Directional Analysis of Image Textures

Ioannis Kontaxakis, Manolis Sangriotis, Drakoulis Martakos
konti@di.uoa.gr sagri@di.uoa.gr martakos@di.uoa.gr
National and Kapodistrian University of Athens
Dept. of Informatics and Telecommunications
Panepistimiopolis, Ilissia, Athens 15784

Abstract

A new approach for texture feature extraction is described, appropriate for unsupervised texture segmentation applications, based on the directional filtering of an image. A new construction of directional filters suitable for Image Analysis which decompose the given image to a set of channels (subimages), each one containing an isolated angular section of the initial image spectrum, is proposed. An envelope detection method is used to estimate the energy locally for each pixel of these channels and the constructed feature vectors are clustered by an unsupervised technique based on an Expectation Maximization Algorithm.

Keywords: Texture Analysis, Image Segmentation, Expectation Maximization

1. Introduction

Texture is the term used to qualify the surface of a given object and is undoubtedly one of the main features used in image processing and pattern recognition. An important task in many image analysis applications is the unsupervised texture segmentation of a picture into homogeneous texture regions. An effective and efficient texture segmentation method is desired in applications like the analysis of aerial images, biomedical images and the automation of industrial applications. Like other segmentation problems, the texture segmentation requires the identification and use of proper texture-specific features with high discriminatory power.

Texture features related to the spectrum of the image can be extracted by using a filterbank consisting of filters with non-overlapping pass-band areas which analyse the initial image $\mathbf{x}$

$$\mathbf{x} = \left\{ x\left(m,n\right), 1 \le m \le \mathrm{M}, 1 \le n \le \mathrm{N} \right\}$$

into a series of 2-D signals, $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_L$,

$$\mathbf{y}_i = \left\{ y_i\left(m,n\right), 1 \le m \le \mathrm{M}, 1 \le n \le \mathrm{N} \right\}, i = 1,2,\ldots,\mathrm{L}$$

called channels or subimages. These L channels are of the same magnitude as the initial image **x** but each of them contains only a distinct spectrum area, A_i, of **x**. Utilizing these L channels, a feature vector with L components can be constructed for each pixel $x(m,n)$ of the initial image, **x**. Each component is related to the pixel energy content in the spectrum area A_i and can be evaluated by detecting the envelope value of the $\mathbf{y}_i$ in the corresponding pixel position $y_i(m,n)$, as Laine and Fan [1] did. A good detection of the envelope can be realised by taking the square of the $\mathbf{y}_i$ and smoothing the result by a properly chosen baseband filter (Figure 1).

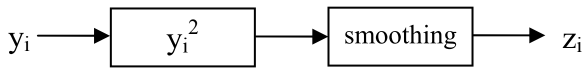


Figure 1. The envelope containing the profile of the spectrum area A_i is obtained by squaring the channel y_i and smoothing the result.

There are many approaches in constructing such filterbanks, with the Gabor series and the Discrete Wavelet Frames the most well known among them.

In the present work a new analysis method, based on a filterbank constructed with directional filters (DF), is proposed (the Directional Filter Analysis-DFA). When an image **x** is analysed by the DF filterbank, each resulting channel has the same size as **x** and contains a distinct wedge-shaped spectrum area placed radially, like the ones depicted in Figure 2. The features extracted from these channels are used either alone or combined with other features extracted from DWF analysis and can form a powerful discriminating feature vector.

The rest of this work is organised as follows: In Chapter 2 the technique for constructing 2-D directional filters is shortly reviewed. In Chapter 3 a new set of directional filters is proposed, suitable for the directional analysis of an image and the construction of equal-sized 2-D channels. In the same chapter the proposed method for the features extraction of the image is described. In Chapter 4 a clustering method based on the Expectation Maximization (EM) algorithm is presented and the unsupervised texture segmentation of an image is described. Finally in Chapter 5 some experiments with texture classification and segmentation are presented and their results are compared with existing classification techniques.

