Image–Based Multiscale Shape Description Using Gaussian Filter

2 authors, including:

Cem Direkoglu
Middle East Technical University
35 PUBLICATIONS   **302** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Time and Cost-Efficient Vein Pattern Recognition and Injection Point Suggestion using Machine Vision Technology View project

Image-based Multiscale Shape Description using Gaussian Filter

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, a multiscale description provides more information about the object, increases discrimination power and immunity to noise. In this paper, we develop a new multiscale Fourier-based object description in 2-D space using a low-pass Gaussian filter (LPGF) and a high-pass Gaussian filter (HPGF), separately. Using the LPGF, at different scales, represents the inner and central part of an object more than the boundary. On the other hand using the HPGF, at different scales, represents the boundary and exterior parts of an object more than the central part. Our algorithms are also organized to achieve size, translation and rotation invariance. Evaluation indicates that representing the boundary and exterior parts more than the central part using the HPGF performs better than the LPGF based multiscale representation, and in comparison to Zernike moments and elliptic Fourier descriptors 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], autoregressive models [2], wavelet descriptors [3], curvature scale space (CSS) [4], and Fourier descriptors [5]. Common region-based shape descriptors are moments [5][6] and Generic Fourier Descriptors (GFDs) [7].

Multiscale shape description is the most promising approach for recognition. Different features of the shape can be obtained at different scales and the combination of these features can increase

discrimination power, so increasing the correct classification rate. In addition, it is more robust to noise. There are many boundary-based multiscale description techniques [4][8][9][10]. One of the most influential techniques is curvature scale space (CSS) [4]. This method uses the scale space framework [11] in 1-D space. The boundary of a shape is filtered by LPGF at varying scales (standard deviation). At each scale, the locations of those curvatures zero crossings are designated as one and otherwise as zero. Finally, the obtained binary CSS image is used for matching. In [8], Fourier descriptors is combined with the wavelet transform to produce multiscale Fourier descriptors for the boundary of a shape. Boundary-based techniques have some limitations. First, they are generally sensitive to noise and variations of shape, since they only use boundary information. Second, in many cases, the object boundary is not complete with disjoint regions or holes. Although many boundary-based multiscale description techniques exist, there is no region-based multiscale description technique in the image space. It is important to note that moments are multiscale approach in the feature space, but not in the image space.

We develop new multiscale Fourier-based object descriptors in 2-D space. We investigate both the LPGF and the HPGF based multiscale representation separately. The LPGF applies smoothing to the object and as scale (standard deviation) decreases, it causes loss of boundary and exterior regions. Therefore using the LPGF at different scales focuses on the inner and central part more than on the boundary of an object. On the other hand, using the HPGF at different scales emphasizes the boundary and exterior parts of an object more than the central part. Part of this work, which uses HPGF for multiscale generation, has been submitted to [12]. In this paper, we extend our approach with LPGF based multiscale description and compare with the HPGF based multiscale description. Our algorithm is organized to achieve size, translation and rotation invariance. First we apply size normalization of the silhouette image and we then

compute a Fourier magnitude image that is translation invariant. At this stage, a LPGF or a HPGF with a selection of different scale parameters is applied to the Fourier magnitude image and produces different Fourier magnitude images. 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 to convert rotation 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 provides 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, increase immunity to noise as well as increase correct classification rate is observed. Figure 1 shows the proposed algorithm to obtain multiscale Fourier-based object descriptors in 2-D space. Evaluation indicates that the HPGF based multiscale representation performs better than the LPGF based multiscale representation, and in comparison to Zernike moments and elliptic Fourier descriptors with respect to increasing noise.
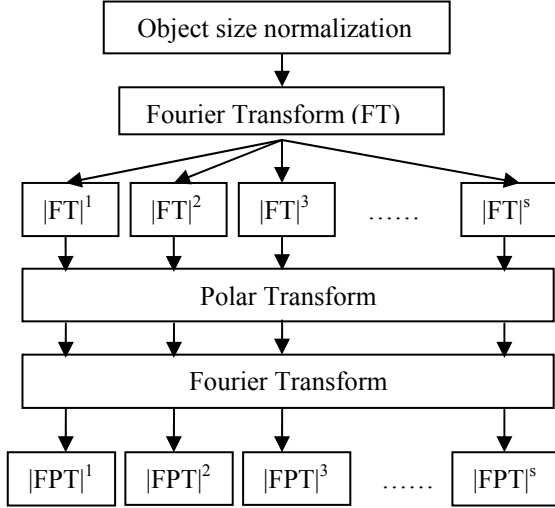

Figure 1. Producing the proposed multiscale Fourier-based object descriptors

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 [13]. The first stage is a 2-D Fourier transform, then the Log-polar transform, and finally another 2-D Fourier transform. The Log-polar transform converts scale and rotation changes to translations. Converting scale change to translation is achieved by logarithmic scaling the radius coordinate of the polar map image [14]. The difference from our new approach is that, we have a filtering approach to create 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.

