Contribution to Image Segmentation and Integral Image Coding

Ioannis N. Kontaxakis[1]

National and Kapodistrian University of Athens
Department of Informatics and Telecommunications
konti@di.uoa.gr

Abstract. Unsupervised image segmentation techniques based on the image textures are studied and developed in this thesis. First, the wavelet filters are considered as the mean for feature extraction and feature vector creation. The creation of the filters is studied as well as the filtered images segmented spectrum. Next, the analysis with directional filters for feature extraction is used. Emphasis is given on the directional filter construction and on the different way of spectrum segmentation for the analyzed images that is achieved. In the next step, after the feature generation, possible ways of feature reduction are considered, ways that do not affect the image separability as it is described by a properly selected separability criterion. Next, known clustering methods and the way they are incorporated in the unsupervised segmentation process are studied. Novel applications of those techniques that point out the contribution of this thesis are presented, and finally, a traversal scheme for the efficient integral image coding is suggested.

Keywords: Image segmentation, feature extraction, feature reduction, clustering, wavelets, directional filters, integral image coding, traversal schemes

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. Moreover, texture is used to identify regions of interest in a scene, for object-based video coding or image content analysis and content-based image retrieval. An important task in many image analysis applications is the unsupervised texture segmentation of a picture into homogeneous texture regions. The process of identifying regions with similar statistical behaviour while separating others is called texture segmentation. 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.

[1] Dissertation Advisor: Manolis Sangriotis, Associate Professor

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 **x**

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

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

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

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).

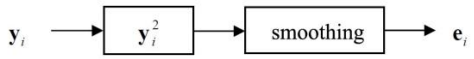

Fig. 1. Envelope Detection

Jain and Farrokhnia [2] used a long series of Gabor filters for the construction of a spectrum sampling filterbank. These filters exhibit a narrow pass-band area that gives the ability to create channels containing a specified small area of the initial image's spatial frequency plane. In Figure 2 the 3dB-contours of such a series of Gabor filters have been designed. From the Gabor series a long feature vector for each pixel of **x** is constructed.

Unser [3] proposed the use of Discrete Wavelet Frames (DWF) instead of the Gabor filter series. In this case, each channel contains a large area of the initial image's spatial frequency plane instead of the narrow areas contained in the Gabor analysis channels. For a great number of applications there is no need for a detailed description of the pixels so the length reduction of the feature vector improves both the algorithmic efficiency and the computational time of the classification procedure. Besides, the computational time for a DWF filterbank analysis of an image is only a small fraction of the time required for a Gabor filter series analysis. In Figure 3, the partitioning of the initial image's spectrum into seven areas by a two-level DWF analysis is shown. However a channel may contain more than one different area of the spectrum with this being a drawback in the case when two or more image areas with symmetric texture construction are to be distinguished. As an example the image in Figure 4 is given. In this example there are two types of texture with symmetric

orientation and the difference in their spectrum cannot be resolved by a DWF filterbank

