UNSUPERVISED TEXTURE SEGMENTATION USING DISCRETE WAVELET FRAMES

S. Liapis, N. Alvertos, and G. Tziritas
Institute of Computer Science - FORTH, and,
Department of Computer Science, University of Crete
P.O. Box 1470, Heraklion, Greece
E-mails: {liapis,alvertos,tziritas}@csd.uch.gr

ABSTRACT

Image segmentation could be based on texture features. In this work, an unsupervised algorithm for texture segmentation is presented. Texture analysis and characterization are obtained by appropriate frequency decomposition based on the Discrete Wavelet Frames (DWF) analysis. Texture is then characterized by the variance of the wavelet coefficients. The unsupervised algorithm determines the regions to characterize each different texture content in the image. For applying the algorithm, it is necessary to know only the number of the different texture contents of the image. Then, based on a distance measure, each point of the image is classified to one of the different contents.

1 INTRODUCTION

Texture information must be often segmented for recognition purposes in several computer vision tasks, including multimedia applications (e.g., [9]). Different statistical methods have been proposed in the past for texture analysis [2], [4], [6], [11]. Inherent disadvantages with those approaches, such as increased computational cost and irreversibility, can be eliminated using the wavelet transform [8], [10].

The problem of texture segmentation is approached in this paper with algorithms based on the concept of wavelet frames. The aim of the analysis is to determine corresponding characteristics for each texture content so that each is uniquely defined. This analysis is performed in the frequency domain, where the input image is decomposed to different frequency levels using the Discrete Wavelet Frames (DWF). Following deduction of these characteristics, statistical properties are applied to conclude those features necessary to describe and classify the texture content.

The philosophy to this approach has been introduced in the past [12], however, our scheme differs in the statistical methodology for evaluating texture parameters and in the criterion by which a texture point is assigned to a particular subregion of the image to be segmented. Also, in order to evaluate texture parameters, without any given information about the region of each texture (unsupervised), an hierarchical clustering algorithm and a criterion to determine whether a region is homogeneous (has the same texture content) are proposed. Then, each point is assigned to one of the different classes in the image to be segmented by applying a distance measure for each parameter. The number of different texture contents in the image (classes) is provided by the user. The supervised version of this algorithm has been proposed in a previous work [7].

2 TEXTURE CHARACTERIZATION

To decompose the frequency domain of the input signal, a lowpass filter H(z) and its complementary highpass G(z) are used. These filters generate more filters by upsampling with a factor of 2, so that the whole range of bands is covered [12]. The following hold true, respectively:

$$\left.\begin{array}{lcl} H(z) & = & \frac{z^2+4z+6+4z^{-1}+z^{-2}}{16} \\ G(z) & = & zH(-z^{-1}) \end{array}\right\} \quad (1)$$

in the frequency domain. In addition, the generated filters can form orthogonal wavelet base functions [8], so the input signal can be decomposed into discrete wavelet frame coefficients, each corresponding to a different frequency band. The previous decomposition can be extended to 2-D signals (images), by forming wavelet bases which result from the cross product of separable bases in each direction. These (four) base functions deduce the following decomposition algorithm:

$$\left.\begin{array}{rcl} d_{1,i+1}(k,l) & = & [h]_{2^i}(k) * [g]_{2^i}(l) * s_i(k,l) \\ d_{2,i+1}(k,l) & = & [g]_{2^i}(k) * [h]_{2^i}(l) * s_i(k,l) \\ d_{3,i+1}(k,l) & = & [g]_{2^i}(k) * [g]_{2^i}(l) * s_i(k,l) \\ s_{i+1}(k,l) & = & [h]_{2^i}(k) * [h]_{2^i}(l) * s_i(k,l) \end{array}\right\} \quad (2)$$

where (k,l) is an image point, $[\ ]_m$ is upsampling by a factor of m, $d_{1,i+1}, d_{2,i+1}, d_{3,i+1}$ are the details of the $i+1$ layer and s_{i+1} the approximation of the decomposition.

The previous analysis can be applied to texture images, yielding the following representative vector:

$$y(k,l) = < y_1(k,l), \ldots, y_{N-1}(k,l), y_N(k,l) > \quad (3)$$

where each element of $y(k,l)$ has been determined according to the analysis in (2) and the dimension of the vector is $N = 3I + 1$, composed of $N-1$ detail components and the approximation at level I component. The texture content is then characterized by the σ_i^2 variances of the $N-1$ detail components of the representative vector ($i = 1, \ldots, N-1$). This characterization is based on the fact that the mean value of the details, as well as the correlation between different components, could be assumed to be zero.

