ACCV2002: The 5th Asian Conference on Computer Vision, 23—25 January 2002, Melbourne, Australia.

Unsupervised Texture Classification of Images using Cortex Filters

M.K.Bashar†
N. Ohnishi‡
†‡*Dept. of Information Engineering, Nagoya University, Japan*
‡*Biomimetic Control Research Centre, RIKEN, Japan*
† *khayrul@ohnishi.nuie.nagoya-u.ac.jp*
‡ *ohnishi@ohnishi.nuie.nagoya-u.ac.jp*

B.K.Mohan¶
*R.K. Shevgaonkar**
¶*Centre of Studies in Resources Engineering*
**Dept. of Electrical Engineering*
Indian Institute of Technology (IIT) Bombay, India
¶bkmohan@csre.iitb.ernet.in
*rks@ee.iitb.ernet.in

Abstract

We present a block-based multi-channel mechanism for unsupervised texture classification of images inspired by Human Visual System (HVS). The proposed approach compresses the large feature space by logical selection of block. We employ 2D Gaussian functions, regarded as cortex filters, to simulate the band pass nature of simple cells in HVS. Within the frequency plane of each data block, filters are defined in various radial bands and orientations and used to obtain a set of feature images whereby texture features are defined by computing average energy. The obtained feature images are thus integrated with 'k-means clustering' for unsupervised classification of homogeneous textural regions. We demonstrate our method by several experiments on real world and synthetic images. Confusion matrix analysis projects the superiority of our method compared to gray level co-occurrence matrix (GLCM) approach.

Keywords: *texture, human visual system, cortex filter, average energy, k-means clustering.*

1. Introduction

Texture analysis is of prime importance to many applications ranging from computer vision, image synthesis to remote sensing. Such analysis actually attempts to identify the regions of uniform textures in a given image. Accurate characterization of texture is a vital issue. But spatial continuity (local and global) and randomness pose a challenge to define it. However, Sklansky's [2] definition appears to be suitable to the segmentation context: "A region in an image has a constant texture if a set of local statistics or other local properties of the picture are constant, slowly varying or approximately periodic".

Last two decades' studies include a variety of techniques namely statistical and non-statistical. Among statistical, Haralick's [1] co-occurrence matrix approach is the most popular. It explores the spatial dependency of gray levels in images. Later modifications and related works were found in [3][6]. Non-statistical techniques include fractal-based [8], and structural [9]. In the early 90's, Gabor and wavelet transform [2][4][5][10] introduce a new thrust among the researchers. However, a recent comparative study [10] between various wavelets (orthogonal, bi-orthogonal and tree structure) and Gabor transform compelled us to adapt the later one for its better performance. Essentially in our study, a multi-channel filtering approach based on the combined HVS and Gabor transform, is discussed

We apply cortex filters within a square block of data. 'Average energy' calculated on the filtered image is regarded as texture feature. Moving the window throughout the entire image and calculating texture features for each we obtain a set of feature images. The feature set so obtained is subsequently classified using popular k-means clustering technique.

The remainder of the paper is organized as follows. In section 2, we explain a simple HVS with detail formation of cortex filters. In section 3, we describe the classification system that uses k-means clustering. In section 4, we present the results and comparison with GLCM approach. Finally we discuss the implications in section 5 and conclude the work through section 6.

2. Human Visual System (HVS)

HVS decomposes the retinal image into a number of filtered images, each of which contains intensity variations over a narrow range of frequencies (size) and orientation. According to various psychophysical experiments [2], it is observed that the radial frequency bandwidth of the simple cells of HVS varies between 0.5 to 2.5 octaves with an average of 1.4 octaves. The average orientation bandwidth of these cells is about 45 degree with a large spread from cell to cell. Such mechanisms of HVS are often referred to as channels, which are interpreted as band pass filter.

