Texture Classification Using Rotation- and Scale-Invariant Gabor Texture Features

Farhan Riaz, Ali Hassan, Saad Rehman, and Usman Qamar

***Abstract*—This letter introduces a novel approach to rotation and scale invariant texture classification. The proposed approach is based on Gabor filters that have the capability to collapse the filter responses according to the scale and orientation of the textures. These characteristics are exploited to first calculate the homogeneous texture of images followed by the rearrangement of features as a two-dimensional matrix (scale and orientation), where scaling and rotation of images correspond to shifting in this matrix. The shift invariance property of discrete fourier transform is used to propose rotation and scale invariant image features. The performance of the proposed feature set is evaluated on Brodatz texture album. Experimental results demonstrate the superiority of the proposed descriptor as compared to other methods considered in this letter.**

***Index Terms*—Gabor filters, pattern recognition, texture analysis.**

I. Introduction

TEXTURE is an important visual cue used for a variety of image analysis applications such as image segmentation, image retrieval, pattern classification etc. Due to its ever increasing importance, it has been an active area of research over several decades now. Texture classification principally involves two major steps: feature extraction and classification. Most research on texture classification focuses on feature extraction given that more powerful and highly discriminative features will yield higher classification accuracies for classifying the underlying texture irrespective of the classifier used for the said purpose. This fact has stimulated research on numerous methods of feature extraction that can be broadly divided into four distinct categories [1]: a) statistical methods, b) model-based methods, c) structural methods and d) filter-based methods.

Most of the statistical and model based approaches [2], [3] consider spatial interaction over relatively small neighbourhoods. Therefore, these approaches are more useful for microtextures. Structural methods are more useful for textures which follow regular or at least semi-regular placement of basic texture primitives [1]. Filter based methods are proposed by many researchers to mitigate these problems [1]. These methods can be grouped into three main categories: a) spatial filtering, b) frequency filtering, and c) spatial-frequency filtering. A weakness shared by the first two is that the image is analyzed in these cases at only a single scale. In addition, the frequency localization in spatial methods and spatial localization in frequency methods is lost. The spatial-frequency methods have been designed to overcome these limitations. An important typical requirement for the design of robust texture features is their invariance to scale and rotation of the images. Several methods to address this issue have been proposed, one of the most widely used methods being the scale invariant feature transform (SIFT). However, it is a local feature and its usage is tied to the existence of reliable key points (e.g., an object) which is not typically the case in texture images [4] (the texture characteristics may not only be encapsulated only by the key points). Additionally, SIFT may not be a generic descriptor given that the interest point detection in SIFT is done by rule based filtering (sensitive to threshold selection) [5]. Moreover, the filtering step in SIFT only uses differential features while ignoring the Gaussian scale-space which tends to be useful in describing local pixel neighborhoods. Another important class of filters used for feature extraction is Gabor filters. They have two distinct characteristics which are not shared by any other method: their similarity with primary visual cortex, and their optimal space-frequency resolution [6]. Also, their capability to capture the filter responses at various scales and orientations can be exploited for designing scale and rotation invariant image descriptors. Several attempts have been made to propose invariant features using Gabor filters. K. Jafari-Khouzani *et al.* [7] used radon transform for rotation invariant texture features. R. Manthalkar *et al.* [8] proposed even symmetric Gabor filters for extracting rotation invariant features. J. Han *et al.* [9] proposed a method to obtain rotation or scale invariant texture features. Some other methods (such as [10]) are more suitable for matching related applications and cannot be used directly for classification using feature space classifiers. The contributions of this letter are as follows: we propose texture features using Gabor filters which are invariant to scale and rotation changes in the images. State-of-the-art Gabor features are first calculated, followed by a rearrangement of the features as a matrix according to filters responses at various scales and orientations. Shift invariance property of the discrete fourier transform is later used to induce scale and rotation invariance to the calculated features. Experimental evaluation of the descriptor is done on Brodatz texture. The letter is organized as follows: In Section II, we describe a brief introduction of Gabor Filters followed by the extension of the state-of-the-art texture features in the proposed methodology (Section III). Later, we discuss our experimental results (Section IV) and conclude the outcomes of this research (Section V).

Manuscript received February 23, 2013; revised April 09, 2013; accepted April 17, 2013. Date of publication April 23, 2013; date of current version May 06, 2013. This work was supported by the National University of Sciences and Technology (NUST), Islamabad. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Qian Du.

The authors are with the National University of Sciences and Technology, Islamabad, Pakistan (e-mail: farhan.riaz@ceme.nust.edu.pk; alihassan@ceme.nust.edu.pk; saadrehman@ceme.nust.edu.pk; usmanq@ceme.nust.edu.pk).

