HIDDEN MARKOV TREE MODEL APPLIED TO THE DETECTION OF MICRO-CALCIFICATION CLUSTERS IN MAMMOGRAMS

CARLOS S. LIMA
MANUEL J. CARDOSO

Department of Industrial Electronics of University of Minho, Campus de Azurém, 4800-058 Guimarães, Portuga

ABSTRACT

This paper is concerned to the application of a relatively new image texture segmentation algorithm named Hidden Markov Tree (HMT) to the detection of micro-calcification clusters in mammograms. The HMT is a wavelet-based tree-structured probabilistic graph that can capture the statistical properties of the coefficients of the wavelet transform. The aim of this approach is, on the one hand, to take advantage of the wavelet coefficients in the characterization of different textures, and on the other hand, to link these coefficients by a tree structure enabling texture change to be detected. The application of the method was evaluated using the Digital Database for Screening Mammography (DDSM) for training purposes and a sample of the Nijmegen database for testing purposes.

INTRODUCTION

Breast cancer continues to be a significant public health problem in the world. Since the causes of this disease still remain unknown primary prevention is not possible. Early detection of breast carcinomas, which can be reliably achieved by screen-film mammography, is the key to improving breast cancer prognosis. It is well known that the presence of micro-calcification clusters (MCCs) is an important sign concerned to the early detection of breast carcinoma. In fact a high correlation between the appearance of micro-calcification clusters and the disease does exist. Therefore Computer Aided Diagnosis (CAD) systems for automatic detection and classification of MCCs will be very useful and helpful for breast cancer control.

The micro-calcifications appear in small clusters of few pixels brighter than the neighboring pixels. This means that the micro-calcifications correspond to high frequency components of the image spectrum. So, the detection of micro-calcifications can be achieved by decomposing the mammograms into different frequency sub-bands, suppressing the low-frequency sub-band, and reconstructing the mammogram from the sub-bands containing only high frequencies. Wavelet-based sub-band image decomposition [1] is one of the most promising approaches for image filtering and reconstruction given the properties of the wavelet transform where short windows at high frequencies and long windows at low frequencies can be used to provide better signal resolution than the resolution that can be achieved by the conventional Short-Time Fourier Transform (STFT). Related multi-scale analysis based on multi-channel wavelet transform can also be found in [2, 3]. However, the main drawback associated with image filtering techniques is the fact that the wavelet filters becomes very dependent on external factors such as tissue density, breast boundary and background noise. Shortening wavelet filters can increases the micro-calcifications sensitivity at a cost of producing more false positives, which shows a natural limitation of the approach.

Alternative approaches come from statistical modeling of the wavelet coefficients. Independent Mixture (IM) Model [4, 5], Hidden Markov Chain (HMC) Model [6] and Hidden Markov Tree (HMT) Model [7] are examples of statistical modeling of the wavelet coefficients of an image. The IM model is the simplest one, which treats wavelet coefficients as independent random variables, while the HMC model captures dependency within each scale and treats the wavelet coefficients as independents from scale to scale. Finally the HMT model consists of a graph with tree-structured dependencies between wavelet coefficients, which emphasizes natural dependencies between parent and child state variables. Apart from that the HMT model appears as a convenient framework in order to match some important properties of the wavelet transform such as the Clustering and Persistence across Scale. Although the HMT model has been suggested for image processing purposes it has also been used in other signal processing fields [8].

Recently an extension of the HMT model named Hierarchical Image Probability (HIP) model [9] has been suggested. The differences between HMT and HIP are that the last uses different numbers of hidden states at each resolution and models the distribution of local vectors of wavelet coefficients by a multivariate Gaussian mixture. While the HIP approach can model more accurately the wavelet coefficients, however at a cost of increasing in computational load which can be prohibitive regarding practical applications. Additionally the effectiveness of the additional complexity of the HIP model concerning image segmentation purposes is not known yet although it presents obviously additional potential relatively to the HMT approach.

