Firm Object Classification using Butterworth Filters and Multiscale Fourier Descriptors

Saravanakumar M
Department of Information Technology,
Bannari Amman Institute of Technology,
Sathyamangalam, Tamilnadu, India

VinothSarun S
Department of Information Technology,
Bannari Amman Institute of Technology,
Sathyamangalam, Tamilnadu, India

Sundaramurthy S
Department of Information Technology,
Bannari Amman Institute of Technology,
Sathyamangalam, Tamilnadu, India

Amitabh Wahi
Department of Information Technology,
Bannari Amman Institute of Technology,
Sathyamangalam, Tamilnadu, India

ABSTRACT
Object classification is an important task in computer vision techniques. In this paper we proposed a region based object classification technique using multiscale Fourier descriptors and Butterworth filters. By using the Butterworth Low pass Filters (BLF) and Butterworth High pass Filters (BHF) at varied scales. While applying the Butterworth Low pass Filter (BLF) at varied parameters, the internal region of the image is more concentrated which results in the smoothed image. Similarly while applying the Butterworth High pass filter (BHF) at varied parameters, the external region of the image is more concentrated which results in the sharpened image. The proposed algorithm is capable of eliminating the variation in size, rotation and translation of the object. The evaluation outcomes show that filtering done with BHF outperforms the BLF and shows better classification results.

Keywords
Shape classifications; Butterworth High pass Filters; Butterworth Low pass Filters; Fourier descriptors; Euclidean distance

1. INTRODUCTION
Object classification and recognition is an important task in computer vision. Humans are capable of correctly identifying the object with translation variation, size variation and also in occluded conditions. In case of computer vision it is a challenging task for identifying and classifying the image devising these complications. There are two widely used methods for identifying the shape descriptors which are descriptors based on boundary of the images and descriptors based on region of the images.

1.1 Boundary based methods
In the boundary based object description the edge of the object is taken and the descriptors are computed. The major drawbacks while using boundary based technique is the object are sensitive to noise and occurrence of incomplete boundaries. The widely recognized boundary based shape descriptors are Fourier descriptors [1], complex coordinates[2], centroid distance[3], curvature signature[4], area function [5] and shape signatures [6,7]. The reason behind using the Fourier descriptors[8] are easy to calculate and more robust. By taking the Fourier transform for a given shape signature Fourier descriptors are obtained. The shape signature is a 1-D function which is obtained from the 2-D boundary of the image. Elimination of geometric variance can be achieved after the shape signature extraction. Computed Fourier transform is having the lower frequency parameters and higher frequency parameters. The general information about the image is enclosed in the lower frequency descriptors and the finer information is enclosed in the higher frequency descriptors.

1.2Region based methods
In the region based methods [9], all the pixels inside the object are considered for shape representation. The widely used region based shape descriptors are moments [10, 11] and Generic Fourier descriptors(GFD’s) [12]. The moments are based upon two types called orthogonal and non-orthogonal moments. For character recognition, license plate recognition geometric moments are used. Low-order moments are used whereas more information is needed. As the order increases the details also increased. Obtaining higher information shows lots of data redundancy as a result of using non-orthogonal basis function.

2. BACKGROUND
2.1 Zernike and Legendre moments
In 1980, Teague[13] proposed Legendre moments that uses Legendre polynomials as its basic function. The polynomials are orthogonal polynomials and uses Legendre moments to extract self-determining features of the object. There is no data redundancy occurs while using Legendre moments. So these methods are used in reconstruction of the image. Zernike moments [14]are initially proposed by Teague[13] which is based on Zernike polynomials. To eliminate the rotational variance in the image Zernike moment[15] are used and it is perfectly resilience to noise and reconstruction features.

2.2 Generic Fourier Descriptors (GFD)
Generic Fourier descriptors [12]can be realistically applicable to general applications. 2-D Fourier transform is applied to the polar shape of the object for extracting the spectral domain. The applications of GFD’s are widely noted in Content Based Image Retrieval (CBIR) [16] systems.