3 TEXTURE SEGMENTATION

If the characteristic (variance vector) of a texture is known, then an image point can be assigned to this texture according to the following distance measure :

$$d_j(y(k,l)) = \sum_{i=1}^{N-1} \left(\frac{y_i^2(k,l)}{\sigma_{i,j}^2} + \log \sigma_{i,j}^2 \right) \qquad (4)$$

where $y(k,l)$ is the representative vector for the point (k,l), $\sigma_{i,j}^2$ is the i component of variance vector of the texture j. This distance is depicted from the Bayesian classifier, assuming Gaussian class conditional probability density function:

$$d_j(y) = (y - \mu_j)^T \Sigma_j^{-1} (y - \mu_j) + \log(\det(\Sigma_j)) \qquad (5)$$

This measure takes the form of (4) observing that the mean value of the features (μ_j) for each filtered image is zero, since filters are zero mean ($G(1) = 0$), and practically the non diagonal elements of the covariance matrix (Σ_j) are also zero due to the minimal correlation between the detail components, as mentioned earlier.

Thus, the problem of segmentation is reduced to estimating the texture parameters (variance vectors) of the different texture contents of the image to be segmented. Fig.1 illustrates the whole procedure of unsupervised segmentation.

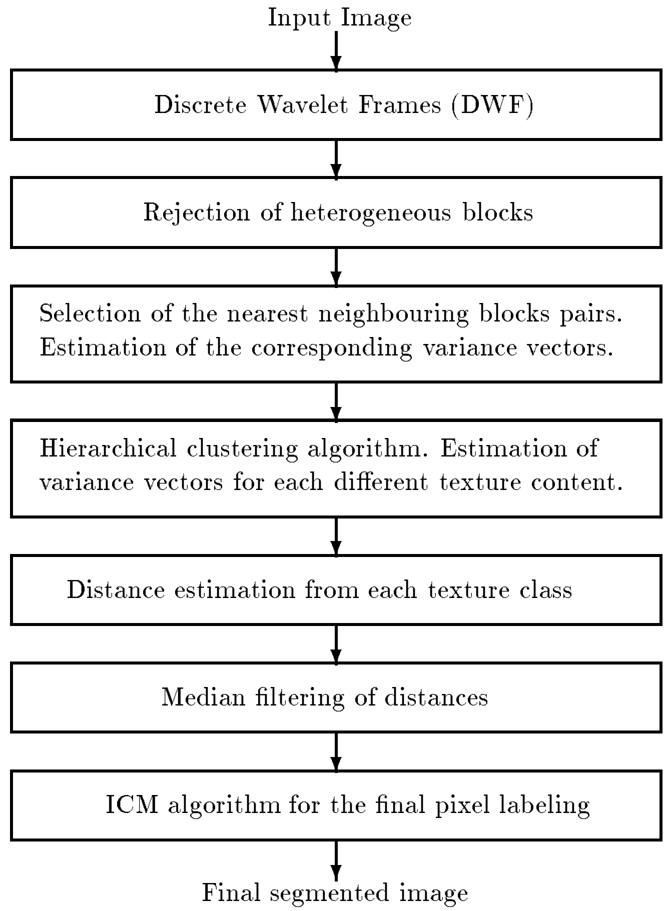

Figure 1: Unsupervised texture segmentation algorithm

3.1 Rejection of Blocks that are Heterogeneous

This algorithm assumes that the number of different texture contents in the image to be segmented is known. The aim of this procedure is to find the regions of the images which yield the best representative texture characteristics (variance vector). For this purpose, the image is divided into blocks of 32×32 pixels. It is evident that the blocks which contain two or more different textures (not homogeneous) are not the ideal ones to estimate the characteristics of a texture. In order to determine whether a block is not homogeneous the following criterion was formed :

$$H_b = \frac{1}{\#pixels} \sum_{p \in block} \sum_{i=1}^{N-1} \frac{(y_{p,i}^2 - \sigma_{b,i}^2)^2}{\sigma_{b,i}^2} \qquad (6)$$

where b is the examined block, p each pixel that belongs to block b, N is the number of the representative vector components, $\sigma_{b,i}^2$ is the variance of the i frequency component of the b block, y_p is the representative vector of a point p in block b. This criterion is based on the idea that, if a block contains different textures, then the mean difference of y_i^2 from the variance σ_i^2 is greater than the mean difference in a block that is homogeneous. This difference at the criterion is valid because the mean value of the details component is zero ($G(1) = H(-1) = 0, \sigma^2 = E\{y^2\}$). The division with $\sigma_{b,i}^2$ is performed for normalization purposes. To conclude that a block is not homogeneous, experimentally is deduced that the sum of H_{b1} and H_{b2} must be less than H_b, where $b1, b, b2$ are neighbouring blocks in the x-axis or y-axis direction.

