Shape Classification using Multiscale Fourier-based Description in 2-D Space

Cem Direkoğlu and Mark S. Nixon
School of Electronics and Computer Science,
University of Southampton, Southampton, SO17 1BJ, UK
{cd05r/msn}@ecs.soton.ac.uk

Abstract

In shape recognition, the boundary and exterior parts are amongst the most discriminative features. In this paper, we propose new multiscale Fourier-based object descriptors in 2-D space, which represents the boundary and exterior parts of an object more than the central part. This representation is based on using a high-pass Gaussian filter at different scales. The proposed algorithm makes descriptors size, translation and rotation invariant as well as increasing discriminative power and immunity to noise. In comparison, the new algorithm performs better than elliptic Fourier descriptors and Zernike moments with respect to increasing noise.

1. Introduction

Silhouette based object description and recognition is an important task in computer vision. The descriptor must be invariant to size, translation and rotation, and it must be effective in adverse conditions such as noise and occlusion. There are two main types of shape description methods: boundary-based methods and region-based methods. The most common boundary-based shape descriptors are chain codes [1], wavelet descriptors [2], Curvature Scale Space (CSS) [3] and Fourier descriptors [4]. Fourier descriptors can also be combined with the wavelet transform to smooth the boundary in one dimension to produce multiscale Fourier descriptors [5]. Common region-based shape descriptors are moments [4][6] and Generic Fourier Descriptors (GFDs) [7].

An object's silhouette can be discriminated mainly by the boundary, and exterior parts more than the central part. If we are to recognize a human from their silhouette, the thorax contributes little to recognition, whereas the limbs and head contribute much more. We produce multiscale Fourier-based object descriptors in 2-D space, which represent the boundary and exterior parts of an object and also allow the central part to contribute to classification slightly. Our algorithm starts with size normalization of the silhouette image and we then compute a Fourier magnitude image that is translation invariant. Representing the boundary in a 2-D space can be achieved by applying a high-pass filter to the Fourier magnitude image. To do this, we use a high-pass Gaussian filter that has a scale parameter (standard deviation), instead of using an ideal filter, and let low frequency components contribute slightly. Representing only the boundary is not sufficient, since it is sensitive to noise and variations in shape. We also need to represent exterior regions of an object, to increase discrimination power and immunity to noise. To achieve this, different scales are selected and as the scale of high-pass Gaussian decreases, filtered Fourier magnitude images will represent the exterior regions of an object. However, the Fourier magnitude images obtained at each scale vary with rotation and are not convenient for matching. To give rotation invariance, each Fourier magnitude image is polar mapped that converts rotation change to translation and then Fourier transform of polar image is computed to obtain the Fourier magnitude. Finally, the obtained Fourier magnitude image is size, translation and rotation invariant and represents descriptors of an object at that scale. For classification, the Euclidean distance is calculated separately at each scale and then the average distance is computed for each object. By classifying with average distance, increased immunity to noise as well as increased correct classification rate is observed. Figure 1 shows the proposed algorithm to obtain multiscale Fourier-based object descriptors.

The Fourier-Mellin transform is similar to our algorithm in terms of achieving rotation, size and translation invariance. The Fourier-Mellin transform is a method for rotation, size and translation invariant image feature extraction in 2-D space [8]. The first stage is a 2-D Fourier transform, then the Log-polar transform, and finally another Fourier transform. The Log-polar transform converts scale and rotation changes to translations. Converting scale change to translation is achieved by logarithmic scaling of the radius coordinate of the polar map image [9]. The difference from our new approach is that, we have a filtering approach to create a multiscale representation and filtering must be applied to the objects of same size. Because of this reason, object size is normalized as a first step and we do not need to apply logarithmic scaling to the radius coordinate of the polar transformed image.

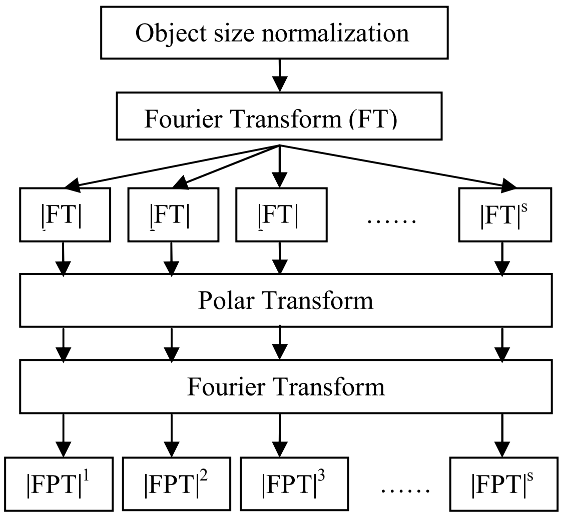


Figure 1: Producing the proposed multiscale Fourier-based object descriptors

The rest of the paper is organized as follows: Section 2 explains the proposed Fourier-based description with multiscale representation. Section 3 represents classification with multiscale Fourier-based object descriptors. Section 4 concerns evaluation and experimental results and finally Section 5 is conclusions.