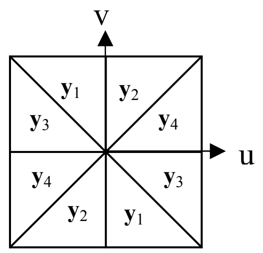


Figure 2. The directional analysis results in 4 channels of the same size with the image, each one containing a wedge shaped area of the initial spectrum.

2. The directional analysis technique

Bamberger and Smith [2] give the basic ideas for the construction of directional filter banks. A detailed description can be found in the work of Do [3]. Here, only the essential theory concerning the directional 2-D filtering is included.

Given a 2-D sampling lattice, a 2×2 non-singular integer matrix M can define a sublattice SBL(M) as

$$\mathrm{SBL(M)} = \left\{\mathbf{m} : \mathbf{m} = \mathrm{M}\cdot\mathbf{n},\ \mathbf{n}\in\mathbb{Z}^2\right\} \tag{1}$$

For a given 2-D signal $x(\mathbf{n})$ the downsampled version $x_d(\mathbf{n})$ of an M-fold downsampling, is derived by

$$x_d[\mathbf{n}] = x[\mathrm{M}\cdot\mathbf{n}] \tag{2}$$

and it contains all the samples of $x(\mathbf{n})$ which lie on the SBL(M). Several authors, for instance Karlsson and M. Vetterli [4], have shown that the relation between the 2-D Fourier Transform, $X_d(\boldsymbol{\omega})$, of the downsampled version $x_d(\mathbf{n})$, and the initial signal's transform, $X(\boldsymbol{\omega})$ is given by

$$X_d(\boldsymbol{\omega}) = \frac{1}{|\det(\mathrm{M})|}\sum_{\mathbf{k}\in\mathrm{N}(\mathrm{M}^T)} X(\mathrm{M}^{-T}\boldsymbol{\omega} - 2\pi\mathrm{M}^{-T}\mathbf{k})\cdot \tag{3}$$

Here, $N(M)$ represents the set of integer vectors of the form $M\mathbf{t}$ where $\mathbf{t} \in [0,1) \times [0,1)$. It can be shown that the number of elements in $N(M)$ is equal to $|\det(M)|$.

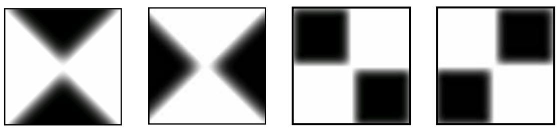
The upsampled version $x_u(\mathbf{n})$ of a 2-D signal $x(\mathbf{n})$ of an M-fold upsampling and the corresponding 2-D Fourier Transform is derived by

$$x_u(\mathbf{n}) = \begin{cases} x[M^{-1}\mathbf{n}], & \text{if } \mathbf{n} \in \mathrm{LAT}(M) \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

$$X_u(\boldsymbol{\omega}) = X(M^T \boldsymbol{\omega}) \tag{5}$$

Two fundamental structures of directional filters are the fan and the quadrant filter types. The frequency responses of these two filter types are shown in Figure 3(a) and 3(b) respectively. A fan filter can be constructed from a simple 2-D lowpass filter with a frequency response like the one in Figure 4(a). This lowpass filter is M-downsampled according to (2), where the non-singular matrix M is given by

$$M = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{6}$$



(a) (b)

Figure 3. Frequency Responses of: (a) a pair of fan filters and (b) a pair of quadrant filters.

By applying (3) it can be easily shown that the resulted M-downsampled version has a diamond-shaped frequency response like the one shown in Figure 4(b). By modulating this filter mask by π, in either the one or the other frequency variable, the pair of fan filters in Figure 4(c) and 4(d) is obtained. The pair of quadrant filters is implemented by M-upsampling a pair of fan filters. The matrix used for the upsampling is the one of (6), the upsampling procedure is described by (4) and the resulted quadrant frequency response is assured by applying (5).

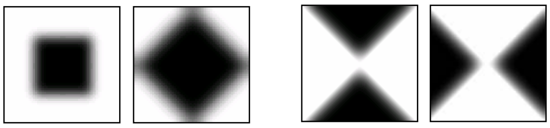


Figure 4. Frequency Responses of: (a) a Lowpass filter, (b) a lowpass filter after M-downsampling (diamond response), (c) and (d) a diamond response after π-modulation in one frequency variable.

