A HYBRID ALGORITHM USING DISCRETE COSINE TRANSFORM AND GABOR FILTER BANK FOR TEXTURE SEGMENTATION

[1]Nezamoddin N. Kachouie, [1,2]Javad Alirezaie and [1]Paul Fieguth
[1]Systems Design Eng. Department, University of Waterloo, Waterloo, Canada
[2]Electrical and Computer Eng. Department, Ryerson University, Toronto, Canada

Abstract

Gabor filters have been widely used for texture segmentation and feature extraction, however there are important considerations regarding filter parameters, filter bank coverage in the frequency domain and feature dimensional reduction. In this paper, a texture segmentation algorithm based on a hybrid filter bank is presented. The proposed method uses a Gabor filter bank and Discrete Cosine Transform (GDCT) to extract the optimal features for texture segmentation. To reduce the feature vector dimension a competitive network is trained to estimate the principal components of the extracted features. The feature vectors composing of both Gabor and DCT features are quantized by estimated eigen vectors. The proposed method enables the use of multiple filter banks or larger filter banks consisting of a higher number of channels.

Keywords: *Gabor filters bank; DCT; Feature extraction; Segmentation; Texture; Neural Networks.*

1. INTRODUCTION

In image analysis, textures have been used to perform scene segmentation, object and shape recognition, surface classification and region discrimination. Texture analysis needs to identify attributes which are useful to discriminate, recognize and segment different texture types and involves accurately partitioning an image into sections by recognizing the textured regions or the borders between different textures.

There are several filter bank choices for textured segmentation. From a practical point of view, some filters may be more useful for specific segmentation tasks but not for others. Gaussian filters modulated by an exponential or by sinusoidal filters, known as Gabor filters, have been proved very useful to analyze textures that contain specific frequency and orientation characteristics [1, 13]. While different filter banks can perform joint spatial/spatial-frequency decomposition, a Gabor filter bank using a Gabor base function is one of the most attractive ones. This set of filters has an optimal localization in the joint spatial/spatial-frequency domain according to the uncertainty principal [1, 4]. These selective band pass filters with different radial spatial frequency and orientation have optimum resolution in the time and frequency domains that resemble the simple visual cortical cell characteristics.

Dunn and Higgins have investigated designing the optimal Gabor filters [4] and argued that Gabor filter outputs can be presented as a Rician model and developed an algorithm to select optimal filter parameters to discriminate texture pairs. Jain and Farrokhnia proposed the optimum radial frequencies and orientations for different channels that have been used widely by many researchers [6]. For further information on regarding previous research using Gabor filter banks please refer to [1,3,4,6,7,12,13].

This paper describes a texture segmentation method according to a hybrid multi-channel decomposition approach. A set of filtered images is generated by applying a Gabor filter bank and a discrete cosine transform to the input image. The multi-channel decomposition is accomplished by estimating the local energy in the filter outputs. Having 20 filters in the Gabor filter bank and 9 DCTs, twenty nine filtered images are generated. A competitive network is trained to estimate the eigen vectors of the training samples and reducing the feature dimensions through principal components. Eventually the resultant features are used to train a multi layer perceptron and segment the input image. In section 2 the multi-channel decomposition by Gabor and DCT is reviewed. Section 3 describes the proposed method using Gabor Filter bank in conjunction with DCT (GDCT). In sections 4 and 5 the results and conclusions are presented respectively.

2. GABOR FILTER BANK

The Gabor base function is a Gaussian function modulated with an exponential or a sinusoid that is defined in terms of the product of a Gaussian and an

CCECE 2004- CCGEI 2004, Niagara Falls, May/mai 2004
0-7803-8253-6/04/$17.00 ©2004 IEEE

exponential. Two-dimensional Gabor functions h(x,y) can be written as:

$$h(x,y) = g(x,y)\ \exp(2\pi j f_0 x) \qquad (1)$$

where,

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\cdot\exp\{-\frac{1}{2}(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2})\} \qquad (2)$$

and its frequency response H(U,V) is:

$$H(u,v) = G(u-f_0, v) = \exp\{-2\pi^2[(u-f_0)^2\sigma_x^2 + v^2\sigma_y^2]\} \quad (3)$$