The rest of the paper is organized as follows: Section 2 explains the proposed multiscale Fourier-based object description with respect to LPGF and HPGF in 2-D space. Section 3 represents classification with multiscale Fourier-based object description. Section 4 concerns evaluation and experimental results and finally Section 5 is conclusions.

2. Fourier-based description with multiscale representation in 2-D space

We produce multiscale Fourier-based object descriptors in 2-D space. For this purpose, we investigate the LPGF and the HPGF, separately. 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 the 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. Multiscale generation is achieved at this stage. To represent the inner and central part of an object more than the boundary, a LPGF with a selection of scale parameters (standard deviation) is applied to the Fourier magnitude image as shown in Equation 2.

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

Where $|FT(u,v)|^s$ and σ_s are respectively Fourier magnitude and scale parameter of scale index s. This method is generating the scale space [11] of the object

in 2-D as shown in Figure 2. It is observed that the LPGF smooths object region and as scale decreases, it causes loss of the boundary and exterior regions. The LPGF emphasizes lower frequency components, but also allows some contribution of higher frequency components.

On the other hand, to represent more the boundary and exterior parts of an object than the central part, a HPGF with a selection of scale parameters (standard deviation) is similarly applied to the Fourier magnitude image as shown in Equation 3.

$$|FT(u,v)|^s = |FT(u,v)|\left(1 - e^{-\sqrt{u^2+v^2}/2\sigma_s^2}\right) \qquad (3)$$

Where $|FT(u,v)|^s$ and σ_s are Fourier magnitude and scale parameter of scale index s, respectively. Filtering with the HPGF at different scales is illustrated in Figure 3. It is observed that the HPGF detects the object boundary and as scale decreases, it represents exterior regions. The HPGF emphasizes higher frequency components, but also allows contribution of lower frequency components slightly.

The obtained Fourier magnitude images are not convenient for matching at this stage, since they still vary with rotation. 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

$$(x,y) = (r * \cos\theta, r * \sin\theta) \qquad (4)$$

Where $r \in \Re$ and $0 \leq \theta \leq 2\pi$. For every point (x,y), there is a unique point (r,θ). Finally, another 2-D Fourier transform is applied, as given in Equation 5, to compute Fourier magnitude, which removes these translations.

$$FPT^s(k,l) = \frac{1}{EF} \sum_{r=0}^{E-1} \sum_{\theta=0}^{F-1} P^s(r,\theta) e^{[-j2\pi(kr/E + l\theta/F)]} \quad (5)$$

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 a shape at scale index s.

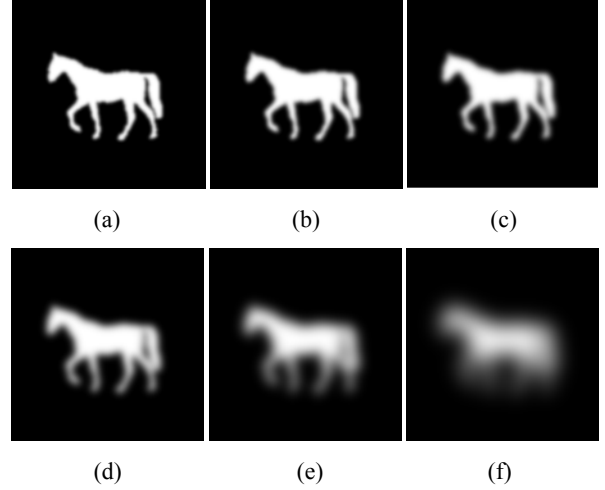(a)　　　　　　(b)　　　　　　(c)

(d)　　　　　　(e)　　　　　　(f)