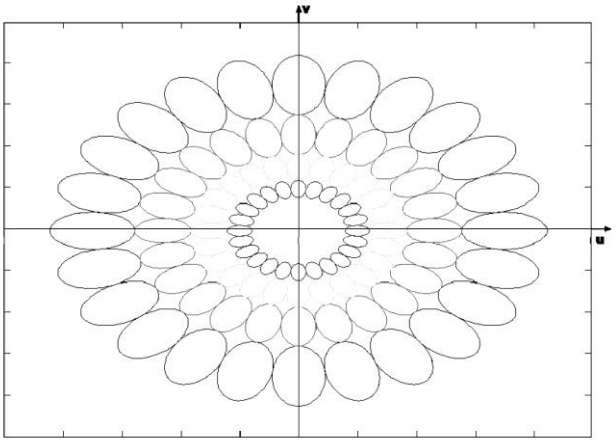

Fig. 2. Gabor Filter Series Spectrum Segmentation

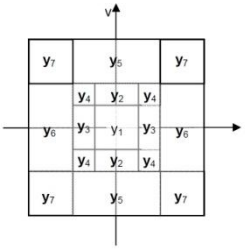

Fig. 3. Wavelet Frames Spectrum Segmentation

2 The Descriptor

Let x(i,j), 0<i,j<M a given image in the region in which different texture samples exist. By applying a real wavelet packet frame transform [3], the subimages $y_m(i,j)$ 0<i,j<M, m=1,2,…,N are constructed. In some of these N images there is a significant difference in the energy in the different texture ranges but in the rest of them, this difference does not exists, or it is rather inadequate. Suppose that in the image x(i,j) there exist two different regions of texture and the first one covers a fraction q of the image surface while the second covers the rest 1-q. After the wavelet transformation of x(i,j), the two textures possess the same pixels in the resulted $y_m(i,j)$ subimages. We can suppose that each one of them has a Gaussian pdf with standard deviations σ and $\lambda \cdot \sigma$ respectively. Zero mean value for both pdfs is imposed by the nature of the wavelet transformation. If $\lambda=1$ in a subimage, The energy of this subimage has similar

values all over it and this subimage cannot offer in the discrimination of the two textures. In contrast, values of λ greater than 2, or less than 0.5 result in different energy values in the two regions and this subimage is selected for a discriminant analysis process.

We propose the following technique for detecting the subimages with proper λ value. The pdf of the whole subimage can be written as:

$$p(x) = \frac{q}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right) + \frac{1-q}{\sqrt{2\pi}\lambda\sigma}\exp\left(-\frac{x^2}{2\lambda^2\sigma^2}\right)$$

As p(x) is a zero mean value pdf, the second μ_2 and fourth μ_4 central moments of the p(x) are evaluated as:

$$\mu_2 = \frac{q}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x^2\exp\left(-\frac{x^2}{2\sigma^2}\right)dx + \frac{1-q}{\sqrt{2\pi}\lambda\sigma}\int_{-\infty}^{\infty} x^2\exp\left(-\frac{x^2}{2\sigma^2}\right)dx$$

and

$$\mu_4 = \frac{q}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x^4\exp\left(-\frac{x^2}{2\sigma^2}\right)dx + \frac{1-q}{\sqrt{2\pi}\lambda\sigma}\int_{-\infty}^{\infty} x^4\exp\left(-\frac{x^2}{2\sigma^2}\right)dx$$

It can be easily shown that:

$$\mu_2 = [q+(1-q)\lambda^2]\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^2\exp\left(-\frac{z^2}{2}\right)dz$$

$$\mu_4 = [q+(1-q)\lambda^4]\frac{\sigma^4}{\sqrt{2\pi}}\int_{-\infty}^{\infty} z^4\exp\left(-\frac{z^2}{2}\right)dz$$

and finally

$$\frac{\mu_4}{\mu_2^2} = 3\cdot\frac{\left[q+(1-q)\lambda^4\right]}{\left[q+(1-q)\lambda^2\right]^2} \quad (1)$$

By plotting relation (1), Figure 4 makes clear that the ratio μ_4/μ_2^2 is a good descriptor for detecting subimages with λ values not near to one.

3 Directional Filters

Various techniques for the construction of directional filters can be found in international literature, like, for instance, Park et al [4] and Bamberger and Smith [5]. 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. Kontaxakis et al [6] propose a Directional Filter Bank construction for the creation of channels of equal size to the analysed image. This bank produces an overall number of eight channels, four for the low and four for the high area of the image spectrum, each one containing a wedge shaped region of the corresponding spectral area.

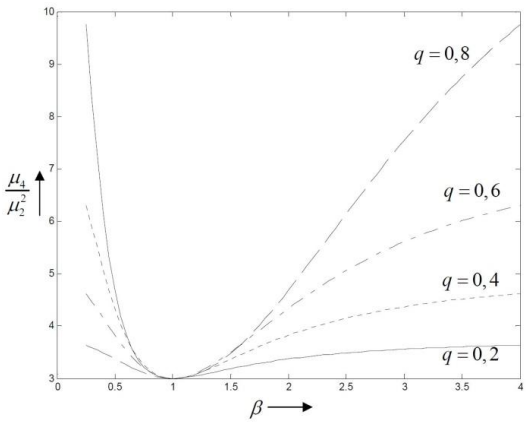

Fig. 4. Parametric Plotting of relation (1)

In this section a simpler Directional Filter Bank dividing the initial spectrum into the four regions showed in Figure 5, is proposed.

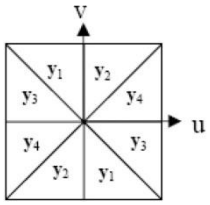

Fig. 5. Directional Filters Spectrum Segmentation

In Figure 6 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 7.