Digital Object Identifier 10.1109/LSP.2013.2259622

II. Gabor Filters

Gabor filters have been widely used for extraction of texture features from images. The foundations for wide usage of Gabor filters for texture description were mainly laid by Manjunath *et al.* in [11] when he proposed homogeneous texture (HT) descriptor, which was later used as one of the visual texture descriptors in MPEG-7. Since then, they have been widely used for texture classification. A two dimensional Gabor function $g(x, y)$ and its Fourier transform $G(u, v)$ can be written as:

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right) \tag{1}$$

$$G(u,v) = \exp\left(\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right) \tag{2}$$

Where $\sigma_u = 1/2\pi\sigma_x$, $\sigma_v = 1/2\pi\sigma_y$ and W is a constant representing the center frequency of the filter bank having the highest frequency. This forms a bandpass filter in the frequency domain, where the bandwidth and center frequency of the filter are controlled by the standard deviation of the Gaussian function and the frequency of complex sinusoid respectively. A Gabor filter bank having a number of bandpass filters, with varying center frequencies, bandwidths and orientations is controlled by the parameters of Gabor wavelets. An input image, $\xi(x, y)$ when filtered by the set of Gabor wavelets $g(x, y)$ is given as:

$$R_{mn}(x,y) = \int \xi(x,y) g^*_{mn}(x-x_1, y-y_1)dx_1dy_1 \tag{3}$$

where $R_{mn}(x, y)$ is the filter response at the spatial location (x, y), $m = 1, 2, \ldots, M$ is the number of scales and $n = 1, 2, \ldots, N$ is the number of orientations. Manjunath *et al.* [11] assumed that local image regions are spatially homogeneous and the mean and standard deviation of the magnitude of the filter responses are used to represent the region for classification and retrieval purposes:

$$\mu_{mn} = \int\int \mid R_{mn}(x,y) \mid dxdy \tag{4}$$

$$\sigma_{mn} = \sqrt{\int\int\int (\mid R_{mn}(x,y) \mid - \mu_{mn})^2 dxdy} \tag{5}$$

A feature vector is constructed using μ_{mn} and σ_{mn} as feature components is known as the HT descriptor:

$$HT = [\, \mu_{11} \quad \sigma_{11} \quad \mu_{12} \quad \sigma_{12} \quad \cdots \quad \mu_{MN} \quad \sigma_{MN} \,] \tag{6}$$

III. Proposed Methodology

An important shortcoming of HT is that it is not invariant to rotation, scale and illumination changes in the images. This happens because if an image is rotated by a certain angle, its response is usually represented by that filter in the filter bank which has been rotated by the same amount as that of the image. When these features are concatenated as a vector, the rotation changes appear as shifts in the feature vectors. Application of shift invariant operators on these features can give us rotation invariant Gabor features. Similar property holds for scale changes in the images. Let us now assume that we have

$$s_\mu = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1N} \\ \vdots & \ddots & \vdots \\ \mu_{M1} & \cdots & \mu_{MN} \end{bmatrix} \tag{7}$$

which represents a collection of means in the HT descriptor in the form of a matrix. The s_μ extracted for a rotated image would be represented by a shift in the columns of this matrix. Similarly, the s_μ of a scaled image would be represented by a shift of the rows of this matrix. It is well known that the magnitude of the Discrete Fourier Transform (DFT) is invariant to shifts [12] and thus the effect of rotation or scaling on the feature vector can be curtailed using 2D DFT of s_μ.

$$S_\mu(k,l) = \frac{1}{\sqrt{MN}} \cdot \sum_{m=0}^{M-1}\left[\sum_{n=0}^{N-1} s_\mu(m,n)e^{-j2\pi(mk/M)}\right]e^{-j2\pi(nl/N)} \tag{8}$$

where $S_\mu(k, l)$ has a collection of coefficients of the DFT of s_μ. Similarly, s_σ is defined as

$$s_\sigma = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1N} \\ \vdots & \ddots & \vdots \\ \sigma_{M1} & \cdots & \sigma_{MN} \end{bmatrix} \tag{9}$$

The application of 2D DFT on s_σ gives us

$$S_\sigma(k,l) = \frac{1}{\sqrt{MN}} \cdot \sum_{m=0}^{M-1}\left[\sum_{n=0}^{N-1} s_\sigma(m,n)e^{-j2\pi(mk/M)}\right]e^{-j2\pi(nl/N)} \tag{10}$$

It is important to note that the magnitude of S_μ and S_σ are invariant to rotation and scale changes in the images due to shift invariance property of DFT. The coefficients obtained from S_μ and S_σ are arranged as a vector giving us:

$$F = [\, S_\mu \quad S_\sigma \,] \tag{11}$$

where F is the feature vector which is obtained by the concatenation of the rotation and scale invariant coefficients of the statistics of the Gabor filter responses of the images.