3. The DFA procedure

Various techniques for the construction of directional filters can be found in international literature, like, for instance, Park et al [5] and Bamberger and Smith [2]. However, all these techniques are derived from the coding applications area, where the samples of the obtained analysis channels are decimated, so these channels and the initial image no longer have the same size. Features extracted from such channels are not shift invariant and cannot be used for a reliable unsupervised texture segmentation procedure.

In this paper a simpler Directional Filter Bank dividing the initial spectrum into the four regions showed in Figure 2, is proposed. In Figure 5 the construction of the corresponding filters using the convolution of a fan and a quadrant filter, is shown. Combining the two types of fan with the two types of quadrant filters, the four directional filters are obtained. In general, directional filters have poor response in the very low frequencies area; they cannot resolve it so this area must be removed from the initial image. For this reason a circularly symmetric 2-D filter, with a bandstop response in the very low frequency area, is used in order to eliminate a narrow region around $(0,0)^T$ frequency. The outline of the DFA procedure is shown in Figure 6.

The feature vector components are formed using an envelope detection technique which is a two-stage process. The first stage includes squaring (or rectification) of each analysis channel while during the second stage, smoothing of each channel is performed. For the smoothing stage b-spline filters are employed.

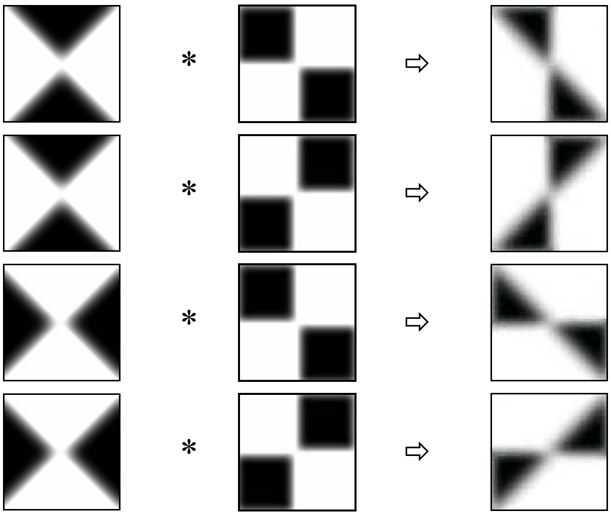


Figure 5. Frequency responses of the DFA filterbank

In order to suitably remove the generation of transient regions at the boundaries of the initial image, during the convolution (filtering), a periodic-symmetric expansion of the image is undertaken, before the analysis takes place.

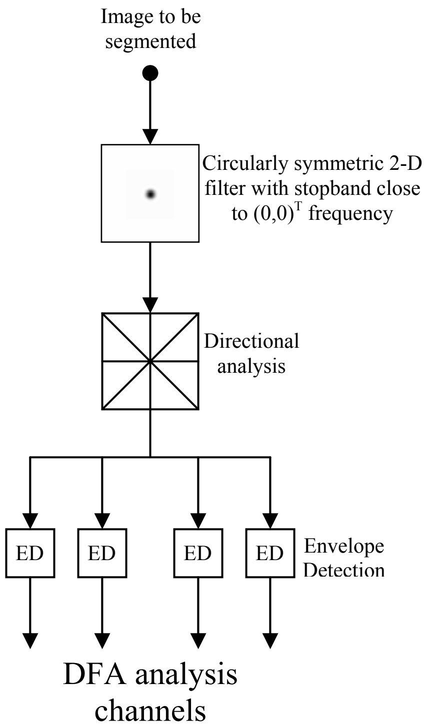


Figure 6. The DFA method.

4. Clustering

The segmentation procedure is completed with the unsupervised classification of the feature vectors $\mathbf{x}_i, i = 1, 2, \ldots, \mathrm{N}$ into the J classes, $C_1, C_2, \ldots C_J$, corresponding to the same number of discrete texture regions. From the feature generation procedure it is anticipated that feature vectors form compact clusters.