2.3 Multiscale Fourier Descriptors
CemDirekogluand Mark S.Nixon [17, 18] introduce multiscaleFourier based object description using Low Pass Gaussian Filters and High Pass Gaussian Filters[19]. This method is a region based object recognition technique which

performs well than using wavelet based multiscale Fourier descriptors [20]. Since there are many boundary based multiscale object recognition techniques exists and slightest techniques available for region based multiscale techniques.

3. PROPOSED METHOD

Shape classification is an important feature and it is a major part of object recognition. Shape representation is not an easy task since the shapes are asymmetrical, translation variance and rotation variance. So it is hard to classify an image according to the similar images of same object. Effectiveness and correct classification percentage is the major concern about shape classification. Here we are using MPEG-7 CE-Shape-1 Part B [21, 22] database for evaluation. MPEG-7 CE-Shape-1 Part B is a most commonly used dataset in shape classification experiments, which are having the silhouette images acquired from real world objects.

Multiscale shape description is the most promising approach for object recognition. The region based multiscale Fourier descriptors using Butterworth Low pass Filter (BLF) and Butterworth High pass Filter (BHF) at various scale index *s*. While applying the Butterworth Low pass Filter (BLF) at varied parameters, the internal region of the image is more concentrated than the external region of the object, which results in the appearance of smoothed image. Similarly while applying the Butterworth High pass Filter (BHF) at varied parameters the external region of the image is more concentrated than the internal region, which results in the appearance of sharpened image. The selected parameters in this work is s_1=1, s_2=5, s_3=10, s_4=15 and s_5=20.

4. PREPROCESSING

Thealgorithm starts with resizing the original imageof size M x N into a fixed dimension of 128x128.

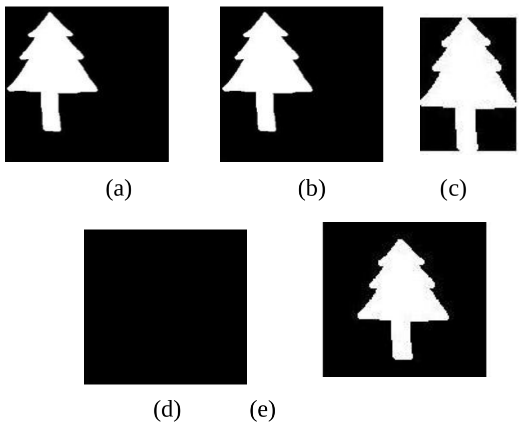


Fig 1: (a)Original input image of size M x N, (b) Binary image, (c) Object cropped image, (d) Blank matrix of size 128X128, (e) Object centralized binary image

In this method we are using bilinear interpolation method for resizing the image. After resizing the image using bilinear interpolation the object in the image is centralized by cropping the unwanted regions surrounding the object. The image is converted into binary form and cropping is done with bounding box method. The bounding box method finds the first white pixel and last white pixel along horizontal axis and similarly by finding the top and bottom white pixel along vertical axis of the object. After finding the coordinates of the pixels, the object is cropped and placed center in an empty zero matrix of 128x128 dimensions. We are using this method to eliminate the scale variance as shown in Fig. 1. Before cropping the image we use region fill method to fill the open areas present in the image.

The next step is to apply Fourier transformto the cropped image *I(x,y)* for eliminating the translation variance from image.

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

Where *FT(u,v)* is the obtained Fourier transform of the centralized binary image *I(x,y)* of size MN. The Fourier transform renovates the image from spatial domain to frequency domain. The obtained result of Fourier transform is in the form of complex numbers. So the complex numbers representsthe phase and magnitude. The phase represents the timing and the magnitude represents frequency of the image*I(x,y)*. The computed Fourier magnitude is translation invariance.

5.FILTERING OF OBJECTS

5.1 Filtering with Butterworth Low pass Filter (BLF)

In Frequency domain the image is filtered using frequency filters. Most widely the frequency filters are used in image reconstruction. Here we are using Butterworth Low pass Filters and High pass Filters separately. Butterworth Low pass Filters are used to smoothing the image. While applying Butterworth low pass filtersto the magnitude of the image will allow only the low frequency components.[

$$|FT(u,v)|^s = \frac{1}{1+[FT(u,v)/\alpha_s]^{2n}} \tag{2}$$

Where *|FT(u,v)|* is the computed Fourier magnitude and α_s is the scale parameters. So the result of this filtering removes all the high frequency components and enhances only the lower frequency components as per the scale parameters. As a result after applying BLF the internal region of the object is more concentrated and appears to be smoothed image which is shown in Fig. 2.