Figure 2. Horse object filtered by LPGF with respect to decreasing scale. (a) $\sigma_1 = 20$, (b) $\sigma_2 = 15$, (c) $\sigma_3 = 11$, (d) $\sigma_4 = 8$, (e) $\sigma_5 = 5$, (f) $\sigma_6 = 3$
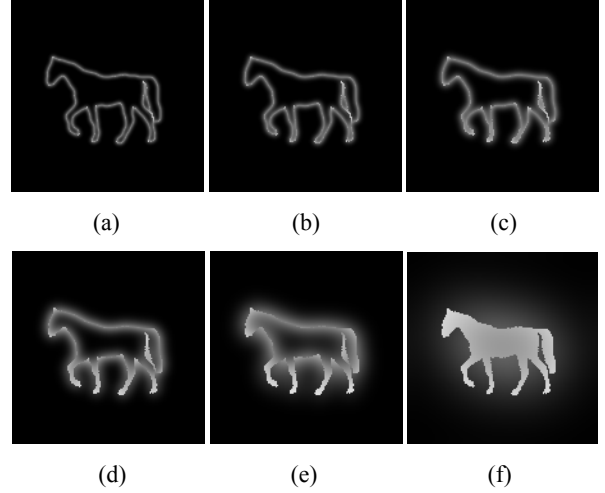

(a)　　　　　　(b)　　　　　　(c)

(d)　　　　　　(e)　　　　　　(f)

Figure 3. Horse object filtered by HPGF 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 description

The classification is done 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,

$$Ed^s(T,D) = \sqrt{\sum_{i=1}^{E} \sum_{j=1}^{F} \left(OD_T^s(i,j) - OD_D^s(i,j)\right)^2} \quad (6)$$

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 (Ad) is computed for each object.

$$Ad = \frac{1}{B}\sum_{s=1}^{B} Ed^s \qquad (7)$$

Where Ad represents average distance and B is the number of scales. Classifying with average distance, instead of single scale distance, increase correct classification as well as increase 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 and are in binary form. The objects are divided into 70 classes with 20 images in each class. The appearance of these silhouettes changes due to view point with respect to objects (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 4. Leave-one-out cross-validation is applied to validate classification. Performance evaluation is employed by comparing the multiscale description using LPGF and multiscale description using HPGF with each other as well as with elliptic Fourier descriptors (EFD) and Zernike moments (ZM). The evaluation is achieved with respect to increasing salt and pepper noise. Figure 5 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 fast and robust boundary-based shape descriptors. The contour is represented with complex coordinates and then the Fourier expansion is performed to obtain the EFD. To evaluate EFD, we use the algorithm given in [5]. 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 multiscale description using LPGF, 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: $\sigma_1 = 20$, $\sigma_2 = 15$, $\sigma_3 = 11$, $\sigma_4 = 8$ and $\sigma_5 = 5$. The size of the object descriptor matrix is 90×90 at each scale.

In multiscale description using HPGF, similarly 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. Table 1 and Figure 6 show the correct classification rate (CCR%) of the multiscale description in 2-D using LPGF, of multiscale description in 2-D using HPGF, of ZM and of EFD, with respect to increasing density (D) of salt and pepper noise. CCR is measured as follows,

$$CCR(\%) = \frac{c}{t}\times100 \qquad (8)$$

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

It is observed that HPGF based multiscale description performs better than LPGF based multiscale description, ZM and EFD. HPGF based multiscale description achieves 95.5% correct classification rate, while LPGF based multiscale description achieves 90.4%, ZM achieves 90% and EFD achieves 82% without adding noise to the database. As noise increases, the performance of all algorithms decreases and their performances degrade similarly. It is also observed that LPGF based multiscale description and ZM have very close performances. The success of HPGF based multiscale description in 2-D appears due to emphasizing the boundary and exterior parts of objects and also allowing the central part contribute slightly to classification. Although we use a multiscale representation (LPGF or HPGF), 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 the Fast Fourier Transform (FFT).