IV. Experimental Results

We demonstrate the efficiency of our proposed descriptor using four datasets. For our experiments, we have used Gabor filters with 6 orientations and 4 scales. This selection is based on our empirical grid search to give optimal classification results while avoiding the problem of curse of dimensionality of the feature vectors. Consequently, a 48-dimensional feature vector is obtained for each pixel. Classification was done using Support Vector Machines (SVM) [13] with polynomial kernel using the Weka toolkit [14]. The optimal parameters were determined using cross validation and used for the classification experiments presented in the letter. The performance of novel image descriptors was compared with some other

TABLE I
SVM CLASSIFICATION PERFORMANCE OF DIFFERENT TEXTURE DESCRIPTORS ON THE FULL BRODATZ TEXTURE ALBUM (LBP – LOCAL BINARY PATTERNS)

Methods	Accuracy
Proposed	**0.98**
HT [11]	0.92
LBP [3]	0.95

state-of-the-art feature extraction methods which show good performance on the Brodatz texture album. To ensure a fair comparison, consistent test bed was used for all the methods. Performance comparison was done using overall classification accuracy which is defined as the total number of correctly classified samples averaged by the total number of samples in the test set. We have evaluated the performance of our proposed descriptor on four datasets.

A. *Dataset 1*

The first dataset used for evaluation of the proposed descriptor is composed of 111 images from the Brodatz texture album [15]. This is a challenging dataset given that it is composed of textures having an impressive diversity and perceptual similarity of some of the textures essentially belonging to the same class but at different scales. Some others are so inhomogeneous that even human observers may not be able to group them correctly [16].

The size of each image in the album is 512×512. For our experiments, we divided each image into 16 non-overlapping regions, each region of size 128×128. Among these 16 regions, 8 were used for training the classifier while the remaining 8 regions were used for testing the performance of the classifier. Thus, for a total of 111 classes we obtained 1776 (111×16) non-overlapping regions out of which 888 were used for training while 888 were used for testing the classifier.

Our experiments demonstrate that the classification results obtained using the proposed feature set outperforms the state-of-the-art texture feature extraction methods considered in this letter (Table I).

B. *Dataset 2*

The second dataset that we used for evaluating the performance of the proposed descriptor is the rotated texture database of the Brodatz texture album. This album consists of 13 textures (D12, D94, D112, D9, D24, D92, D84, D29, D15, D38, D16, D68 and D19) rotated at discrete angles of 30, 60, 90, 120, 150 and 200 degrees (USC-SIPI image database [17]). We used this album to analyze the rotation invariance properties of the proposed descriptor as compared to its state-of-the-art counterparts. For these experiments, the original texture images were used for training (13 textures) the classifier while the feature set from images rotated at 30, 60, 90, 120, 150 and 200 degrees were used for testing the performance of the classifier.

Our experiments show that the proposed feature set shows good stability to rotation changes in the images (Table II). The instability of HT to the image rotations is evident from the fact that for the rotated dataset, the average accuracy for classifying the images is just around 0.48. The rotation invariant local binary patterns show good rotation invariance characteristics giving an average classification accuracy of 0.92.

TABLE II
STABILITY OF CLASSIFICATION RESULTS WHEN IMAGES ARE ROTATED AT DIFFERENT ANGLES (LBP – LOCAL BINARY PATTERNS)

Methods	Accuracy						
	30°	60°	90°	120°	150°	200°	Mean
Proposed	**1**	**1**	**1**	**1**	**1**	**1**	**1**
HT [11]	0.62	0.38	0.31	0.31	0.62	0.62	0.48
LBP [3]	0.92	0.92	0.92	0.92	0.92	0.92	0.92

TABLE III
STABILITY OF CLASSIFICATION RESULTS WHEN IMAGES ARE SCALED BY VARIOUS FACTORS (LBP – LOCAL BINARY PATTERNS)

Methods	Accuracy				
	×0.66	×0.8	×1.25	×1.5	Mean
Proposed	**0.86**	**1**	**0.86**	**0.86**	**0.90**
HT [11]	0.70	1	0.70	0.70	0.78
LBP [3]	0.23	0.69	0.23	0.23	0.35

C. *Dataset 3*

The objective of the third dataset is to analyze the scale invariance characteristics of the proposed descriptor. For this purpose, the original images in the *dataset 2* (13 images) were scaled by certain factors ($\times$ 0.66, $\times$ 0.8, $\times$ 1.25 and $\times$ 1.5) and used for the underlying experiments. For evaluation purposes, we used the original texture images (no rotation) for training the classifier while testing was done on the scaled images. Our experiments show (Table III) that with the scale changes, the classification performance depreciates for all the image descriptors. However, the proposed features are most stable among all other feature extraction methods.