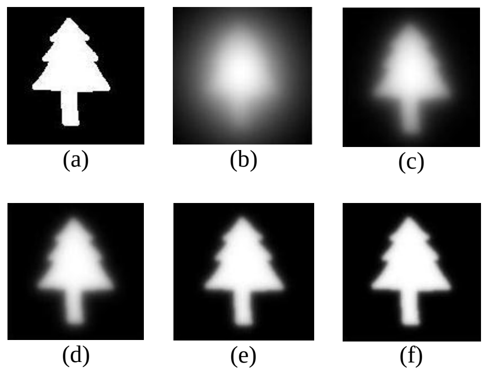


Fig 2: Object filtered by using BLF at various scale index (a)Original image, (b) s_1=1, (c) s_2=5, (d) s_3=10, (e) s_4=15 and (f) s_5=20

5.2 Filtering with Butterworth High pass Filter (BHF)

Likewise to represent the external region of the object we apply the Butterworth High pass Filter as shown in (3). The lower frequency components are removed and higher frequency components are enhanced as per the scale parameters.

$$|FT(u,v)|^s = \frac{1}{1+[\alpha_s / FT(u,v)]^{2n}} \quad (3)$$

The resultant image is reconstructed by taking an inverse Fourier transform. The obtained image at various scale parameters are shown in Fig. 3.

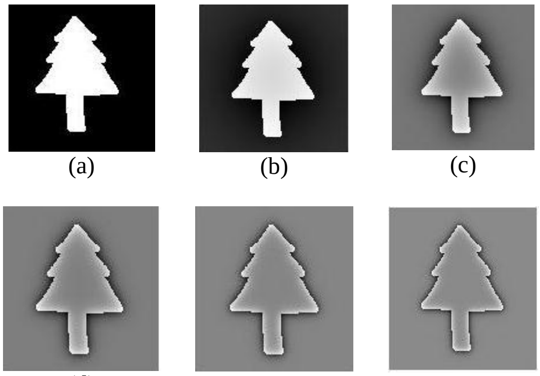
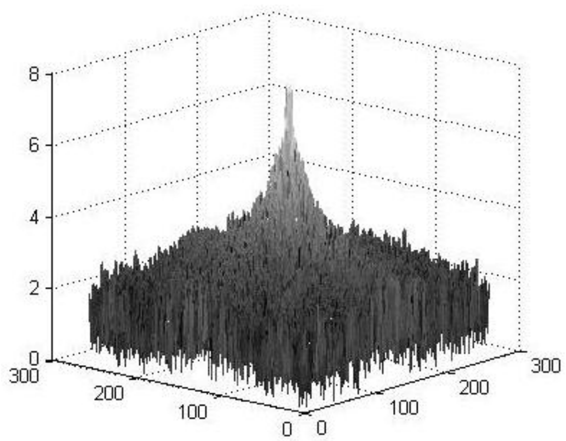


Fig 3: Object filtered by using BHF at various scale index (a)Original image, (b) s_1=1, (c) s_2=5, (d) s_3=10, (e) s_4=15 and (f) s_5=20

The image shows the external region is more concentrated than the internal region. The sharpening of the object increases as the scale parameter increases. In some cases, if noise is present in the object, this high pass filter will also sharpen the noise and the object information is loosed. The computed Fourier magnitude is still having rotation variance. So it is not suitable for matching.

6. CARTESIAN TO POLAR TRANSFORM

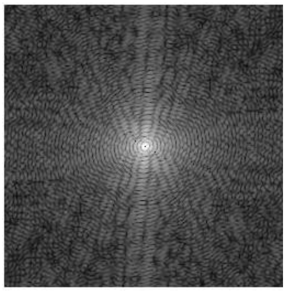
The rotation variance is the major problem in classifying the image since similar images are separated by rotation variance. The obtained Fourier magnitude is in the Cartesian form which is shown in Fig. 4. (a) (b). So the image is transformed into polar mapped by fitting the obtained Fourier magnitude inside an empty circle.