2.1 Filter description

Channels are specified by a set of band pass cortex filters in the frequency domain. 2D gaussian functions are used to achieve optimal resolution as explained in [7]. The filters are described as follows:

2.1.1 Radial Band Filter

The radial band filter $f(r)$, described by 2D Gaussian function, is used to extract the radial property of texture in digital images as follows:

$$f(r) = \frac{1}{\sqrt{2\pi}\sigma_r}\exp\left\{-\frac{(r-m_r)^2}{2\sigma_r^2}\right\}, \quad (1)$$

$$m_r = Int\,[(r_1 + r_2)/2], \quad (2)$$

$$\sigma_r = \sqrt{\left\{-\frac{(r_1 - m_r)^2}{2\log_e(R_r)}\right\}}, \quad (3)$$

$$R_r = \frac{f(r_1)}{f(m_r)}, \quad (4)$$

$$r = \sqrt{u^2 + v^2} \quad (5)$$

where Int[.] calculates the integer value of the argument, r is the frequency variable in polar co-ordinate, r_1 and r_2 are the lower and upper boundary of frequency variable. m_r, σ_r are the mean and effective width and R_r is a ratio parameter of the Gaussian. The frequency spectrum of a sample radial band filter is shown in **Fig. 1**.

2.1.2 Orientation Filter

This filter is also Gaussian type in the angular direction θ and is used to extract the orientation features of digital image texture,

$$g(\theta) = \frac{1}{\sqrt{2\pi}\sigma_\theta}\exp\left\{-\frac{(\theta - m_\theta)^2}{2\sigma_\theta^2}\right\}, \quad (6)$$

$$m_\theta = (index)\pi/4 + \pi/8, \quad (7)$$

$$\sigma_\theta = \sqrt{-\frac{(\pi/8)^2}{2\log_e(R_\theta)}}, \quad (8)$$

$$R_\theta = \frac{g(\pi/8 + m_\theta)}{g(m_\theta)}, \quad (9)$$

$$\theta = \tan^{-1}\left(\frac{v}{u}\right) \quad (10)$$

where m_θ and σ_θ are the mean and effective width of the gaussian respectively and ***index*** represents integers having a range 0-3. θ is a frequency variable in polar co-ordinate. In general θ can be limited to any two values θ_1 and θ_2. To be consistent with the experimental evidence of cortex transform, we will allow the filter to cover 45°. Thus the different bands are 0°→45°, 45°→90°, 90°→135° and 135°→180°. In addition they are symmetrical to inversion, e.g., the filter containing the 0°→45° angular frequencies also contain 180°→225°

angular frequencies. Using indices in **Fig.2** we can calculate m_θ by Eq. (7). The frequency spectrum of the orientation filter is shown in **Fig. 3**.

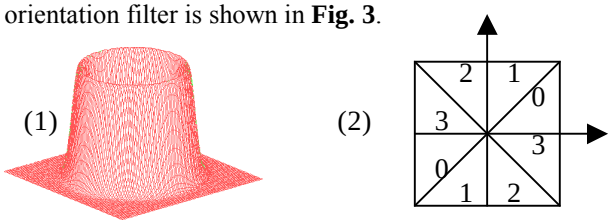


Fig. 1: Frequency Spectrum of Radial band filter.

Fig. 2: Indices for the orientation frequency plane.

2.1.3 Cortex Filter

We obtain cortex filter by the product of two Gaussian filters described above. That is

$$C(r,\theta) = f(r)g(\theta). \qquad (11)$$

This is a frequency domain filter used to compute the textural properties of images at different radial bands and orientations. The frequency spectrum of a sample cortex filter is shown in **Fig. 4**.

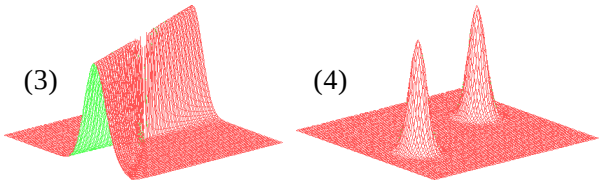


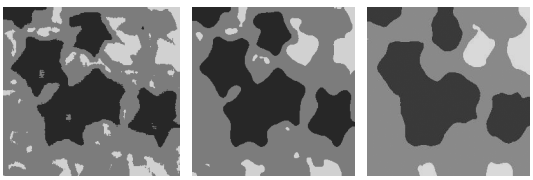
Fig. 3: The frequency spectrum of orientation filter.
Fig: 4: Frequency response of Cortex filter.

3. Classification System

A detail classification system consists of the following sub-sections.

3.1 Window selection

In our analysis, window size is very important as it determines the feature space dimension.



(a) Window (8), (b) window (16), (c) window (32)

Fig. 5: Effect of window sizes on class boundary.