HMT MODEL

The Hidden Markov Tree model provides a flexible stochastic modelling of wavelet coefficients, which are associated with states that represent the signal energy. The approach takes advantage of several properties of wavelet theory regarding to multiresolution analysis and uses some wavelet properties of real signals such as persistency across scale to build a stochastic model that allows capturing the statistical properties of the wavelet transform coefficients. Multiscale segmentation applied to 2-D images means that a square image is sequentially segmented in four square subimages (dyadic squares) of equal size, until blocks of only one pixel are obtained. Each dyadic square can be associated with a sub-tree of wavelet coefficients where the Haar and Daubechie wavelets are the most usual. These blocks which represent image regions are random fields for the which it is required to estimate the class they belong to, in order to achieve image segmentation purposes. This estimation requires a pixel pdf model for each class that is suited to the dyadic squares. This pdf can be learned from examples in the pattern recognition framework.

Figure 1 shows the quad-tree structure of dyadic squares and the 2-D wavelet hidden Markov tree model, where the states are vertically connected across scale in Markov-1 chains to capture the persistence property of the wavelet transform, which occurs for the most practical signals.

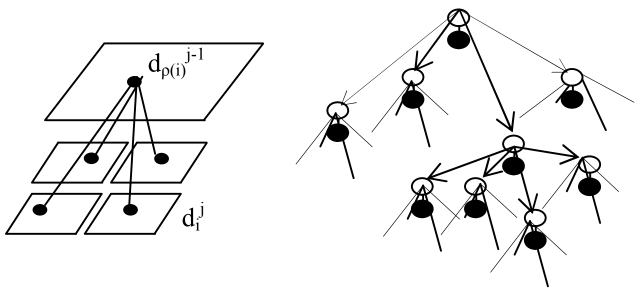


Figure 1- Quad-tree structure of dyadic squares (left) and 2-D wavelet Hidden Markov Tree model.

Figure 1 shows the parent-child dependencies in the tree. Each arrow points from a parent wavelet coefficient to its four children at the next finer scale. Each wavelet coefficient (black node) is modelled as a Gaussian mixture controlled by a hidden state variable (white node). In the HMT model the states are connected vertically across scale in Markov-1 chains, which captures the persistence across scale property of wavelet transforms. It is well known that some algorithms model the wavelet coefficients as independent, which is not accurate for singularity-rich images because of the strong residual correlations between wavelet coefficients. Each state to state link has an underlying state transition matrix that controls statistically the persistence of large and small states down the three. This persistency is related with image singularities which manifest themselves as cascades of large wavelet coefficients through scale. Conversely, smooth regions lead to cascades of small coefficients. This fact is the most important concerning image segmentation.

The HMT algorithm differentiates between large and small wavelet coefficients by associating with each coefficient a binary state variable that controls its size. If the distribution of each coefficient is Gaussian when conditioned on its state models the marginal distribution of each coefficient is a Gaussian mixture.

The wavelet transform of most real-world images consists of a small number of large coefficients and a large number of small coefficients, then the pdf of each wavelet coefficient is well approximated by a two-density Gaussian mixture model. To each wavelet coefficient w_i we associate a discrete hidden state S_i that takes values S (small) or L (large) with probability mass function p_{Si} (m) with m={S, L}. Conditioned on S_i=m, w_i is Gaussian with mean $\mu_{i,m}$ and variance $\sigma_{i,m}^2$. Thus its overall pdf is given by