3.2 Clustering Phase for Estimating Variance Vectors

With the procedure described in the previous section, all blocks that are not homogeneous are rejected. Then, to obtain the best representative blocks from the remaining homogeneous ones, the K nearest neighbouring blocks are selected, where K is one fourth of the total number of blocks in the image. This value of K has been determined empirically. The distance measure between two blocks is described by the following equation :

$$d_{b_1,b_2} = \sum_{i=1}^{N-1} \frac{|\sigma_{1,i}^2 - \sigma_{2,i}^2|}{\sigma_{1,i}^2 + \sigma_{2,i}^2} \quad (7)$$

where N is the number of the representative vector components and σ_1^2, σ_2^2 are the variance vectors of b_1, b_2, respectively. The division in the distance is performed, as previously, for normalization purposes.

Then, from the K pairs, K variance vectors are estimated for each block pair. In order to estimate the variance vectors of the different textures in the image, a hierarchical clustering algorithm [5] is applied to the K variance vectors. This algorithm, in each step, merges the two nearest variance vectors by estimating the new variance vector from the corresponding blocks. Thus, at each step, the number of vectors is reduced by one. The procedure terminates when the number of vectors becomes equal to the number of the different textures in the image to be segmented. The texture parameters are estimated at the end of this procedure.

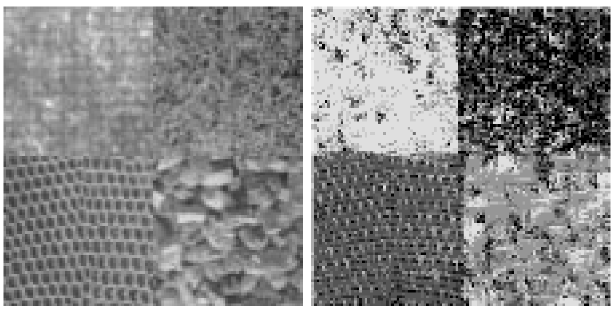

Figure 2: Left: Initial synthetic image from D19, D9, D3, D5 of the Brodatz album. Right: Labeling with distances only.

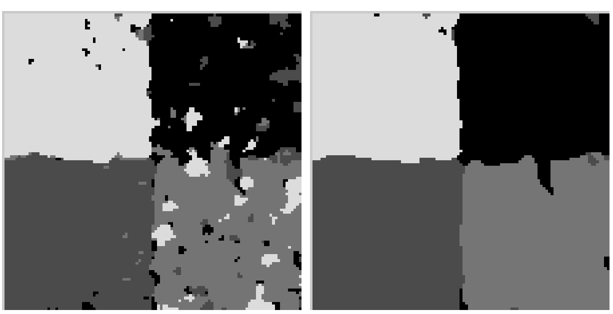

Figure 3: Left: Assignments after smoothing with median filter. Right: Final segmented image after the applying ICM algorithm.

3.3 Pixel Labeling

Having estimated the parameters of each different texture in the image, each pixel in the image is assigned to one texture-class by using the distance measure given by (4). Following these assignments, due to statistical errors on the distance measure, a median filter of a 15×15 pixel window is applied to each distance array of pixels from each texture-class (variance vector). This yields smoothed distance arrays, thus compensating for the statistical errors. After the procedure, some very small regions remain in the labeled image which are removed by applying an iterative algorithm (ICM) [1] for labeling noisy images, where for each pixel assignment the assignments of the neighbouring pixels are considered, according to a Markov random field.

D19	D9	D3	D5
5.96	2.03	2.94	0.95
0.08	0.37	0.96	0.12
0.03	0.06	0.07	0.05
0.89	3.42	6.00	2.29
0.61	1.86	1.98	0.39
0.03	0.05	0.05	0.02
1.24	3.62	7.70	4.62
1.46	3.12	7.60	1.62
0.06	0.19	0.17	0.10
2.61	2.24	12.43	8.47
1.78	2.33	3.30	2.68
0.07	0.19	0.44	0.22

Table 1: Variance vectors for each different texture content for supervised segmentation

D19	D9	D3	D5
6.13	1.75	2.92	0.90
0.08	0.41	0.41	0.14
0.02	0.05	0.06	0.04
0.91	3.42	5.84	2.23
0.61	1.77	1.76	0.34
0.02	0.05	0.04	0.01
1.32	3.42	8.19	4.30
1.47	3.26	7.31	1.26
0.06	0.20	0.16	0.10
3.13	2.01	13.48	9.20
1.76	2.42	2.61	2.67
0.07	0.19	0.42	0.23

Table 2: Variance vectors for each different texture content for unsupervised segmentation