2. Proposed Fourier-based description with multiscale representation

We produce Fourier-based object descriptors, which represent the boundary and exterior parts of an object more than the central part in 2-D space. The new algorithm starts with size normalization of an object using bilinear interpolation. Then, the 2-D Fourier transform is applied, as given in Equation 1, to compute Fourier magnitude image.

$$FT(u,v)=\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}I(x,y)e^{[-j2\pi(ux/M+vy/N)]} \quad (1)$$

Where $FT(u,v)$ is Fourier transfom of the silhouette image $I(x,y)$. $M\times N$ is the size of the silhouette image. The computed Fourier magnitude image, $|FT(u,v)|$, is translation invariant, however it retains rotation variance. Representing the boundary and exterior parts of an object is achieved at this stage. To represent the boundary, a high-pass Gaussian filter with scale parameter (standard deviation) is applied to the Fourier magnitude image as shown in Equation 2.

$$|FT(u,v)|^s=|FT(u,v)|\left(1-e^{-\sqrt{u^2+v^2}/2\sigma_s^2}\right) \quad (2)$$

Where $|FT(u,v)|^s$ and σ_s are Fourier magnitude and scale parameter of scale index s, respectively. By this way, low frequency components also contribute to classification slightly. However, using only the boundary information is both sensitive to noise and variations of shape. We also create other Fourier magnitude images by decreasing the scale of the high-pass Gaussian filter and represent exterior regions of an object. This process is shown in Figure 2. The proposed approach can be considered as multiscale representation of an object with respect to high-pass Gaussian filter in 2-D space. However, the obtained Fourier magnitude images are not convenient for matching, since they still have rotation variance. To remove rotation variance, the coordinates of each Fourier magnitude image are polar mapped to make rotations appear as translations in the new image. Consider a point $(x,y)\in\Re^2$ and define

$$(x,y)=(r*\cos\theta, r*\sin\theta) \quad (3)$$

Where $r\in\Re$ and $0\le\theta\le 2\pi$. For every point (x,y), there is a unique point (r,θ). Finally, another 2-D Fourier transform is applied, as given in Equation. 4, to compute Fourier magnitude, which removes these translations.

$$FPT^s(k,l)=\frac{1}{EF}\sum_{r=0}^{E-1}\sum_{\theta=0}^{F-1}P^s(r,\theta)e^{[-j2\pi(kr/E+l\theta/F)]} \quad (4)$$

Where $FPT^s(k,l)$ is the Fourier transform of the polar mapped image $P^s(r,\theta)$ of size $E\times F$ and at scale index s . The resultant Fourier magnitude image, $|FPT(k,l)|^s$, is translation, size and rotation invariant and represents object descriptors, OD^s , of shape at scale index *s*.

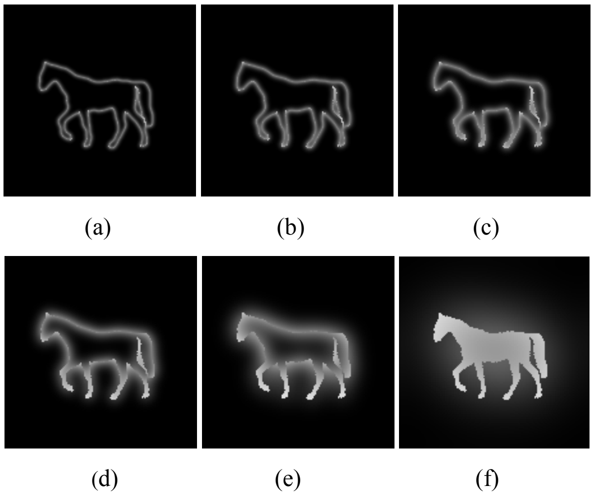


Figure 2: High-pass Gaussian filtered images with respect to decreasing scale. (a) $\sigma_1 = 15$, (b) $\sigma_2 = 11$, (c) $\sigma_3 = 8$, (d) $\sigma_4 = 5$, (e) $\sigma_5 = 3$, (f) $\sigma_6 = 1$