(a)

Fig 4: (a) Obtained Fourier Magnitude of the object

For mapping the image to circle two different approaches are used. Placing the square image inside the circle and another method is placing the square image outside the circle. In our method we followed the first method, placing the obtained magnitude of the image inside the circle shown in Fig. 5.



(b)

Fig 3: (b) Obtained Fourier Magnitude of the object

For mapping the image to circle two different approaches are used. Placing the square image inside the circle and another method is placing the square image outside the circle. In our method we followed the first method, placing the obtained magnitude of the image inside the circle shown in Fig. 5.

Our final magnitude image Fig. 4. (b)is having 128x128dimensions. To place an image inside the circle, construct an empty circle of radius *r*as follows

$$r = \sqrt{x^2 + y^2} \quad (4)$$

$$\theta = \tan^{-1}(y/x) \quad (5)$$

Where *x*and y are the dimensions of the magnitude image.By using the r and θ generate a circle and padding is done before place the magnitude image inside the circle shown in Fig. 5.

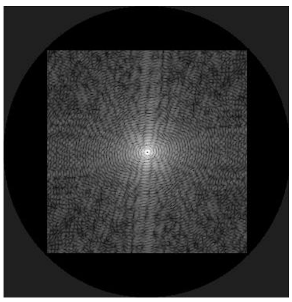


Fig 5: Cartesian form of the magnitude image which is placed inside the circle.

For each point *(x,y)* in the image*I(x,y)* is having unique point *(r, θ).* Converting the image to polar form can be represented by the following matrix.

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos\partial & -\sin\partial \\ \sin\partial & \cos\partial \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad (6)$$

Where (x_1, y_1) the point is in Cartesian form (x_2, y_2) is the point in polar form and θ is the angle of rotation.
Consider $(x_1,y_1)=(r\cos\theta, r\sin\theta)$ and substitute in (6) we get the co-ordinates as$x_2=r\cos(\theta+\partial)$ and $y_2=r\sin(\theta+\partial)$. So when a rotation occurs in Cartesian coordinate will cause translation in polar coordinates.

$$(x_1,y_1)\leftrightarrow(r, \theta) \quad (7)$$
$$(x_2,y_2)\leftrightarrow(r, \theta+\partial) \quad (8)$$

The Cartesian form of the magnitude image is now converted into polar coordinates (7) (8) as shown in Fig. 6.The null pixels in the polar image are considered as zeros and having size of 92x92 dimensions.

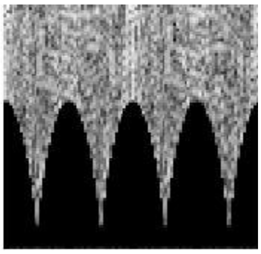


Fig 7: Polar form of the Cartesian image placed within the circle.

Finally to remove the rotation variance another 2-D Fourier transform (1) is taken and the computed resultant magnitude of the image shown in Fig. 7.is invariance to size, translation and rotation. After this the object is ready for matching and classifying.

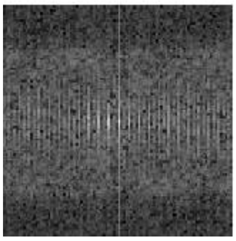


Fig 8: Resultant magnitude of the image

7. CLASSIFICATION

For classification, we are using the minimum distance classifier based on the Euclidean distance(Ed) method. Because our obtained resultant magnitude is having 8464 values and to classify these values statistical classifiers are used widely for better classification. Euclidean distance is used to measure the resemblance between the objects and the minimum distance is obtained.

$$Ed(T,D) = \sqrt{\sum_{x=1}^{C}\sum_{y=1}^{E}(RD_T(X,Y) - RD_D(X,Y))^2} \quad (9)$$

Where $Ed(T,D)$ is the Euclidean distance between test image T and resultant descriptors RD present in the database D respectively.

