International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 6 (2018) pp. 3321-3326

Rotation-invariant Texture Classification by Spectral Transformation of Gabor Filter Features

Gouchol Pok

Division of Computer and IT Education,
Pai Chai University, Daejeon, Republic of Korea.

Abstract

Gabor filters have been proved to be powerful for texture analysis. Gabor filters are constructed as a filter bank in order to capture local textural properties characterized by different frequencies and spatial orientations. Conventionally, the set of power spectrum obtained by Gabor filters are used as texture features. This simple approach has drawbacks that, among others, fine tuning of parameters is required for optimal setting and good performance, and Gabor filter itself does not provide rotation-invariant texture classification scheme. In this paper, we propose a rotation-invariant texture classification method in which texture features are obtained with two processing stages: first, Gabor features are computed in a conventional way, and then Fourier transform is applied to the Gabor features to obtain rotation-invariant texture features. The resulting texture features allow for researchers to select outstanding components of textures and to visually inspect textural characteristics. The proposed method was compared with other methods and showed superior classification results.

Keywords: texture classification, rotation-invariant texture descriptor, Gabor filter, Fourier descriptor

INTRODUCTION

Texture analysis has played an important role in a lot of computer vision and image processing applications [1]. Methods of texture analysis can be divided into four categories: statistical, geometrical, model-based and signal processing types [2].

Statistical methods consider local features from the spatial distribution of gray level intensities and compute a number of statistics from the distributions of the local features. According to the number of pixels defining the local features, statistics are further divided into first-order (one pixel), second-order (two pixels) and higher-order (three or more pixels) statistics [3]. First-order statistics include average and variance, and thus compute properties of individual pixels and ignore the spatial relationship between pixels. Second- order and higher-order statistics consider correlative properties of two or more pixels occurring at specific locations.

Statistical methods include cooccurrence matrix [4], gray level differences [5], signed differences of grey-level intensities [6], and the Local Binary Pattern (LBP) operator [7]. LBP operator has shown superior performance in texture classification and face recognition [8] by combining statistical and structural properties within local areas.

Geometrical methods try to describe texture as a collection of texture primitives called textons and find the rules governing the spatial arrangement of the texture primitives. The texture primitives can be obtained by edge detection with a Laplacian-of-Gaussian filter [9] or by adaptive region extraction [10]. The texture primitives typically include intensity, area, elongation, and orientation. With these primitives, the texture analysis proceeds to compute statistics of the primitives. The structure and organization of the texture primitives can be considered using Voronoi tessellations [11].

Model-based methods attempt to hypothesize the underlying texture process by use of generative image model and stochastic model [12]. The parameters of the image model are estimated and then used as texture features. Stochastic spatial interaction models consider the intensity generation process as a stochastic process where the observed intensity is modeled as a linear combination of intensities and additive noise. Depending on different neighborhood systems and noise sources, various models can be constructed. Autoregressive (AR), moving-average (MA), and autoregressive-moving-average (ARMA) models represent statistical relationships of pixels along raster scanning direction. Random field models consider the spatial correlations of pixels in 2-D space. Global random field models view the entire image as a realization of a random field, while local random field models consider local relationships of pixels in local neighborhoods. A Gibbs random field model is a representative example of the global model [13], and Markov random field model [14] is the representative local random field model.

Signal processing methods are focused on the frequency domain properties of textures. Spatial domain filters include Laws masks [15], local linear transforms [16], and various edge detection masks [17]. Authentic frequency analysis can be conducted by the Fourier transform (FT) because FT is able to describe the global frequency content of the texture as a set of filter coefficients. However, FT does not consider localization of frequency content in the spatial domain, and hence result in poor performance for some applications. In order to remedy this drawback, a window function can be imposed on the input signal to form the short-time Fourier transform. Wavelet transform can achieve multi-resolution analysis of textures by changing widow width as the frequency changes [18]. If the Gaussian function is employed as the window function, the transform is called the Gabor transform [19]. A 2-D Gabor filter consists of filter bank each of which is sensitive to a particular frequency and orientation. Texture features are obtained by filtering the texture with a bank of Gabor filters, each filter

having a specific frequency and orientation. Typically a lot of scales and orientations are involved, and hence Gabor filtering results in high-dimensional texture features. Dimension reduction method can be applied by taking only those bands with high energy [20], or by optimizing filter design to cover the desired frequency range [21].