However, at present, the size of the window is determined as (16x16) by trial and error method based on the principle of preserving undistorted class boundary. **Fig. 5** shows how various window sizes affect the class boundary.

3.2 Filter selection

We choose an octave scale during filter selection within the image block. This is justified by the radial frequency bandwidth of the simple cells in the visual cortex of HVS. According to this scale, number of cortex filters,

$$\textbf{NOF} = 4\log_2[(C/2)/IR] \qquad (12)$$

where C and IR represent image width and initial radial frequency. For a square window of size (128x128), number of cortex filters calculated as 16, where IR is 4 units. Two gaussian isotropic filters may also be used to include information relating to the coarsest and finest texture in the block. Therefore the total number of filters becomes (NOF+2). In our example it is 18.

3.3 Filter kernel generation

We constructed the filter kernel using equations (1)-(10) with the help of following equation.

$$u = i - N/2, \quad v = N/2 - j, \qquad (13)$$

where i, j are the column and row indices for a (NxN) window same as the sliding image block.

Table: 1 Filter Kernel for (8x8) window and for a ratio $R_r / R_\theta = 0.01$ and initial radius 2 units.

v,u	-4	-3	-2	-1	0	1	2	3
4	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0.220	0
1	0	0	0	0	0	0.218	2.848	0
0	0.038	.0004	0	0	0	.0004	0.038	.0004
-1	2.848	0.218	0	0	0	0	0	0
-2	0.220	0	0	0	0	0	0	0
-3	0	0	0	0	0	0	0	0

Table 1 shows the filter kernel for Cortex filter constructed and used in our implementation.

3.4 Feature Images formation

Feature images are computed based on the average energy on the filtered images computed as

$$E_k = \frac{\sum_{(u,v)\in W} |F(u,v)|^2}{N^2}, \qquad (14)$$

where N^2 represents number of pixels within the window W, $F(u,v)$ is the filtered image in the frequency domain, and k indicates the number of feature images. Algorithm for feature image calculation is as follows:

Step 1: A square block of data (corresponding to first pixel) is windowed from the original image, take Fourier transform, and compute magnitude spectrum with frequency scaling. **Step 2:** Determine the number of filters and compute their kernels. **Step 3:** Multiply each kernel with the transformed data to obtain filtered images and then calculate and store average energy from each filtered image as first data element on various 2D arrays.

Step 4: Repeat step 1 to step 3 till the scanning of whole image in a step of 1 pixel shifting of the window in horizontal and vertical direction.

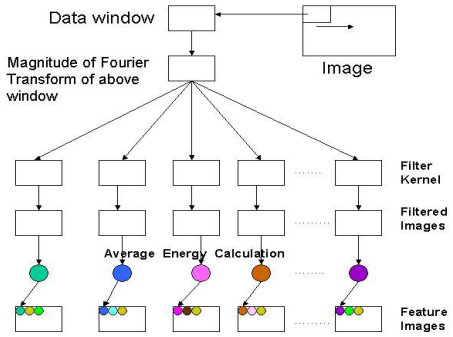


Fig. 6: Feature images formation

Fig.6 shows the block diagram of the technique described.

3.5 Feature images integration

3.5.1 Unsupervised

Popular **K-means Clustering** method is used for unsupervised classification. The number of decision regions or clusters k in this method is determined by the visual inspection of the original image. However