Gabor functions are band-pass filters which are Gaussians, centered on f_0 in the spatial-frequency domain. The parameters f_0, σ_x and σ_y determine the sub-band Gabor filter where f_0 is center frequency and σ_x and σ_y are the bandwidths of the filter. Equation (2) defines a complete Gabor function consisting of both real and imaginary (or even and odd) components. Rotation by θ in the spatial domain (x-y plane) or in the spatial-frequency domain (u-v plane) provides selective arbitrary orientation for different channels. We can implement a Gabor filter bank by using only even-symmetric or real components, as suggested by Jain and Farrokhnia [6], represented by,

$$h(x,y) = g(x,y)\ .\ \cos\{2\pi f_0 x\} \qquad (4)$$

The multi-channel decomposition method as shown in Fig. 1 consists of 3 major stages: applying the filter bank, estimating the local energy and segmenting the extracted features into different regions. In the first step a textured input image is decomposed into filtered images. In the second stage, a local energy function consisting of nonlinearity and smoothing is applied to the filtered images [1,6,3,12] and a set of feature images or a feature vector corresponding to each pixel in the input image is generated.

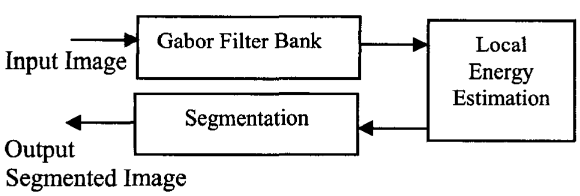

Fig.1. Multi-channel decomposition

Finally the feature vectors are segmented according to the different textures and assigned to regions.

3. DISCRETE COSINE TRANSFORM

The DCT transforms a signal from a spatial representation into a frequency representation. Because of the fast implementation and good results, DCT is widely used in image compression algorithms such as JPEG. The DCT can decompose the image into spectral sub-bands having different importance with respect to the visual quality of the image. After applying the two dimensional DCT to the input image *I*, the coefficients of the output image *G* are obtained by:

$$G_{vu} = \frac{1}{4}A_v A_u \sum_{y=0}^{N-1}\sum_{x=0}^{N-1} I_{yx}\cos\left(v\pi\frac{2y+1}{2N}\right)\cos\left(u\pi\frac{2x+1}{2N}\right)$$

$$A_u = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } u = 0 \\ 1 & \text{else} \end{cases},\quad A_v = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } v = 0 \\ 1 & \text{else} \end{cases} \qquad (5)$$

The input image is *N* pixels wide by *N* pixels high; *I* (x,y) is the intensity of the pixel in row *i* and column *j*; *G*(u,v) is the DCT coefficient in row *u* and column *v* of the DCT matrix. All DCT multiplications are real. This lowers the number of required multiplications, as compared to the discrete Fourier transform.

4. PROPOSED METHOD (GDCT)

Our proposed method (GDCT) as shown in Fig. 2 consists of three main stages:

4.1 Decomposition by Gabor and DCT

In this stage two filter banks, i.e., Gabor filter bank and DCT, are applied to the input image in parallel. Then non-linearity and smoothing functions are applied to the filtered images to estimate the local energy.

4.1.1 Filtering by Gabor

The proposed Gabor filter bank uses 5 radial frequencies: $4\sqrt{2}, 8\sqrt{2}, 16\sqrt{2}, 32\sqrt{2}, 64\sqrt{2}$

The radial frequency bandwidth is one octave. For each radial frequency 4 orientations 0°, 45°, 90° and 135° are used that generate the total number of 20 channels in the filter bank. A square function is used as nonlinearity and it causes the sinusoidal modulations in the output of the filter bank to be transformed to the square modulation. Then to smooth out the fluctuations in the specific texture or noise in the image *I* (x,y), a Gaussian low pass filter is applied. The size of the smoothing function is determined according to the size of the Gabor band pass filter.