In the present work, when J=2 or J=3, the unsupervised classification has been based on the Expectation Maximization Algorithm. This method has been suggested in the recent scientific literature as in the work of Pereira et al [6] and Sanjay-Gopal and Hebert [7]. Although no feature reduction is performed its complexity remains comparable to the MDA procedure and the 2-D clustering technique associated with it. Moreover, using the EM method the classification that follows is based on the well-known Bayes classification rule, mentioned in the next paragraph, which results in minimum probability of segmentation error.

According to the EM method the distribution of feature vectors is considered to be a mixture of J distributions, each one characterized by a PDF, $p(\mathbf{x} \mid C_j)$, and a probability, P_j, $j = 1, \ldots, \mathrm{J}$. If the parameters' vector $\mathbf{\Theta}$, consisting of the J PDF parameters along with the P_j, was known, the classification would be implemented based on Bayes classification rule, i.e. $\mathbf{x}_i$ would be classified to the class C_j if and only if it was valid that $P(C_j \mid \mathbf{x}_i) > P(C_k \mid \mathbf{x}_i)$ for every $k = 1, 2, ..., \mathrm{J}$ and $k \neq j$.

Each feature vector $\mathbf{x}_i$ corresponds to a pixel of the analyzed image which, as a matter of fact, belongs to one of the texture regions. Consequently there is an integer label j_i with a value in the interval [1,J] that denotes the actual class to which this vector belongs. Of course no value of these labels is known during an unsupervised classification procedure, which is performed exactly for the value designation of these labels.

In this work the EM algorithm estimates the parameters' vector $\mathbf{\Theta}$ in such a way that the Expected Value of the likelihood function $E\left[p\left(\mathbf{x}_1, j_1, \ldots, \mathbf{x}_N, j_N\right)\right]$ over the $\mathbf{\Theta}$ and the label set j_i, $i = 1, \ldots \mathrm{N}$ is maximized. As $\mathbf{x}_i$ are statistically independent, the initial likelihood function is given as the product

$\prod P(\mathbf{x}_1, j_1, ..., \mathbf{x}_N, j_N)$. However it is valid that $p(\mathbf{x}_k, j_k) = p(\mathbf{x}_k \mid j_k) P_{j_k}$ or, if the loglikelihood function is used:

$$L(\mathbf{\Theta}) = \sum_{i=1}^{N} \ln\left(p\left(\mathbf{x}_i \mid j_i, \mathbf{\Theta}\right) P_{j_i} \right) \tag{7}$$

and by taking the expectation:

$$\begin{aligned} Q\left(\mathbf{\Theta};\mathbf{\Theta}(t)\right) &= E\left[\sum_{i=1}^{N} \ln\left(p(\mathbf{x}_i, j_i) \right) \right] = \\ &= \sum_{i=1}^{N} E\left[\ln\left(p\left(\mathbf{x}_i \mid j_i;\mathbf{\Theta}\right) P_{j_i} \right) \right] = \\ &= \sum_{i=1}^{N} \sum_{j_i=1}^{J} P\left(j_i \mid \mathbf{x}_i;\mathbf{\Theta}(t)\right) \ln\left(p\left(\mathbf{x}_i \mid j_i;\mathbf{\Theta}\right) P_{j_i} \right) \end{aligned} \tag{8}$$

The estimation of $\mathbf{\Theta}$ for the maximization of $Q\left(\mathbf{\Theta};\mathbf{\Theta}(t)\right)$ is performed iteratively until a small enough change of $\mathbf{\Theta}$ is reached and each iteration stage includes two steps (given below with slightly simpler notation):

- E-step: At the (t+1)th stage of the iteration, where $\mathbf{\Theta}(t)$ is available, compute the expected value of $Q(\mathbf{\Theta};\mathbf{\Theta}(t))$, that is:

$$Q\left(\mathbf{\Theta};\mathbf{\Theta}(t)\right) = \sum_{i=1}^{N} \sum_{j=1}^{m} P\left(C_j \mid \mathbf{x}_i;\mathbf{\Theta}(t)\right) \ln\left(p\left(\mathbf{x}_i \mid C_j;\mathbf{\Theta}\right) P_j \right)$$