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.

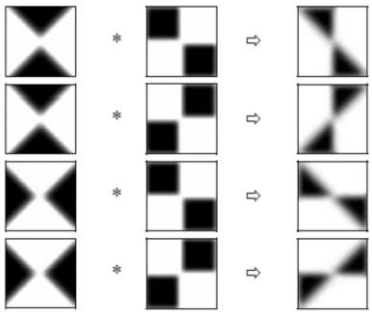

Fig. 6. Directional Filter Bank construction

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.

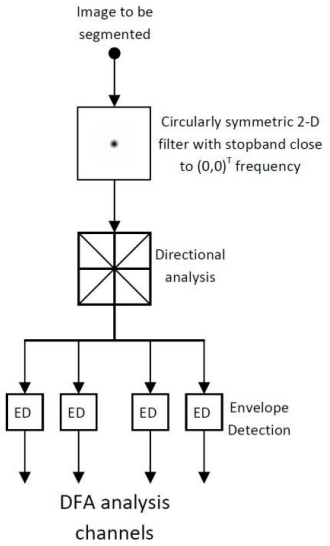

Fig. 7. Outline of the DFA procedure

4 Integral Image Coding

In this section an image traversal scheme based on the Hilbert curve is proposed, in order to increase the efficiency of encoders where EI rearrangement is performed. This is achieved by maximizing a 2D locality property of the EIs that are encoded as a

single entity and therefore increasing the correlation between jointly encoded EIs. As the encoded EIs are reassembled in a static InIm at the display and projected to the viewer, large variations in the quality of different areas in the InIm can reduce the 3D effect. For this reason a quality assessment metric is introduced for evaluating the optimality of a number of different traversal schemes previously proposed in the literature [8,9], taking into account the variability in the quality of different regions of an InIm. A number of different test InIms are used for evaluating the effect of the different traversal schemes for the encoders proposed in [8,9].

There are many types of curves [10] that can be used in order to traverse a set of points placed on a regular 2D grid and transform them to a 1D sequence of points. One of the most discussed properties of these curves is their ability to retain the locality properties of the reordered points. For this reason a number of different measures are used [11] to evaluate their efficiency. In omni-directional InIm coding, these curves are used in order to rearrange the EIs in a 1D stream of images preserving the highest possible correlation among jointly EIs.

In this section the performance of previously proposed traversal schemes [9] along with the one based on the Hilbert curve is evaluated. The different traversal schemes are depicted in Fig. 8

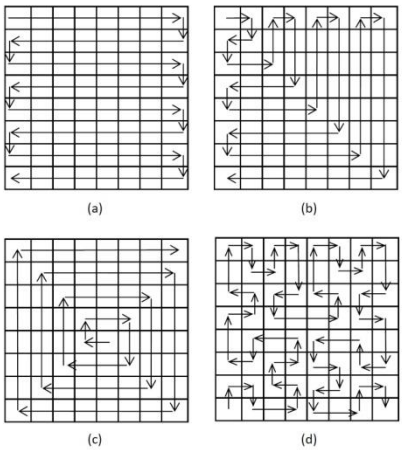

Fig. 8. Traversal schemes for rearranging the EIs of the InIm 2D structure into a 1D stream of EIs. (a) Parallel, (b) Perpendicular, (c) Spiral and (d) Hilbert.

The locality preservation properties of each of these space filling curves can be evaluated by estimating a large number of different metrics as the ones proposed in [11]. The Hilbert curve outperforms all previously proposed curves achieving performance very close to the boundary values of these metrics. In this section we propose an appropriate metric for evaluating the efficiency of the different traversal schemes, which is defined as the ratio between the EI distances in the 2D grid, and their distance in the 1D stream as given by Eq. (2).

$$r_{ij} = \frac{d(e_i, e_j)^2}{|i-j|}, \quad \forall i, j : i > j \quad (2)$$

In this equation *i, j* are the indices of the EIs in the constructed 1D stream and d(·) is their Euclidean distance in the 2D InIm structure. Specifically if (w_i,c_i) and (w_j,c_j) are the row and column indices of the two EIs in the 2D InIm structure then the distance is defined by Eq. (3).

$$d(e_i, e_j) = \sqrt{(w_i - w_j)^2 + (c_i - c_j)^2} \quad (3)$$

Next, we calculate the mean value, $\bar{r}$, and the standard deviation, σ, of this quantity for each traversal scheme. A simulation performed for InIms with different number of EIs, showed that the Hilbert curve achieves the minimum mean and standard deviation values for r_{ij} that are almost invariant regardless of the number of EIs contained in an InIm. On the contrary, as shown in Fig. 9, the corresponding values for the other traversal schemes dramatically diverge as the number of EIs increases. In the same figure, a non linear increase of the standard deviation is observed for traversals other than the Hilbert curve, which increases the possibility for jointly encoding low correlated EIs and thus reducing the overall encoder performance.