This happens because the scaling of images is not optical (through camera lens) as it is done using signal processing methods (interpolation). This introduces some artifacts in the images which are responsible for a certain amount of deterioration in the recognition performance of the image descriptors. Experiments show that the proposed descriptors show good scale invariance characteristics as compared to the other state-of-the-art feature extraction methods (Table III).

D. *Dataset 4*

In addition to analyzing the stability of the proposed descriptor to scale or rotation changes, it is important to evaluate the invariance of the descriptor if the images are subjected to both rotation and scale changes. For this purpose, all the images (rotated at various angles) in the *dataset 2* are subjected to scale changes. The original texture images were used for training (13 textures) the classifier while the images rotated by 30, 60, 90, 120, 150 and 200 degree and scaled by $\times$ 0.66, $\times$ 0.8, $\times$ 1.25 and $\times$ 1.5 were used for testing the classifier performance. Experiments show that as compared to other state-of-the-art algorithms, the proposed feature set shows good stability characteristics even when the images are subjected to both rotation and scale changes (Table IV).

V. CONCLUSION

This letter proposes a novel texture descriptor that is invariant to rotation and scale changes in the images. The feature extraction is done using Gabor filters. At first, the state-of-the-art

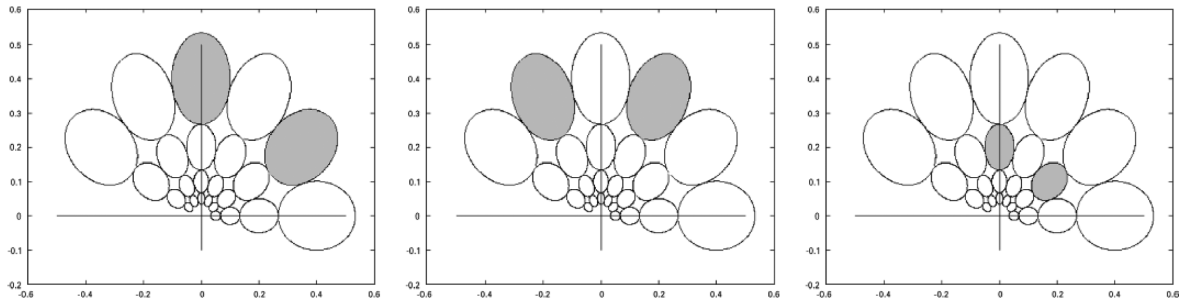

Fig. 1. Frequency spectrum of a Gabor filter bank: Filter response for original image (left), rotated image (middle) and scaled image (right).

TABLE IV
STABILITY OF CLASSIFICATION RESULTS WHEN THE IMAGES ARE SUBJECTED TO BOTH ROTATION AND SCALE CHANGES. THE ROWS FOR EACH METHOD CORRESPOND TO THE SCALING FACTORS OF $\times$ 0.66, $\times$ 0.8, $\times$ 1.25 AND $\times$ 1.5 (LBP – LOCAL BINARY PATTERNS)

Methods		Accuracy						
		30°	60°	90°	120°	150°	200°	Mean
Proposed	×0.66	**0.69**	**0.69**	**0.69**	**0.69**	**0.69**	**0.69**	**0.69**
	×0.8	**0.77**	**0.69**	**0.69**	**0.77**	**0.69**	**0.77**	**0.73**
	×1.25	**0.69**	**0.77**	**0.62**	**0.69**	**0.62**	**0.77**	**0.69**
	×1.5	**0.69**	**0.77**	**0.69**	**0.69**	**0.77**	**0.69**	**0.72**
HT [11]	×0.66	0.15	0.15	0.23	0.23	0.23	0.15	0.19
	×0.8	0.54	0.38	0.30	0.30	0.38	0.38	0.38
	×1.25	0.30	0.23	0.30	0.30	0.30	0.30	0.29
	×1.5	0.30	0.23	0.30	0.30	0.30	0.30	0.29
LBP [3]	×0.66	0.23	0.23	0.30	0.30	0.15	0.15	0.23
	×0.8	0.54	0.61	0.54	0.61	0.54	0.54	0.56
	×1.25	0.07	0.23	0.15	0.15	0.07	0.15	0.14
	×1.5	0.07	0.15	0.07	0.07	0.15	0.23	0.12