- M-step: Compute the next (t+1)th estimate of $\mathbf{\Theta}$ by maximizing $Q(\mathbf{\Theta};\mathbf{\Theta}(t))$,

$$\mathbf{\Theta}(t+1) = \arg\max_{\mathbf{\Theta}} Q\left(\mathbf{\Theta};\mathbf{\Theta}(t)\right)$$

The maximization is achieved by evaluating the gradient of $Q\left(\mathbf{\Theta};\mathbf{\Theta}(t)\right)$ and setting it equal to **0.** This maximization with respect to ***P*** is a constraint optimization problem since

$$P_j \geq 0, \quad j = 1, ..., J \text{ and } \sum_{j=1}^{J} P_j = 1$$

In the case of Gaussian distributions, the parameters vector $\mathbf{\Theta}$ includes the mean value and the scatter matrix along with the probability of each class. The above two steps are realized in the next equations, as they were used by the algorithm in the book of Theodoridis and Koutroumbas [8]. In every iteration stage the probability $P(C_j \mid \mathbf{x};\mathbf{\Theta}(t))$ is calculated for every feature vector and every class:

$$P(C_j \mid \mathbf{x};\mathbf{\Theta}(t)) = \frac{\left|\sum_j(t)\right|^{-1/2} \exp\left(-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}_j(t)\right)^T \sum_j^{-1}(t)\left(\mathbf{x}-\boldsymbol{\mu}_j(t)\right)\right) P_j(t)}{\sum_{k=1}^{m}\left|\sum_j(t)\right|^{-1/2} \exp\left(-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}_k(t)\right)^T \sum_k^{-1}(t)\left(\mathbf{x}-\boldsymbol{\mu}_k(t)\right)\right) P_k(t)}$$

and then the new parameters' estimation is performed:

$$\boldsymbol{\mu}_j(t+1) = \frac{\sum_{k=1}^{N} P\left(C_j \mid \mathbf{x}_k;\mathbf{\Theta}(t)\right)\mathbf{x}_k}{\sum_{k=1}^{N} P\left(C_j \mid \mathbf{x}_k;\mathbf{\Theta}(t)\right)}$$

$$\sum_j(t+1) = \frac{\sum_{k=1}^{N} P\left(C_j \mid \mathbf{x}_k;\mathbf{\Theta}(t)\right)\left(\mathbf{x}_k-\boldsymbol{\mu}_j(t)\right)\left(\mathbf{x}_k-\boldsymbol{\mu}_j(t)\right)^T}{\sum_{k=1}^{N} P\left(C_j \mid \mathbf{x}_k;\mathbf{\Theta}(t)\right)}$$

$$P_j = \frac{1}{N}\sum_{i=1}^{N} P\left(C_j \mid \mathbf{x}_i;\mathbf{\Theta}(t)\right) \quad j = 1,\dots,J$$

In order to make the clustering process even faster, the EM algorithm was applied to a much smaller portion of the initial population, formed by taking one out of 20 feature vectors. This reduction did not affect the accuracy of the parameters' estimation. The EM-estimated parameters then fed the Bayesian classifier for the final classification of the whole population.

5. Experimental Results

These experiments have been worked on images consisting of two or three texture samples. Pixels features have been extracted using the envelope detection technique of channels created by the DFA analysis method, as described in Chapter 3. An unsupervised class separation, based on EM clustering parameter estimation, and Bayesian classification have been performed.

The first texture segmentation example is given in Figure 7. The initial image includes two texture samples. The first texture sample presents directional structure along the 0^{o} and the 90^{o} direction, while the second texture type along the 45^{o} direction. The DFA method achieves an accurate distinction between the two texture samples. In Figure 8, a similar example is given, where three texture samples are present, two of them presenting loose directional structure. Again the DFA method achieves an acceptable segmentation.