5. Conclusion

We have presented novel multiscale Fourier-based object descriptors in 2-D space for the purpose of

recognition. We have used a low-pass Gaussian filter (LPGF) and a high-pass Gaussian filter (HPGF) for multiscale generation. Our algorithm also makes descriptors size, translation and rotation invariant. Multiscale description using HPGF, which represents the boundary and exterior parts of an object more than the central part, outperforms multiscale description using LPGF, Zernike moments (ZM) and elliptic Fourier descriptors (EFD) with respect to increasing salt and pepper noise in the database. Classifying objects with this new multiscale Fourier-based object description using the HPGF in 2-D space increases immunity to noise and discrimination power.
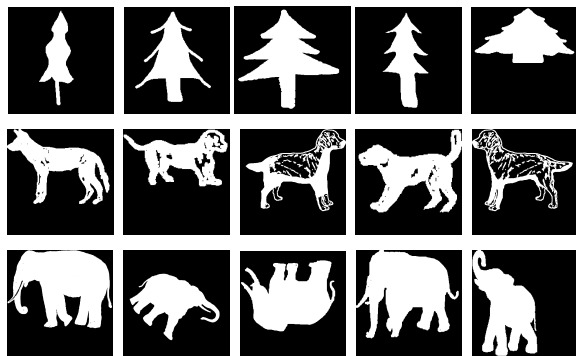

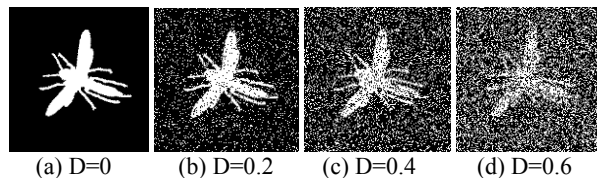**Figure 4**. Some objects from the database and their variations

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

Figure 5. Fly object with increasing density (D) of salt and pepper noise

Table 1. CCR% of HPGF and LPGF based multiscale description in 2-D, ZM and EFD, with respect to increasing density (D) of salt and pepper noise in the database.

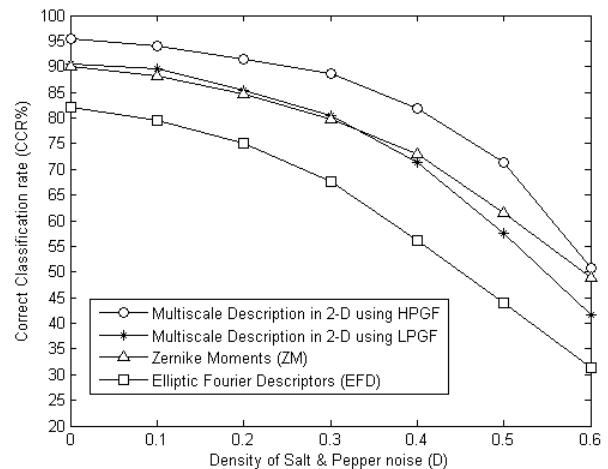| D	Multiscale HPGF	Multiscale LPGF	ZM	EFD
0	95.5 %	90.4 %	90 %	82 %
0.1	93.6 %	89.5 %	88.2 %	79.5 %
0.2	92.2 %	85.2 %	84.7 %	75 %
0.3	88.5 %	80.5 %	79.8 %	67.5 %
0.4	82 %	70.3 %	73 %	56 %
0.5	71.3 %	55 %	61.5 %	44 %
0.6	52 %	41.5 %	48.8 %	31.2 %

Figure 6. Classification performance of LPGF and HPGF based multiscale description in 2-D, ZM and EFD, with respect to increasing density (D) of salt and pepper noise.

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] S.R Dubois and F.H. Glanz, "An Autoregressive Model Approach to Two-Dimensional Shape Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, January 1986, Vol. 8, pp. 55-66.

[3] 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.

[4] 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.

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

[6] W.Y. Kim and Y.S Kim, "A Region-based Shape Descriptor using Zernike moments", *Signal Processing: Image Comm.*, 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] 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.

[9] T. Adamek and N.E. O'Connor, "A Multiscale Representation Method for Nonrigid Shapes with a Single Closed Contour," *IEEE Trans. on Circuit and Systems for Video Technology*, May 2004, Vol. 14, No: 5, pp. 742-753.

[10] R.S. Torres, A.X. Falcao and L.F. Costa, "A Graph-based Approach for Multiscale Shape Analysis", *Pattern Recognition,* June 2004, Vol. 37, No:6, pp. 1163-1174.

[11] A. Witkin, "Scale-space Filtering", *Int. Joint Conf. Art. Intell.*, Karlsruhe, Germany, 1983, pp.1019-1022

[12] Cem Direkoglu and Mark S. Nixon, "Shape Classification using Multiscale Fourier-based Description in 2-D space", accepted by *IEEE International Conference on Signal Processing*, Beijing, China, 2008.

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

[14] 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.