In this paper, we first employ a Gabor filter to extract the spatial frequency domain property of the input textures and to form texture features. Then, Fourier transform is applied to the texture features to obtain the rotation-invariant features of the input texture. The coefficients of the Fourier transform are called Fourier descriptors, and rotation invariance property is achieved by ignoring the phase information and by taking only the magnitude values of the Fourier descriptors. The overall process of the proposed method can be summarized as illustrated in Figure 1.

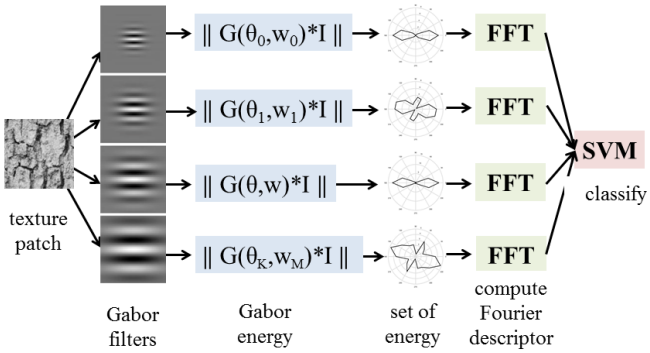

Figure 1. System flow of the proposed method

GABOR FILTER BANK

Gabor filters are motivated by the mathematical and the biological properties of Gaussian functions and useful for texture features extraction. The Gabor functions form a complete, non-orthogonal basis which allows any given function to be expanded in terms of these basis functions. Daugman expanded the Gabor function to 2-D form to model the receptive-field profiles of simple cells in the striate cortex [22]. A Gabor function in the spatial domain is a sinusoidal-modulated Gaussian function. The real impulse response of the 2-D Gabor filter is given by,

$$G(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e\left\{\frac{-1}{2}\left[\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right]\right\}\cdot\cos(2\pi\mu_0 x) \quad (1)$$

where σ_x and σ_y are the spreads in the x and y directions, respectively, and μ_0 is a modulating frequency. The 2-D Gabor impulse response function is illustrated in Figure 2.

A Gabor filter bank is constructed with varying values of frequency μ_0 and rotation angle θ of the impulse response function, which define the center location of the filter in the spatial frequency domain. We can obtain arbitrary Gabor filter by tuning frequency μ_0 and rotation angle θ so that the resulting filters cover the spatial-frequency domain.

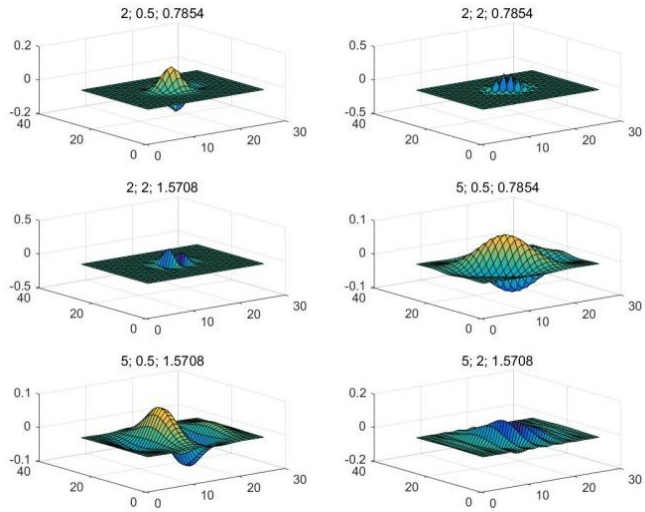

Figure 2. Real part of impulse response functions of a Gabor filter bank. Numbers refer to scale, orientation, and frequency, respectively.