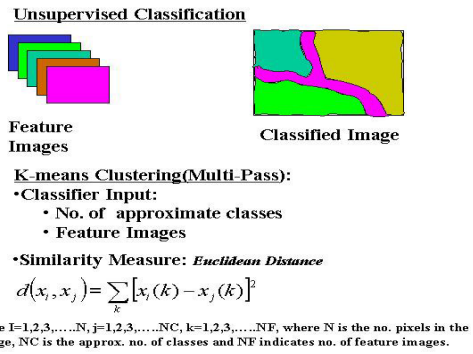


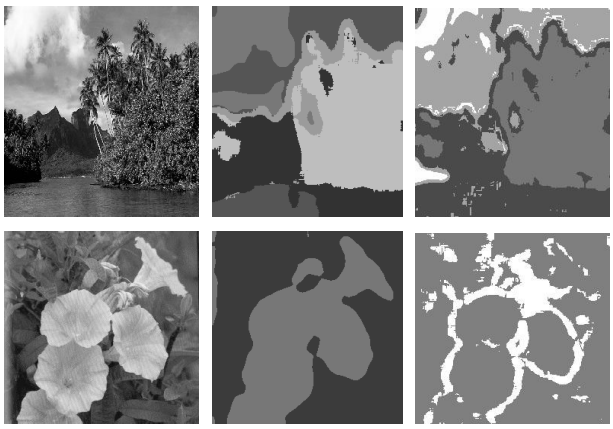
Fig.7: Unsupervised Classification

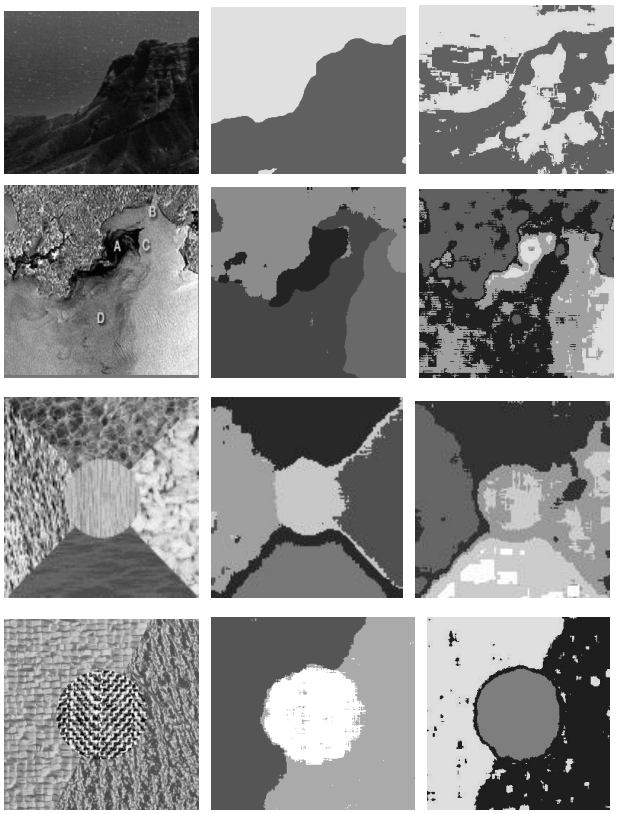
this algorithm can be extended for iterative selection of k based on some error computation criterion. The classifier uses euclidean distance as a similarity measure approach. We used a value 0.01 as the threshold. **Fig. 7** shows the pictorial representation of classification algorithm.

4. Results and Comparison

4.1 Results

Our algorithm is tested on images from camera, radar sensors and Brodatz's album and appears to be promising with respect to classification accuracy. **Fig. 8** shows a set of original and classified images (names are given in table 4) obtained from the cortex filtering and GLCM





(a) (b) (c)

Fig. 8: (a) Original images (256x256) (except mosaic#1, 128x128) (b) Classified images from unsupervised method (c) Classified images from GLCM approach

methods. From the scene image it is observed that our algorithm can distinguish between cloudy and non-cloudy sky, soil, water and vegetation whereas GLCM approach failed to do so. Flower image is perfectly classified as leafy background and flowers as a foreground objects but GLCM can only extract the approximate flower boundary. In the case of radar (oily & valley water) and mosaic images, a clear distinction is observed too between our cortex filtering and popular GLCM approaches. Specifically classification of oily water image as severe (black), lesser, and non-oily regions clearly demonstrate the potentiality of our approach.

4.2 Comparison with GLCM approach

We compared the classification results of our cortex filtering approach with that of Haralick's [1] gray level co-occurrence matrix (GLCM) approach using confusion matrix analysis. Four practically significant features namely **ENERGY, CONTRAST, HOMOGENEITY,** and **ENTROPY** are used with one pixel co-occurrence length and (13x13) sliding window in our analysis. We compute confusion matrices for a number of samples from the classified images where the sample positions were fixed by visual interpretation of the original images. Ideally main diagonal elements of the error matrix are non-zero.

Table 2: Confusion matrix for mosaic-2 image by Filtering Approach.

	C1	C2	C3	Total	ICA (%)	OA (%)
C1	768	0	0	768	100.0	99.00
C2	0	768	0	768	100.0	
C3	0	23	745	768	97.00	
Total	768	791	745			

Table 3: Confusion matrix for mosaic2 image from GLCM approach.