8. EVALUATION AND RESULTS

For this work,we are using MPEG-7 CE-Shape-1 Part B [34] database for evaluation. MPEG-7 CE-Shape-1 Part B is a most commonly used dataset in shape classification experiments, which are having 1400 silhouette images acquired from real world objects which are represented in binary form. The 1400 images are classified into 70 classes with 20 images in each class. The experiment is carried out in MATLAB R2010a running on Core i5 processor.

The correct classification percentage (CCP) is obtained by the following formula.

$$CCP = \frac{c_0}{t_0} \text{ x } 100 \quad (10)$$

Where$c_0$is thecorrectly classified object and t_0 is the total number of objects in the database. In our evaluation we are using selected scale index s_1=1, s_2=5, s_3=10, s_4=15 and s_5=20 and apply these scales for BLF and BHF separately. The images present in the MPEG-7 CE-Shape-1 Part B database is already existed with noise. To remove these noises we apply morphological flood-fill operation to fill the regions to get silhouette images.

Table 1.Correct Classification Percentage.

S. No	Varied Scale Index	Butterworth Low pass Filter (BLF)	Butterworth High pass Filter (BHF)
1	S=1	43.2	82.3
2	S=5	55.4	97.2
3	S=10	68.4	86.3
4	S=15	91.2	74.6
5	S=20	82.5	64.3

For this experiment the 1400 images present in the database is grouped in to 70 classes with 20 objects in per class. Among these 1400 pre-segmented images many objects are geometrically different. Each and every object is different and having variance in rotation, translation and shape. So this database is more complex to classify. By using this region based multiscale Fourier descriptor the classification is done with more efficiency and accuracy. For the testing we take 15% and remaining 85% in the database for comparison. Table 1 shows the correct classification results of this method. Here there are five selected scale index are used for BLF and BHF. The experiment is carried out by filtering the object with BLF with each selected scale values and similarly for the BHF with each selected scale values. For each class the object is computed for the resultant descriptors and the descriptors are compared with all the other class objects by using the Euclidean minimum distance classifier. The minimum distances between the test objects and the object present in the database are computed and the classification rate is calculated as per the CCP%.

Table 1 shows the selected scale parameters and the correct classification percentage between the BLF and BHF. The selected scale parameters ares_1=1, s_2=5, s_3=10, s_4=15 and s_5=20. The size of the resultant magnitude is 92x92. As from the table 1 it shows that while using the BLF the classification percentage increases as the scale index increase. It also shows that the poor classification rate at scale parameters s_1=1. This happens because the object gets more smoothed and the details preset in that object are not observed. As the scale value increases the classification percentage also increases. But at the scale index s_5=20 the percentage starts to decreases. This happens because from this scale the object is represented by region itself. The best classification percentage 91.2% which is obtained at scale *S=15.*

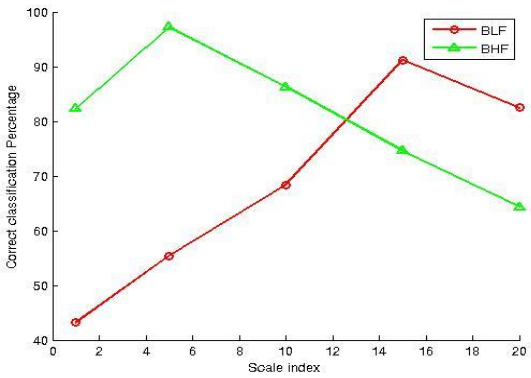


Fig 9: Comparison between BLF and BHF Correct Classification Percentage.

Similarly using the BHF shows better classification percentage when compared with BLF shown in Fig 9. The best classification percentage 97.2% is achieved at scale *S=5.*This rate is achieved because using BHF focusing on the external region is more than the internal region. The data gathered is comprises of both the external boundary data and some of the internal region data. As the scale increases not like BLF the BHF starts to decrease in CCP. This shows that using the Butterworth High pass Filter is the most efficient way of classifying the objects based upon the region.