In addition to the frequency and rotation angle, fine tuning of a Gabor filter involves an optimal setting of the frequency bandwidth B_F and the orientation bandwidth B_θ. As 2-D Gabor functions formulated as models of the receptive fields of simple cells, neurophysiological research has reported that the half-response spatial-frequency bandwidths of simple cells fall in the range of 0.4 to 2.6 octaves for mammals [23]. In practice, a frequency bandwidth from frequency f_1 to f_2 is set to one octave which is defined as $log_2(f_1/f_2)$ so that the frequency bandwidth increases in a logarithmic fashion. For the orientation angle, a bandwidth of 30° is recommended for good performance [24], while other researchers proposed 45° [20]. Finer quantization of orientation angle expectedly results in better performance with sacrifice of computational time. Figure 3 shows the spatial-frequency domain covered by the Gabor filters which are apart with 45° angular orientation from each other and with one octave of radial frequency. Texture features can be defined as the magnitude of the energy computed by convolving Gabor filter with texture patch,

$$E_{\theta,\omega} = ||G(\theta,\omega) * I|| \quad (2)$$

where θ refers to angular orientation and ω radial frequency.

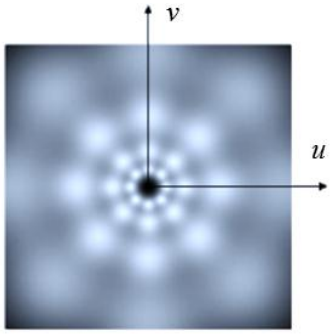

Figure 3. Spatial-frequency domain covered by Gabor filters with 45° apart in angular orientation and radial frequencies 1 octave apart.

SHAPE SIGNATURE

Before processing shapes, the shape at hand first needs to be represented as a 1-D function, called a *shape signature* which describes the boundary of the shape. Among many different shape signatures, the most popular ones include centroid distance, complex coordinates, curvature function, and cumulative angles [25].

1. Centroid distance

The centroid distance is measured by a function of the distance between the boundary points and the centroid (x_c, y_c) of the shape,

$$s(t) = \sqrt{(x(t) - x_c)^2 + (y(t) - y_c)^2} \quad (3)$$

The advantage of the centroid distance signature is that the Fourier descriptors computed from the distance are robust against noise [26].

2. Complex coordinates

Complex coordinates refer to the complex numbers generated from the boundary points,

$$z(t) = x(t) + iy(t) \quad (4)$$

This representation is dependent on the starting point on the boundary, and hence it needs to eliminate the effect, the following shifted coordinates are used to achieve translation invariance,

$$z(t) = [x(t) - x_c] + i[y(t) - y_c] \quad (5)$$

where (x_c, y_c) is the centroid of the shape.

3. Curvature signature

Curvature is the second derivative of the boundary function, or the first derivative of the boundary tangent function. The curvature function is defined as the differentiation of successive angles calculated in a window *w*,

$$k(t) = \theta(t) - \theta(t-1) \quad (6)$$

where

$$\theta(t) = tan^{-1}\ \frac{y(t)-y(t-w)}{x(t)-x(t-w)} \quad (7)$$

4. Cumulative angular function

The cumulative angular function at a boundary point is defined as the amount of angular change from the starting position. Angular change is defined as the derivative of the angular function $\varphi(t)$, which is actually the curvature described in Equation (6) above. Because the angular change function, a cumulative angular function is introduced to overcome the problem. The cumulative angular function is defined as,

$$\gamma(t) = \int_0^t \kappa(r)dr - \kappa(0) \quad (8)$$

where parameter *t* takes values from 0 to the length *L* of the boundary curve. The cumulative angular function still has problems: it suffers the discontinuity problem at the boundary, and its value depends on the length of the curve. These problems can be addressed by introducing normalized function $\gamma^*(t)$,

$$\gamma^*(t) = \gamma\left(\frac{L}{2\pi}t\right) + t \quad (9)$$

where *t* takes values from 0 to 2π. The factor $L/2\pi$ normalizes the angular function in such a way that its value does not change when the curve is scaled.

FOURIER DESCRIPTOR

Fourier descriptor (FD) has been widely used for shape signatures [27] as a kind of remedy of the weak discrimination capability of the moment-based descriptors. FD aims to capture global shape features in low frequency terms, and finer features of the shape with higher frequency terms. Advantages of FD-based method include easy normalization and rotation invariant property. Although wavelet descriptors achieve localization in both spatial and frequency domains and hence have several advantages over FD for feature extraction, they lack rotation invariance property.