	C1	C3	C3	Total	ICA (%)	OA (%)
C1	741	0	27	768	96.48	94.3
C2	0	768	0	768	100.0	
C3	104	0	668	768	86.98	
Total	845	768	695			

Table 4: Comparison of overall classification accuracy.

Images	Sample pixels	Overall Accuracy (%)	
		CT	GLCM
Scene	3840	87.6	67.76
Flower	1536	100	72.92
Oily water	2048	100	61.70
Mosaic #1	1280	97.85	62.03
Mosaic #2	2304	99.0	94.3

However, practically some entries in the off-diagonal elements are observed. The individual (*ICA*) and overall class accuracy (*OA*) are calculated using Eq. (15) and Eq.(16).

$$ICA = \frac{C_{ii}}{\sum_{j=1}^{n} C_{ij}} \quad (15)$$

$$OA = \frac{\sum_{i=1}^{n} C_{ii}}{\sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}} \quad (16)$$

where $C_{ij}'s$ are the various entries of confusion matrix. **Fig. 8** and **Table 2-4** show the superiority of our approach compared to GLCM approach.

5. Implications

In our implementation**,** we selected the window size as (8x8) or (16x16) by trial and error to compromise between the number of filters and accurate class boundary. However, image frequency analysis may allow automatic window size selection. Initial radial frequency, *IR* and filter design parameter ratio, $R_r \, or \, R_\theta$ are observed to be suitable between 2-4 units and 0.01-0.03 for various types of images. At present, we did neither use overlapping of filter kernels nor include the phase information of images. But the inclusion of such features will surely improve the overall performance. However due to pixel-wise operation our method takes about 15 minutes to obtain a classified image of size 256x256 using Pentium III, 800 MHz, 128 MB RAM machine.

6. Conclusion

We have explored a texture-based image classification algorithm. Our algorithm reduces feature space by proper selection of the window size (trial and error) and number of filters within it. This also avoids an optimal filter selection scheme, which is generally an inevitable step in any multi-scale analysis.

Though the feature vector contains limited spectral information due to scale reduction, our approach permits the use simple texture measure like average energy.

However, the effect of noise and non-uniform illuminations has not taken into consideration in our study. Therefore automatic window selection and texture analysis in the noisy environment are our future concern.

References

[1] R.M. Haralick, K. Shanmugam and I. Dinstein, “Textural features for Image classifications”, IEEE Trans. Syst. Man. Cybern., vol.3, pp. 610-621, 1973.

[2] Anil K. Jain and farshid Farrokhnia, “Unsupervised Texture segmentation using Gabor Filters”, Pattern Recognition, vol. 24, no. 12, pp. 1167-1186, May 1991.

[3] Chung-Ming Wu and Yung-Chang Chen, “Statistical feature matrix for texture analysis”, CVGIP: Graphical Model and Image Processing, vol. 54, no. 5, pp. 407-419, Sep. 1992.

[4] Dennis Dunn and William E. Higgins, “Optimal Gabor Filters for texture segmentation”, IEEE Trans. Image Processing, vol.4, no.7, pp.947-963, July 1995

[5] Chun S. Lu, Pau C. Chung and Chih F. Chen, “Unsupervised Texture segmentation via Wavelet Transform”, Pattern Recognition, vol. 30, no. 5, pp. 729-742, 1997.

[6] Pasaresi, M., “Texture Analysis for Urban Pattern Recognition using Fine-resolution Panchromatic Satellite Imagery, “Geographical and Environmental Modelling”, vol. 4, no.1, pp.43-63, 2000.

[7] A. Papoulis, System and Transforms with Application in Optics, McGraw-Hill, New York, 1968.

[8] A.P. Pentland, “Fractal based description of natural scenes,” IEEE Trans. Patt. Recog. And Mach. Intell., Vol.6, No.6, pp.661-674, 1984.

[9] R.M. Haralick, “Statistical and structural approaches to texture”, Proceedings of THE IEEE, Vol. 67, no. 5, pp.787-804, 1979.

[10] W.Y. Ma and B.S. Manjunath, “A comparison of wavelet transform features for texture image annotation,” In Proc. IEEE Int. conf. on Image Proc., 1995.