9. CONCLUSION

The region based multiscale Fourier descriptors using Butterworth filters shows an efficient method in classifying the objects and this method also shows that this approach very promising towards shape variance, translation variance and rotation variance. This method uses both Butterworth Low pass Filters (BLF) and Butterworth High pass Filters (BHF). Among these two BHF shows better classification rate of 97.2%. Here we are using Euclidean distance for classifications since there is huge number of values are computed during the resultant magnitude. In future work, improvements are made towards reducing the number of values in resultant matrix and different classifiers are used in order to increase the accuracy.

7. REFERENCES

[1] E. Persoon, K. Fu, "Shape discrimination using Fourier descriptors",IEEE Transactions on Systems, Man and Cybernetics (1977) 170–179.

[2] D. S. Zhang, G. Lu, "A comparative study of Fourier descriptors for shape representation and retrieval", Asian Conference on Computer Vision (2002) 646–651.

[3] H. Ling, D. W. Jacobs, "Shape classification using the inner-distance", IEEE Transactions on Pattern Analysis and Machine Intelligence (2007) 286–299.

[4] H. Kauppinen, TapioSeppanen and M. Pietikainen, "An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D Shape Classification". IEEE Trans. 201-207

[5] I. E. Rube, N. Alajlan, M. Kamel, M. Ahmed, G. Freeman,"Robust multiscale trianglearea representation for 2-D shapes", IEEE International Conference on Image Processing (2005) 545–548.

[6] I. Yong, J. Walker and J. Bowie. "An Analysis Technique for Biological Shape". Computer Graphics and Image Processing, Vol.25, pp.357-370, 1974.

[7] Peter J. van Otterloo. "A contour Oriented Approach to Shape Analysis". Prentice Hall International (UK) Ltd. C1991.

[8] C. T. Zahn, R. Z. Roskies, "Fourier descriptors for plane close curves", IEEE Transactions on Computers (1972) 269–281.

[9] M. S. Nixon, A. Aguado, "Feature Extraction and Image Processing", 2nd ed., Elsevier, 2007.

[10] M. K. Hu, "Visual pattern recognition by moment invariants", IRE Transactions on Information Theory (1962)179–187.

[11] R. J. Prokop, A. P. Reeves, "A survey of moment-based techniques for un-occluded object representation and recognition", CVGIP: Graphical Models and Image Processing (1992)438–460.

[12] D. S. Zhang, G. Lu, "Generic Fourier descriptor for shape-based image retrieval", IEEE International Conference on Multimedia and Expo (2002) 425–428.

[13] M. R. Teague, "Image analysis via the general theory of moments", Journal of the Optical Society of America (1979) 920–930.

[14] W. Y. Kim and Y. S. Kim. "A Region-based Shape Descriptor Using Zernike Moments", Signal Processing & Image Communication, 95-102, 2000.

[15] Michael Reed Teague. "Image Analysis via the General theory of Moments", Journal of Optical Society of America, 920-930, 1980.

[16] T. Kato. "Database Architecture for Content-based Image Retrieval in Image Storage and Retrieval Systems", Proc SPIE 1662, pp.112-123, 1992.

[17] C. Direkoglu, M. S. Nixon, "Image based multiscale shape description using Gaussian filter", IEEE Indian Conference on Computer Vision, Graphics and Image Processing (2008)673–678.

[18] C. Direkoglu, M. S. Nixon, "Shape classification using multiscale Fourier-based description in 2-D space", IEEE International Conference on Signal Processing (2008)820–823.

[19] CemDirekoglu and Mark S. Nixon, "Shape classification via image-based multi-scale description", Elsevier: Pattern Recognition letters, Vol.44, No.9, 2134-2146, Sep- 2011.

[20] T. Adamek, N.E.O Connor, "A multiscale representation method for non-rigid shapes with a single closed contour", IEEE Transactions on Circuit and Systems for Video Technology (2004)742–753

[21] L. J. Latecki, R. Lakamper, U. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contour", IEEE International Conference on Computer Vision and Pattern Recognition (2000)424–429.

[22] http://www.imageprocessingplace.com/root_files_V3/image_databases.htm