Once shape signature has been computed, Fourier transform is applied to the shape signature. Hence continuous form of shape boundary needs to be first sampled at fixed number of points. As shapes have different sizes, the normalization process is employed so that the shape signature must be sampled to have the same number of data points. Moreover, applying the fast Fourier transform (FFT) requires the number of sampled points to be power-of-two integer.

Suppose shape signature $s(t)$ has been computed for normalized points t=0,1, ..., L. The discrete Fourier transform of $s(t)$ is expressed by,

$$Z(k) = \frac{1}{N}\sum_{t=1}^{N-1} s(t)\exp(\frac{-j2\pi kt}{N}), k = 0,1,\dots,N-1 \quad (10)$$

The coefficient $Z(k)$ are called Fourier descriptors of the shape analyzed. For the sake of simplicity and clarity of notation, Fourier descriptors are denoted by FD_k instead of $Z(k)$ for k=0, 1,..., N-1. As aforementioned, rotation invariance property by FDs is achieved by ignoring the phase information and by taking only the magnitude values of the FDs.

1. FDs for complex coordinate signature

For the complex coordinate signature, the DC component is usually discarded because its value depends on the position of the shape, and the remaining FDs are used in describing the shape. Scale normalization is accomplished by dividing the magnitude of FDs by the magnitude of the second FD to obtain the rotation-invariant features,

$$f_{FD} = (\frac{|FD_2|}{|FD_1|}, \frac{|FD_3|}{|FD_1|}, \dots, \frac{|FD_{N-1}|}{|FD_1|}) \quad (11)$$

2. FDs for centroid distance and curvature signature

Centroid distance signature and curvature signature have only real values, so that Fourier transform takes only *N*/2 different frequencies, and therefore, only half of the FDs are needed to represent the shape. Scale invariance is achieved by dividing the magnitude of the first half of FDs by the DC component,

$$f_{FD} = (\frac{|FD_1|}{|FD_0|}, \frac{|FD_2|}{|FD_0|}, \dots, \frac{|FD_{N/2}|}{|FD_0|}) \qquad (12)$$

This feature vector is invariant under scales, translations, and rotations.

EXPERIMENTS AND CONCLUSIONS

We evaluated the proposed method by classifying 30 texture classes from the Brodatz image database as shown in Figure 5.

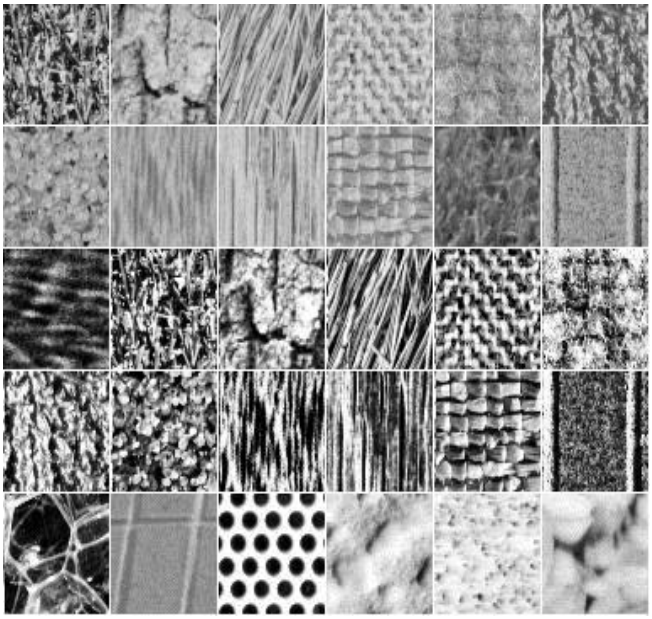

Figure 5. Test images obtained from the Brodatz database.

For each image, non-overlapping sub-images were cropped and then rotated by a number of angles to construct the training and test set. The training set comprised of textures rotated by 0º, 30º, 45º, and 60º, and the test set comprised of textures rotated by 20º, 70º, 90º, 120º, and 135º. Consequently, the training set consists of 600 samples (30 classes x 4 angles x 5 samples) and the test set of 750 samples (30 classes x 5 angles x 5 samples).