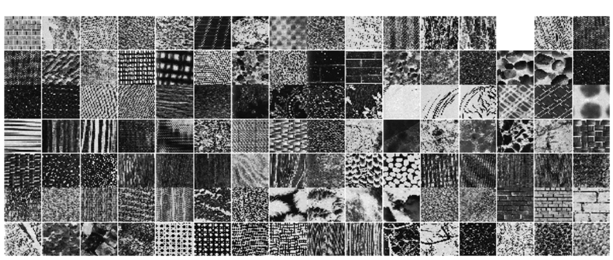

Fig. 2. Texture images from the full Brodatz album.

texture features (Homogeneous Texture – HT) are extracted. This is followed by the arrangement of the HT features in the form of two distinct matrices (for means and variances separately) that interpret the rotation and scale changes in the images as shifting of rows and columns respectively. The 2D Discrete Fourier Transform is later applied to the matrices given its shift invariance property thus giving us features which are invariant to scale and rotation changes in the images. The features extracted from the images using the proposed feature set are used for classification of Brodatz album images. For classification purposes, we have used Support Vector Machines. We have compared the performance of our feature extraction methodology with other state-of-the-art feature extraction methods giving very good performance in the used datasets. Experiments show the superiority of our proposed feature set as compared to other texture feature extraction methods considered in this letter.

Although our proposed features are invariant to rotation and scale changes in the images as compared to the other state-of-the-art methods, there are several scenarios which require the features to be invariant to illumination gradients in the images. In the future, we intend to explore this issue and try to incorporate invariance in the texture features pertaining to illumination gradients in addition to the rotation and scale changes in the images.

REFERENCES

[1] S. Selvan and S. Ramakrishnan, “Svd-based modeling for texture classification using wavelets transformation,” *IEEE Trans. Image Process.*, vol. 16, no. 11, 2007.

[2] R. M. Haralick, K. Shanmugan, and I. Dinstein, “Texture features for image classification,” *IEEE Trans. Syst., Man Cybern.*, vol. 3, no. 6, 1973.

[3] T. Ojala, M. Pietikinen, and T. Maenpaa, “Gray scale and rotation invariant texture classification with local binary patterns,” in *Proc. ECCV*, 2000.

[4] C. Siagian and L. Itti, “Rapid biologically-inspired scene classification using features shared with visual attention,” *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 29, no. 2, pp. 300–312, 2007.

[5] B. Li, R. Xiao, Z. Li, R. Cai, B.-L. Lu, and L. Zhang, “Rank-Sift: Learning to rank repeatable local interest points,” in *Proc. 2011 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 1737–1744.

[6] J. G. Daugman, “Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters,” *J. Opt. Soc. Amer.*, vol. 2, no. 7, 1985.

[7] K. Jafari-Khouzani and H. Soltanian-Zadehl, “Radon transform orientation estimation for rotation invariant texture analysis,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 27, no. 6, 2005.

[8] R. Manthalkar, P. K. Biswas, and B. N. Chatterji, “Rotation invariant texture classification using even symmetric Gabor filters,” *Patt. Recognit. Lett.*, vol. 24, no. 12, 2003.

[9] J. Han and K. K. Ma, “Rotation-invariant and scale-invariant Gabor features for texture image retrieval,” *Image Vis. Comput.*, vol. 25, no. 9, 2007.

[10] X. Xie, Q. Dai, K. M. Lam, and H. Zhao, “Efficient rotation- and scale-invariant texture classification method based on Gabor wavelets,” *J. Electron. Imag.*, vol. 17, no. 4, 2008.

[11] B. S. Manjunath and W. Y. Ma, “Texture features for browsing and retrieval of image data,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 18, no. 8, 1996.

[12] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.

[13] V. Vapnik, *The Nature of Statistical Learning Theory*. Berlin, Germany: Springer, 1995.

[14] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten, “The Weka data mining software: An update,” *ACM SIGKKD Explorations Newslett.*, vol. 11, no. 1, 2009.

[15] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York, NY, USA: Dover, 1966.

[16] S. Lazebnik, C. Schmid, and J. Ponce, “A sparse texture representation using local affine regions,” *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1265–1278, 2005.

[17] *T. U.-S. I. Database*, , 2003 [Online]. Available: http://sipi.usc.edu/database/?volume=rotate