognition by moment invariants", IRE Trans. on Information Theory, 1962, 8, pp. 179~187.

[4] S.O. Belkasim, M. Shridhar, "Pattern recognition with moment invariant: a comparative study and new results", Pattern Recognition, 1991, 24, pp. 1117~1138.

[5] C.T. Zahn, R.Z. Roskies, "Fourier descriptors for plane closed curves", IEEE Trans. on Computers, 1972, 21, pp. 269~281.

[6] E. Persoon, K.S.Fu, "Shape discrimination using Fourier descriptor", IEEE Trans. on PAMI, 1986, 8, pp. 388~397.

[7] Andre Folkers, Hanan Samet, "Content-based image retrieval using Fourier descriptors on a logo database", ICPR , 2002, 3 , pp. 521-524.

[8] R.C. Gonzalez, R.E. Woods, "Digital Image Processing (second edition)", Prentice Hall, Beijing China. 2002, pp 532~535.

[9] H. Kauppinen, T. Sepanen, "An experiment comparison of autoregressive and Fourier-based descriptors in 2D shape classification", IEEE Trans. on PAMI, 1995, 2, pp.201~207.

[10] WANG Tao, LIU Wen-Yin, "Using fourier descriptor to recognize object's shape (In Chinese)", Journal of Computer Research and Development, 2002, 12, pp. 1714~1719.

[11] Laura Keyes, Adam Winstanley, "Fourier descriptors as a general classification tool for topographic shapes", In: Proceedings of Irish Machine Vision and Image Processing Conference. 1999, pp.193~203.

[12] Douglas.D.H, Peucker.T.K, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature", Canadian Cartographer, 1973,10(2), pp. 112~122.

[13] John Hershberger, Jack Snoeyink, "Speeding up the Douglas-Peucker line-simplication algorithm", In: processing of 5th Symposium on Data Handling, 1992, pp.134-143.