Gabor filter bank is constructed with four radial frequencies (4, 8, 16, and 32 cycles per patch width) and 12 orientations (0º, 30º, 60º, ..., 270º, 300º, and 330º). The reason why we employed more orientations than typical setting is that Gabor filter energy values are used as the input to the Fourier transforms, which results in the rotation-invariant features. For each texture patch in the training and test dataset, we applied Gabor filters and obtained energy of each Gabor filter. Figure 6 shows the magnitude of Gabor energy at the 12 orientations with radial frequency of 32 cycles for corresponding texture in Figure 5.

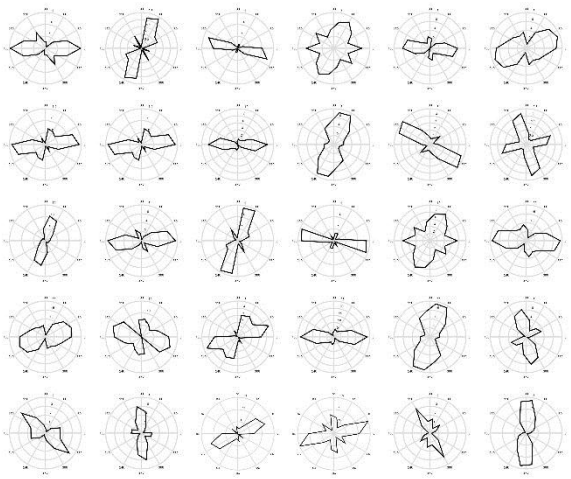

Figure 6. Magnitude of Gabor energy at 12 orientations with radial frequency of 32 cycles for textures in Figure 5.

In Figure 6, one can see that Gabor features are excellent indicators for discriminating different texture patterns. These features well preserve for varying scales as illustrated in Figure 7 which shows the magnitude of Gabor energy at the 12 orientations with radial frequency of 16 cycles for corresponding texture in Figure 5.

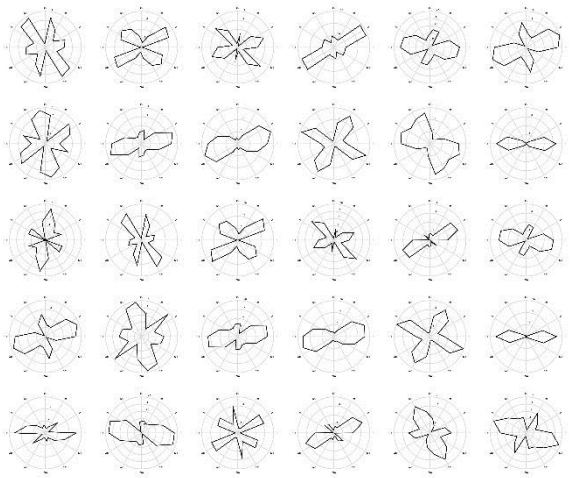

Figure 7. Magnitude of Gabor energy at 12 orientations with radial frequency of 16 cycles for textures in Figure 5.

The rotation-invariant features that characterize each texture class are obtained by applying the Fourier transform to the Gabor energy magnitudes and collecting the FDs. Classification tests have been conducted on the FDs using the multiclass SVM classifier [28]. Table I shows the experimental results for 10-fold tests with various size of Gabor functions. One can see that Gabor filter plus FD based method consistently outperforms over the plain Gabor filter based texture classification for rotation-invariant classification problem.

Table I: Texture classification results (correct classification rates) of 10-fold experiments on the 30 textures of the Brodatz database

Filter Size	Features	Mean (%)	Std.
16 x 16	Gabor + FD	**82.43**	±1.12
	Gabor	80.12	±1.64
32 x 32	Gabor + FD	**82.61**	±0.98
	Gabor	81.89	±1.55
64 x 64	Gabor + FD	**83.27**	±0.91
	Gabor	82.58	±1.27

ACKNOWLEDGEMENTS

This work was supported by the research grant of Pai Chai University in 2017, and partially supported by the Korea National Research Foundation with the research grant of NRF-2017R1D1A2B03028954.

REFERENCE

[1] T. Randen and J.H. Husoy, "Filtering for texture classification: a comparative study," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.21(4), pp. 291-310, 1999.

[2] Tuceryan, M. and Jain, A.K., "Texture analysis," In: Chen, C.H., Pau, L.F. and Wang, P.S.P., Eds., Handbook of Pattern Recognition & Computer Vision, World Scientific Pub Co Inc., 1993.