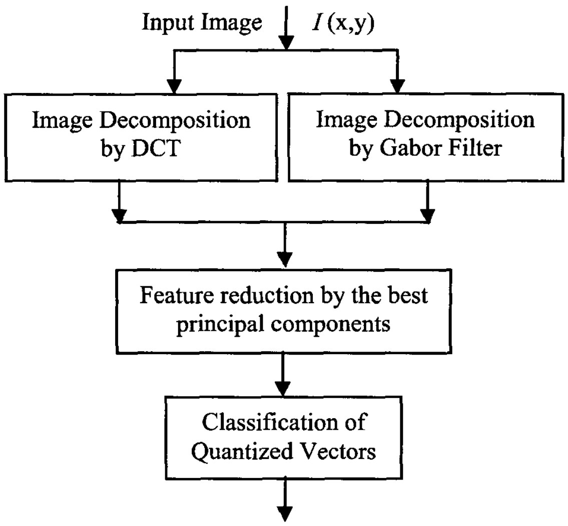

Fig.2. GDCT method

4.1.2 Applying DCT

A two dimensional DCT can be evaluated more quickly using a series of one dimensional DCTs, This means that instead of performing a 2D DCT, a 1D DCT on each row and then on each column can be performed. Ng et al. [11] introduced a 3 by 3 DCT to extract features from textured images. The suggested filter bank is separable and includes the following masks: h1 = [1; 1; 1], h2 = [1; 0; 1], h3 = [1; 2; 1]. Applying the suggested filter bank to the input image generates nine filtered images.

4.2 Feature Vector Reduction

To reduce the feature vector dimension, the best principal components of the extracted features by Gabor and DCT are estimated using training samples. A competitive network is trained to estimate the principal components of the extracted features. The feature vectors composing of both Gabor and DCT features are quantized by estimated eigen vectors. In the training phase the neurons are adjusted to be closer to the specific inputs [8] and the weights of the winners will be adjusted with the unsupervised learning rule:

$$\Delta w = \lambda(p - w) \qquad (6)$$

where p, w and λ are input vector, weight vector and learning rate respectively. The neuron whose weight vector is closest to the input vector will be updated to be even closer. During training, as more inputs are presented, each neuron in the layer closest to a group of input vectors adjusts its weights to more closely resemble the inputs. After the competitive network is trained, its weight vectors are used as quantizing mask coefficients to apply to the extracted feature vectors for dimensionality reduction. Different number of neurons was considered in the proposed competitive network. According to the two criteria, i.e., the classification error and dimension of quantized feature vectors, a network with 16 neurons is selected as the optimum.

4.3 Classification by Multi-Layer Perceptron

A Multi-Layer Perceptron (MLP) has multiple layers with nonlinear transfer functions and can learn nonlinear relationships between input and output vectors. Feedforward networks often have one or more hidden layers of sigmoid neurons followed by an output layer [5]. The MLP is trained by adjusting the weights using Least Square Error (LSE) that minimizes the mean square error between the desired and the actual outputs. Since the proposed MLP should approximate a nonlinear function, at least two layers of neurons are required. To select the proper number of layers regarding complexity and overfitting, two-layer and three-layer perceptrons are implemented, trained and tested. According to the training convergence curve, the required time and classification error, the three-layer perceptron is selected to accomplish the segmentation task. A sigmoid transfer function is used in all three layers of the proposed MLP. During training, random selected quantized feature vectors are assigned to the proper classes. After MLP is trained, quantized feature vectors corresponds to each pixel of the image are assigned to the proper regions.

5. RESULTS

In the presented approach, textured images from MIT vision and modeling database [10], Brodatz album [2], and MeasTex image texture database [9] are used to derive the input data set. The image in Fig. 5a is a textured image consisting of Fabric.0000, Fabric.0017, Flowers.0002, Leaves.0006 and Leaves.0013 selected from [10]. Fig. 5b-5d show the segmented result by our proposed GDCT method, Laws filter bank [14] and DCT [11].

Table 1:Classification Errors for Fig.5a.

Method	Error (percent)
GDCT	5.91
Gabor	27.5
DCT	15.27
Laws	25.7

The comparison of classification results using GDCT and the other methods i.e. Gabor, Laws and DCT is presented in Table 1. Our GDCT method has better classification performance in comparison with the other methods. The increase in the feature dimension using 29 features extracted by Gabor and DCT in comparison with 20 filters in the Gabor filter bank is compensated by feature quantization using a competitive layer. The dimension of the quantized feature vectors is 16.