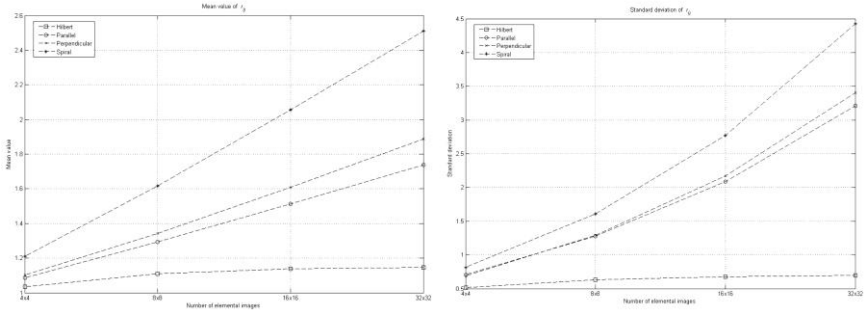

Fig. 9. (a) Mean value and (b) Standard deviation of r_{ij} for different traversal schemes and InIm sizes.

There are primarily two different approaches for evaluating the performance of image encoding techniques based on subjective or objective evaluation metrics [12,13]. The most used metric for objective 2D image quality assessment is the *peak-signal-to-noise-ratio* (*PSNR*) which is given by Eq.(4)

$$PSNR_{dB} = 10 \cdot \log_{10}\left(\frac{255^2}{MSE}\right) \quad (4)$$

where the mean squared error (MSE) is calculated using Eq.(5).

$$MSE = \frac{1}{N \cdot M} \sum_{i=1}^{N} \sum_{j=1}^{M} [I(i,j) - I_r(i,j)]^2 \quad (5)$$

where *I(i,j)* are the intensity values of an 8-bit grayscale *N* x *M* pixels original uncompressed InIm, and $I_r(i,j)$ are the intensity values of the decompressed InIm.

However this measure estimates the overall image quality and do not take into account the special structure of the displayed InIm. In detail, as neighboring EIs are simultaneously displayed to produce the 3D representation, they should exhibit similar quality characteristics in order to produce viewable results. In this case large variations of the *PSNR* values can diminish the 3D effect while *PSNR* values remain high. It is therefore necessary to calculate the fluctuation of the PSNR value over the InIm. In this work the mean *PSNR* value ($\overline{PSNR}$) and its standard deviation (σ_{PSNR}) are calculated using the *PSNR* values found by applying Eq.3 for each of the EIs in the InIm. The values for the $\overline{PSNR}$ and σ_{PSNR} for an InIm assembled of K x L EIs are given by Eqs. (6) and (7) respectively.

$$\overline{PSNR} = \frac{1}{K \cdot L} \sum_{s=1}^{K} \sum_{t=1}^{L} PSNR_{s,t} \quad ...(6)$$

$$\sigma_{\text{PSNR}} = \sqrt{\frac{\sum_{s=1}^{K} \sum_{t=1}^{L} PSNR_{s,t}^2}{K \cdot L} - \left(\frac{\sum_{s=1}^{K} \sum_{t=1}^{L} PSNR_{s,t}}{K \cdot L} \right)^2} \quad (7)$$

Whenever an MPEG-2 strategy is employed to encode a video sequence, the quality of the reproduced sequence varies between frames as a result of the encoding process. In the case of video sequences these fluctuations are not perceived by the observer due to the rapid succession of frames. On the other hand, InIms encoded using an MPEG-2 encoder, also exhibit this kind of *PSNR* variation between EIs resulting in a poor 3D image quality. The use of the Hilbert curve traversal for the InIm decomposition also manages to substantially decrease these variations in regard to previously proposed traversal schemes. In our context we introduce the relative standard deviation, given as $\sigma_{\text{PSNR}} / \overline{PSNR}$ which quantifies the fluctuation of the $\overline{PSNR}$ value for each traversal scheme and acts as an index of homogeneity characterisation for the displayed InIm. The results for all traversal schemes for both $\overline{PSNR}$ and $\sigma_{\text{PSNR}} / \overline{PSNR}$ values for a representative InIm are depicted in Figure 10 (a)-(b) respectively.