[3] Ojala T, Pietikäinen M & Mäenpää T, "Multiresolution gray-scale and rotation invariant texture classification with Local Binary Patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24(7), pp. 971-987, 2002.

[4] R. M. Haralick, "Textural Features for Image Classification," Studies in Media and Communication SMC-3(6), pp. 610-621, 1973.

[5] B. V. Dasarathy and E. B. Holder, "Image characterizations based on joint gray level—run length distributions," Pattern Recognition Letters, vol. 12(8), pp.497-502, 1991.

[6] T. Ojala, K. Valkealahti, E. Oja, M. Pietikainen, "Texture discrimination with multidimensional distributions of signed gray level differences," Pattern Recognition vol.34 (3), pp. 727–739, 2001.

[7] T. Ojala and M. Pietikainen, "A comparative study of texture measures with classification based on feature distributions," Pattern Recognition vol.29 (1), pp.51–59. 1996.

[8] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol.28(12), pp.2037 -2041, 2006.

[9] H. Voorhees and T. Poggio, "Computing texture boundaries from images," Nature vol.333, pp.364–367, 1988.

[10] F. Tomita and S. Tsuji, "Computer Analysis of Visual Textures," Kluwer, Norwell, MA, 1990.

[11] M. Tuceryan and A. K. Jain, "Texture segmentation using Voronoi polygons," IEEE Trans. Pattern Anal. Mach. Intell. vol.12, pp. 211-216, 1990.

[12] R. Chellappa and A. Jain, Markov random fields. Theory and application, Academic Press, 1993.

[13] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," IEEE Trans. Pattern Anal. Mach. Intell., vol.6(6), pp. 721-741, 1984.

[14] S. Z. Li, Markov Random Field Modeling in Image Analysis, Springer, 2001.

[15] K. Laws, "Rapid texture identification," Proceedings SPIE Image Processing for Missile Guidance, vol. 238, pp.376-380, 1980.

[16] M. Unser and M. Eden, M. "Multiresolution feature extraction and selection for texture segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol.11(7), pp.717-728, 1989.

[17] G. N. Chaple, R. D. Daruwala, and M. S. Gofane, "Comparisions of Robert, Prewitt, Sobel operator based edge detection methods for real time uses on FPGA," 2015 Int'l Conf. Tech. for Sustainable Development (ICTSD), 2015.

[18] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," " IEEE Trans. Pattern Anal. Mach. Intell., vol.11 (7), pp.674-693, 1989.

[19] M. Clark and A. C. Bovik, "Texture segmentation using Gabor modulation/ demodulation," Pattern Recognition Letters, vol.6(4), pp. 261-267, 1987.

[20] A.K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," 1990 IEEE Int'l. Conf. Systems, Man and Cybernetics, 1990.

[21] B. S. Manjunath and W.Y. Ma, "Texture features for browsing and retrieval of image data," IEEE Trans. Pattern Anal. Mach. Intell., vol.18(8), pp. 837-842, 1996.

[22] J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," Journal of the Optical Society of America A, vol.2(7), pp. 1160-1169, 1985.

[23] N. Petkov, "Biologically motivated computationally intensive approaches to image pattern recognition," Future Generation Computer Systems, vol.11 (4-5), pp.451-465, 1995.

[24] D. Clausi, M. Ed Jernigan, "Designing Gabor filters for optimal texture separability," Pattern Recognition, vol.33, pp. 1835-1849, 2000.

[25] D. Zhang and G. Lu, "A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures," Proc. of 5th Asian Conference on Computer Vision (ACCV), pp. 646-651 , 2002.

[26] M. Sidram and N. Bhajantri, "A Novel Shape Signature of Geometric Mean of Segmented Centroid Distance Function to Track the Object through Fourier Descriptors," International J. of Computer Applications, vol.83(14), 2013.

[27] H. Kauppinen, T. Seppanen and M. Pietikainen, "An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D Shape Classification," IEEE Trans. PAMI-17(2):201-207, 1995.

[28] J. Weston and C. Watkins, "Multi-class support vector machines," Proceedings of 7th European Symposium on Artificial Neural Networks (ESANN99), 1999.