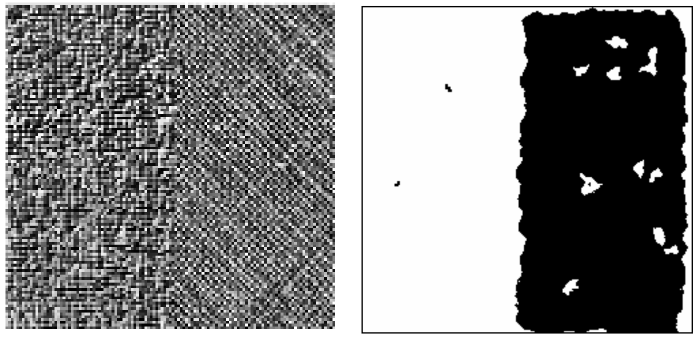


Figure 7. Test image with two texture types and the segmentation results.

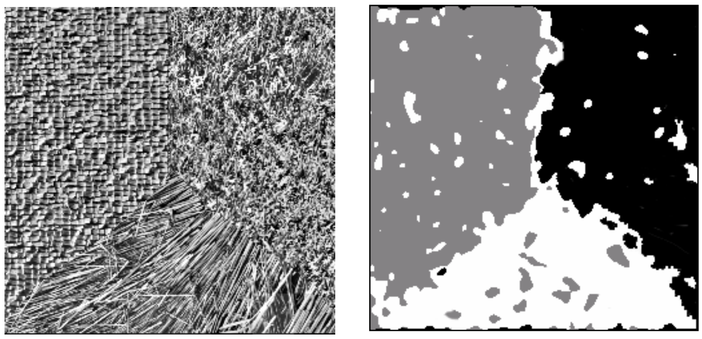


Figure 8. Test image with three texture types and the segmentation results.

6. Conclusions

In this work, a new approach for texture feature extraction suitable for unsupervised segmentation tasks, based on the directional filtering of an image, was described. The new technique achieved a novel spectrum partition, which enabled the exploitation of the directional structure in the image, when present. In general the proposed technique has comparable performance with the DWF method, although the emphasis in the results section is given in cases where the directional structure is present.

7. Acknowledgement

The authors would like to acknowledge the General Secretariat of Research and Technology of Greece for supporting this research in the context of the "i-Byzantine" project.

8. References

[1] A. Laine and J. Fan, "Frame Representation for texture Segmentation" *IEEE Transactions on Image Processing*, Vol. 5, No. 5, May 1996, pp. 771-780

[2] R. H. Bamberger and M. J. T. Smith, "A Filter Bank for the Directional Decomposition of Images: Theory and Design," *IEEE Trans. Signal Processing*, Vol. 40, April 1992. pp. 882–893,

[3] Minh N. Do, "Directional Multiresolution Image Representations", *Ph.D. Thesis*, Department of Communication Systems, Swiss Federal Institute of Technology Lausanne, November 2001

[4] G. Karlsson and M. Vetterli, "Theory of Two - Dimensional Multirate Filter Banks," *IEEE Trans. Acoust., Speech, and Signal Proc.*, Vol 38, No 6, June 1990 pp. 925–937

[5] Chul-Hyun Park, Joon-Jae Lee, Mark J. T. Smith, Sang-il Park, and Kil-Houm Park, "Directional Filter Bank-Based Fingerprint Feature Extraction and Matching," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 14, No 1, January 2004, pp. 74-85

[6] Manuela S. Pereira, Lucio Andrade, Sameh El Difrawy, Barry L. Karger, Elias S. Manolakos, "Statistical Learning Formulation of the DNA Base-calling Problem and its Solution in a Bayesian EM Framework," *Discrete Applied Mathematics*, 104 (2000) pp.229-258

[7] S. Sanjay-Gopal and Thomas J. Hebert, "Bayesian Pixel Classification Using Spatially Variant Finite Mixtures and the Generalized EM Algorithm," *IEEE Transactions on Image Processing*, Vol. 7, No. 7, July 1998, pp. 1014-1028

[8] S. Theodoridis, K. Koutroumbas, *Pattern Recognition*, Academic Press, 1999.