3. Classification with multiscale Fourier-based object descriptors

Classification is achieved using the nearest neighbour algorithm. Euclidean distance (Ed) is used to measure similarity between objects and is computed separately in each scale as given below,

$$Ed^s(T,D) = \sqrt{\sum_{i=1}^{E}\sum_{j=1}^{F}\left(OD_T^s(i,j) - OD_D^s(i,j)\right)^2} \quad (5)$$

Where $Ed^s(T,D)$ is the Euclidean distance between the object descriptors, OD_T^s, of the test image T and object descriptors, OD_D^s, of an image from database D, at scale index s. Then average distance is computed for each object.

$$Ad = \frac{1}{B}\sum_{s=1}^{B} Ed^s \quad (6)$$

Where Ad represents average distance and B is the number of scales. Classifying with average distance, instead of single scale distance, makes a significant increase in correct classification as well as increasing immunity to noise.

4. Evaluation and experimental results

For evaluation, we use MPEG-7 CE-Shape-1 Part B database. There are 1400 images in this dataset, which are pre-segmented into binary form. The objects are divided into 70 classes with 20 images of each class. These silhouettes change due to view point (scale, translation and rotation variance), due to non-rigid object motion (e.g. people walking and fish swimming), and due to noise inside shape (e.g digitization and segmentation noise). Some objects from the dataset and their variations are shown in Figure 3. Leave-one-out cross-validation is applied to validate classification. The performance evaluation compares the new algorithm with elliptic Fourier descriptors (EFD) and Zernike moments (ZM), with respect to increasing salt and pepper noise. Figure 4 illustrates salt and pepper noise corrupted binary images with increasing density. Although some objects in the dataset contain noise inside the shape, adding salt and pepper type noise cause noise outside the shape as well. Salt and pepper noise is added to all objects in the database; therefore noisy test image is matched with the noisy images from database.

EFD are one of the fast and robust boundary-based shape descriptors. The contour is represented with complex coordinates (position function) and then the Fourier expansion is performed to obtain EFD. To evaluate EFD, we use the algorithm given in [4]. Note that we describe the boundary of the biggest region in the image, since there will be many regions after adding noise.

Zernike moments (ZM) are region-based shape descriptors. They are an orthogonal moment set, which makes optimal utilization of shape information and allows accurate recognition. It is the most potent moment technique for shape description. To evaluate ZM, we use the algorithm given in [6].

In our algorithm, the object size is normalized to be 2500 in a 151×151 image. 5 different scales are selected for multiscale representation. The selected scales are: $\sigma_1 = 11$, $\sigma_2 = 8$, $\sigma_3 = 5$, $\sigma_4 = 3$ and $\sigma_5 = 1$. The size of the object descriptor matrix is 90×90 at each scale. Figure 5 shows the correct classification rate (CCR%) of the proposed algorithm, of EFD and of ZM, with respect to increasing salt and pepper noise. CCR is measured as follows,

$$CCR(\%) = \frac{c}{t}\times100 \quad (7)$$

where c is the total number of correctly classified images and t is the total number of classified images.

It is observed that proposed multiscale Fourier-based description in 2-D space performs better than EFD and ZM. Our algorithm achieves 95.5% correct classification rate, while ZM achieves 90% and EFD achieves 82% without adding noise to the database. As noise increases, the performance of all algorithms decreases and their performance degrade similarly. The success of our algorithm appears due to the multiscale representation of Fourier-based description in 2-D space, which emphasizes the boundary and exterior parts of objects and also lets the central part contribute slightly to classification. Although we use a multiscale representation, the proposed algorithm is easer to compute in comparison to Zernike moments. To obtain the Zernike moments, Zernike polynomials are computed, which are difficult and complex. On the other hand, in our algorithm, we rely on a polar transform and two Fourier transforms that are computed by Fast Fourier Transform (FFT) technique.