4 EXPERIMENTAL RESULTS

The above algorithm was applied to a synthetic image containing four different textures (Fig.2-left) derived from the Brodatz Album [3]. Fig.2-right illustrates the segmented image with assignments to labels deduced using only the distance measure in eq.4. In Fig.3 (left) the label assignments resulting from distance smoothing using the median filter are shown. The final segmented image after applying the ICM algorithm is presented in Fig.3-right. The initial image was analyzed to 4 frequency levels yielding 12 detail coefficients. In the following tables, the variance vectors of each texture content are presented, both for the supervised (Table 1) and unsupervised (Table 2) procedures. An additional example was considered, where the segmentation algorithm was applied to the image illustrated in Fig.4-left which contains five different texture types. As in the previous example, the label assignments based on the distance measure are shown in Fig.4-right, the result after smooothing of the distance values is presented in Fig.5-left and the result following application of the ICM algorithm is illustrated in Fig.5-right. The decomposed frequency layers in this example were 3, thus producing 9 detail coefficients.

5 CONCLUSION

In this work an efficient method for segmenting images based on the different texture content is presented. Texture is characterized by the variances of the details which are depicted from the Discrete Wavelet Frames analysis. The benefit in using DWF is the improved texture characterization, since it remains invariant under translation and preserves its localization properties. A new unsupervised segmentation algorithm is proposed based on the previous analysis. This algorithm efficiently rejects the heterogeneous blocks of the image, in terms of texture content. The different texture parameters are estimated after applying a hierarchical clustering procedure in the remaining blocks. The described scheme assumes that only the number of different texture types is known.

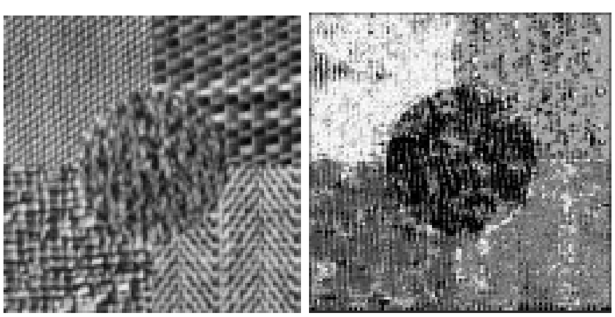

Figure 4: Left: Initial synthetic image from D77, D55, D84, D17 and D24 of the Brodatz album. Right: Labeling with distances only.

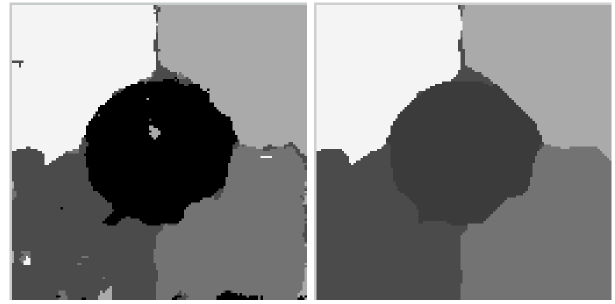

Figure 5: Left: Assignments after smoothing with median filter. Right: Final segmented image after applying ICM.

References

[1] J. Besag. On the statistical analysis of dirty images. *J. of Royal Statistics Society*, 48:259–302, 1986.

[2] A.C. Bovic, M. Clark, and W.S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Patt. Anal. Machine Intell.*, 12:55–73, Jan. 1990.

[3] P Brodatz. *A Photografic Album for Artists and Designers*. Dover, New York, 1966.

[4] P.C. Chen and T. Pavlidis. Segmentation of texture using correlation. *IEEE Trans. on Patt. Anal. Machine Intell.*, 5:64–69, January 1983.

[5] R. Duda and P. Hart. *Pattern Classification and Scene Analysis*. New York: J. Wiley & Sons, 1973.

[6] R.L. Kashyap, R. Chellappa, and A. Khotanzad. Texture classification using features derived form random field models. *Pattern Recognition Letters*, 1:43–50, 1982.

[7] S. Liapis, N. Alvertos, and G. Tziritas. Maximum likelihood texture classification and bayesian texture segmentation using discrete wavelet frames. *Intern. Conf. on Digital Signal Processing*, 2:1107–1110, July 1997.

[8] S. G. Malat. A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Patt. Anal. Machine Intell.*, 11:674–693, January 1989.

[9] A. Pentland, R. W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases. *M.I.T. Media Lab. TR-255*, Nov. 1993.

[10] O. Rioul and M. Vetterli. Wavelets and signal processing. *IEEE Signal Processing Mag.*, 8:11–38, October 1991.

[11] M. Unser. Local linear transforms for texture measurements. *Signal Processing*, 11:61–79, July 1986.

[12] M. Unser. Texture classification and segmentation using wavelet frames. *IEEE Trans. on Image Processing*, 4:1549–1560, November 1995.