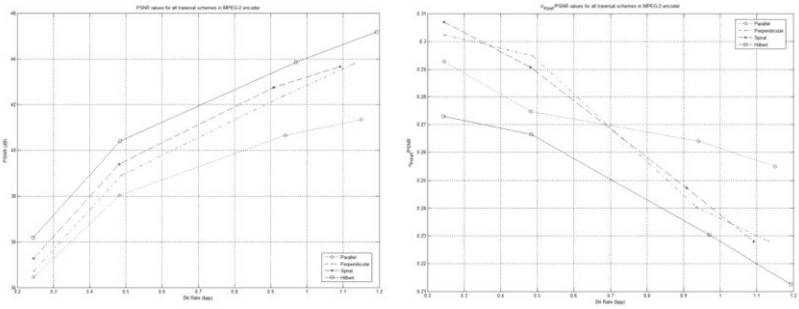

Fig. 10. (a) $\overline{PSNR}$ and (b) $\sigma_{\text{PSNR}} / \overline{PSNR}$ for an encoded InIm as a function of bitrate for the MPEG-2 encoder

References

1. A. Laine and J. Fan, "Frame Representation for Texture Segmentation," *IEEE Transactions on Image Processing*, Vol. 5, No 5, May 1996
2. A. Jain, F. Farrokhnia, "Unsupervised Texture Segmentation Using Gabor Filters," *Pattern Recognition*, 24 (12): 1167 1186, 1991
3. Michael Unser, "Texture Classification and Segmentation Using Wavelet Frames," *IEEE Transactions on Image Processing*, Vol. 4, No 11, November 1995
4. J.-H. Park, Y. Kim, S.-W. Min and B. Lee,"Three Dimensional Display Scheme Based on Integral Imaging with Three Dimensional Information Processing," *Optics Express*, Vol. 12, 2004
5. R. H. Bamberger and M. J. T. Smith, "A Filterbank for the Directional Decomposition of images: Theory and Design," *IEEE Transactions on Signal Processing*, Vol. 40, April 1992
6. I. Kontaxakis, M. Sangriotis, D. Martakos, "Descriptor Detecting the Most Non-Uniform Subimages in Multiresolution Texture Segmentation Procedure," *Electronics Letters*, Vol. 38, No 4, 14th February 2002
7. I. Kontaxakis, E. Sangriotis, D.Martakos, "Directional Analysis of Image Texture for Feature Extraction and Segmentation,"*3rd International Symposium* on *Image and Signal Processing and Analysis* (ISPA 2003), Rome, Italy, September 18-20 2003
8. N. P. Sgouros, D. P. Chaikalis, P. G. Papageorgas and M. S. Sangriotis, "Omnidirectional Integral Photography Images Compression Using the 3D-DCT," in *Adaptive Optics… Topical Meeting on CD-ROM*, OSA Technical Digest (CD), Optical Society of America, 2007 (DTuA2)
9. S. Yeom, A. Stern, B. Javidi, "Compression of 3D Color Integral Imaging," *Optics Express*, Vol. 12, 2004
10. H. Sagan, *Space Filling Curves*, Springer-Verlag 1974
11. Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, Vol. 13, 2004
12. A. Stoica, C. Vertan, C. Fernandez-Maloigne, "Objective and Subjective Color Image Quality Evaluation for JPEG 2000 Compressed Images," *IEEE International Symposium on Signal, Circuits and Systems*, 2003
13. Kontaxakis I., Sangriotis M., Angelopoulou R., Plastira K., Sgouros N., Mavroudi S., "Automatic Analysis of TUNEL Assay Microscopy Images," 7th *IEEE International Symposium on Signal Processing and Information Technology* (ISSPIT 2007), Cairo, Egypt, December 15-18 2007
14. Nicholas Sgouros, Ioannis Kontaxakis and Manolis Sangriotis, "Effect of Different Traversal Schemes in Integral Image Coding," *Applied Optics*, Vol. 47, Issue 19, 2008