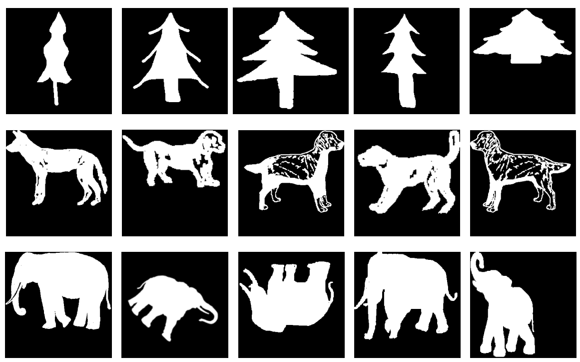
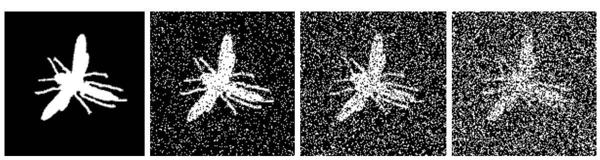


Figure 3: Some objects from the database and their variations



(a) D=0 (b) D=0.2 (c) D=0.4 (d) D=0.6

Figure 4: Fly object with increasing density (D) of salt and pepper noise

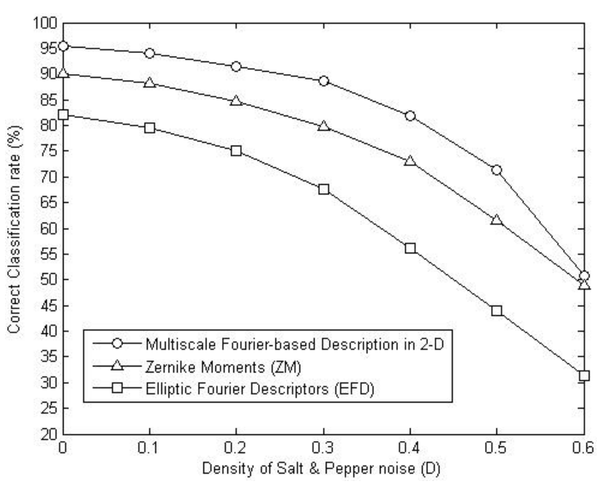


Figure 5: Classification performance of proposed algorithm, ZM and EFD, with respect to increasing salt and pepper noise in the database

5. Conclusion

We have presented novel multiscale Fourier-based description in 2-D space for shape classification. Multiscale object descriptors mainly represent boundary and exterior regions of an object, as well as allowing the central part to make a small contribution to classification. This representation is based on using a high-pass Gaussian filter at different scales. Our algorithm makes descriptors size, translation and rotation invariant. Evaluation indicates that our algorithm can perform better than Zernike moments (ZM) and elliptic Fourier descriptors (EFD) with respect to increasing salt and pepper noise. Classifying objects with these new multiscale Fourier descriptors increases immunity to noise and discrimination power.

6. References

[1] H. Freeman and L.S Davis, "A Corner Finding Algorithm for Chain Coded Curves," *IEEE Transactions on Computers*, March 1977, Vol. 26, pp. 297-303.

[2] G.C.H. Chang and C.C.J. Kuo, "Wavelet Descriptor of Planer Curves: Theory and Applications," *IEEE Trans. on Image Processing*, January 1996, Vol. 5, pp. 56-70.

[3] F. Mokhtarian and A.K. Mackworth, "A Theory of Multiscale, Curvature-based Shape Representation for Planer Curves," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992, Vol. 14, pp. 789-805.

[4] M. S. Nixon and A. Aguado, "*Feature Extraction and Image Processing*", 2nd Edition, Elsevier, 2007.

[5] I. Kunttu, L. Lepisto, J. Rauhamaa and A. Visa, "Multiscale Fourier Descriptor for Shape Classification," *IEEE International Conference on Image Analysis and Processing*, Mantova, Italy, September 2003, pp. 536-541.

[6] W.Y. Kim and Y.S Kim, "A Region-based Shape Descriptor using Zernike moments", *Signal Processing: Image Communication*, September 2000, Vol. 16, No:1, pp. 95-102.

[7] D.S. Zhang and G. Lu, "Generic Fourier Descriptor for Shape-based Image Retrieval", *IEEE International Conference on Multimedia and Expo*, Lusanne, Switzerland, August 2002, Vol. 1, pp. 425-428.

[8] J. Wood, "Invariant Pattern Recognition: A Review" *Pattern Recognition,* 1996, Vol. 29 , No:1, pp. 1-17.

[9] G. Wolberg, S. Zokai. "Robust Image Registration using Log-polar Transform", *IEEE Int. Conf. on Image Processing*, Vancouver, Canada, September 2000, Vol, 1, pp. 493-496.