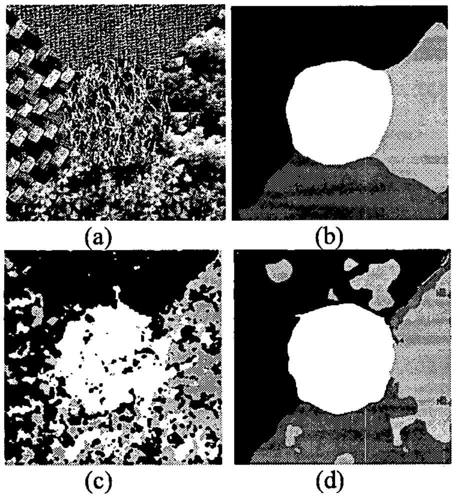

Fig.5.(a) Textured image. (b) Segmentation result by GDCT method. (c) Segmentation result by Laws filters. (d) Segmentation result by DCT.

6. CONCLUSIONS

In this paper, we have presented a new method for segmenting textured images. The classification error is reduced by optimizing Gabor filter bank and utilizing neural networks for both feature reduction and segmentation. A competitive network is combined with a MLP to learn the nonlinear relationship between quantized vectors and desired outputs. The segmentation problem is formulated as a combinational optimum problem by obtaining a better filter bank and reducing the feature dimension. The proposed method enables the use of multiple filter banks or larger filter banks consisting of a higher number of channels. The proposed feature quantization is a generic approach and could be employed to learn and process different kinds of input feature vectors. This algorithm can also be extended to problems of segmenting the synthetic aperture radar (SAR) images. Our current goal is to improve the performance of this algorithm and future research will be conducted on the application of the proposed technique for segmentation of satellite textured images.

Acknowledgments

This work was funded in part by a grant from Natural Sciences and Engineering Research Council of Canada.

References

[1] A. Bovik, M. Clark and W. Geisler, "Multichannel texture analysis using localized spatial filters", *IEEE Trans. on PAMI,* Vol. 12, No. 1, July 1990.

[2] P. Brodatz, Textures: A Photographic Album for Artists and Designers. New York: Dover, 1966.

[3] D. A. Clausi and M. E. Jernigan," Designing Gabor filters for optimal texture separability", *The journal of pattern recognition society,* Elsevier, Vol. 33, 2000.

[4] D. Dunn and W. E. Higgins, "Optimal Gabor filters for texture segmentation", *IEEE Transactions on Image Processing,* Vo4., No. 7, July 1995.

[5] J. A. Freeman and D. M. Skapura, *Neural networks algorithms, applications, and programming techniques,* Addison-Wesley, Massachusetts, 1992.

[6] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using gabor filters", *Pattern Recognition,* Vol. 24, No. 12, pp. 1167-1186, 1991.

[7] A. K. Jain and K. Karu, "Learning texture discrimination masks", *IEEE Trans. on PAMI,* Vol. 18, No. 2, pp.195-205, February 1996.

[8] T. Kohonen, T. S. Huang and M. R. Schroder, *Self organizing maps,* Springer-Verlag, Berlin, 1997.

[9] MeasTex image texture database, http://www.cssip.elec.uq.edu.au/~guy/meastex/meastex.h, 1998.

[10] MIT vision and modeling group, http://www.media.mit.edu/vismod/, 1998.

[11] I. Ng, T. Tan, & J. Kittler, On local linear transform and Gabor filter representation of texture, *Proc. Int'l Conf. on Pattern Recognition*, 1992, 627-631.

[12] T. Randen and H. Husoy, "Filtering for texture classification: A comparative study", *IEEE Trans. on PAMI,* vol. 21, no. 4, 1999.

[13] T. P. Weldon, W. E. Higgins and D. F. Dunn, "Gabor filter design for multiple texture segmentation", *SPIE Journal of OPT. ENG.,* Vol. 35, No. 10, 1996.

[14] K.I. Laws, Rapid texture identification, *Proc. SPIE Conf. on Image Processing for Missile Guidance*, 1980, 376-380.