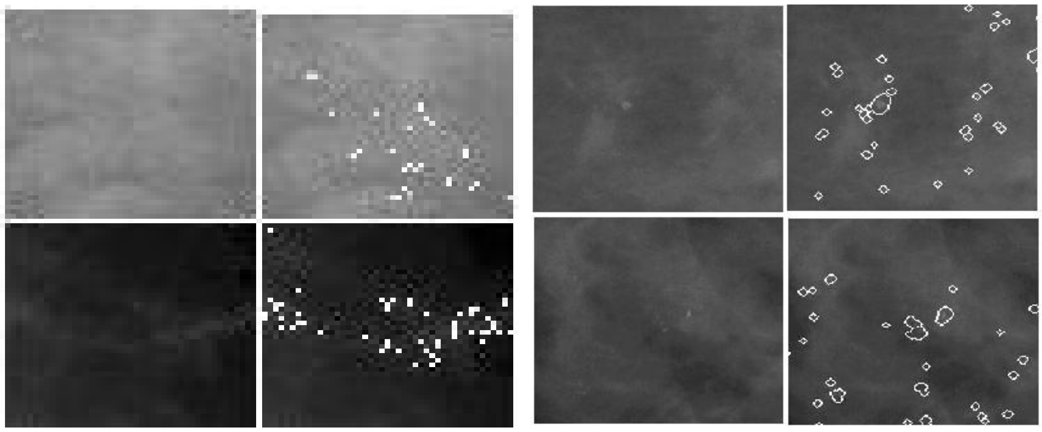
$$f(w_i) = \sum_{m=S,L} p_{S_i}(m) f(w_i / S_i = m) \quad and \quad f(w_i / S_i = m) \approx N(\mu_{i,m}, \sigma_{i,m}^2) \quad (1)$$

Once the model of the marginal density of each wavelet coefficient is established, dependencies between the wavelet coefficients can be modelled by the joint probability mass function of the hidden states. As the coefficients correlations propagate across scales, the dependency structure of the wavelet coefficients in each subband has a quad-tree structure. In the most common case of Haar wavelet transform for each parent-child pair of hidden states $\{S_{\rho(i)}, S_i\}$ the state transitionprobabilities $\varepsilon_{im'}^{\rho(i),m}$ para m=m'=S,L represent the probability for w_i to be small/large when its parent is small/large. In this case the state transition probability matrix is

$$\begin{bmatrix} \varepsilon_{i,S}^{\rho(i),S} & \varepsilon_{i,L}^{\rho(i),S} \\ \varepsilon_{i,S}^{\rho(i),L} & \varepsilon_{i,L}^{\rho(i),L} \end{bmatrix} = \begin{bmatrix} \varepsilon_{i,S}^{\rho(i),S} & 1-\varepsilon_{i,S}^{\rho(i),S} \\ 1-\varepsilon_{i,L}^{\rho(i),L} & \varepsilon_{i,L}^{\rho(i),L} \end{bmatrix} \quad (2)$$

EXPERIMENTAL RESULTS

Experimental results were evaluated by using training images from the Digital Database for Screening Mammography (DDSM) database. Almost all files in the directory labelled “case 1374” from the set of benign pathology were used for training purposes, since when the number of training images is small the use of intra-scale tying is required to avoid overfitting the models. From each image homogeneous regions were picked out for training of the representative textures, respectively microcalcification clusters and normal tissue. Then the algorithm was tested under several images from the Nijmegen database. Figure 2 shows the results of the segmentation of files c02c.pgm, c11c.pgm, c04c.pgm e c08c.pgm.



DISCUSSION

Hidden Model Tree is a framework for Bayesian image segmentation which takes advantage of the properties of wavelet image decomposition such as the properties of persistence and clustering, decorrelation, locality multiresolution and energy compaction. All these properties are known as making part of the most natural images and can be taken into consideration to build elaborated but tractable statistical models. In this context HMT model was tailored for modelling wavelet coefficients in natural images. By concisely modelling and fusing the statistical behaviour of textures at multiple scales the HMT algorithm produces a robust and accurate segmentation of texture images. One important advantage of the HMT algorithm is that as it operates directly on the wavelet coefficients it is able to segment wavelet-compressed images without expanding them to the space domain.

REFERENCES

[1] Wang, T. C. and Karayiannis, N. B. (1998). Detection of Microcalcifications in Digital Mammograms Using Wavelets. IEEE Transactions on Medical Imaging, vol. 17, nº4 pg. 498-509.

[2] Song, D., Qian, W. and Clarke, L. P. (1996). Digital mammography: hybrid M-channel wavelet transform for microcalcification segmentation. 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Amsterdam pp. 1069-1070.

[3] Dippel, S., Stahl, M., Wiemker, R., and Blaffert, T. (2002). Multiscale contrast enhancement for radiographies: Laplacian pyramid versus fast wavelet transform. IEEE Trans. Med. Imag. 21 (4) 343-353.

[4] Pesquet, J.-C., Krim, H., and Hamman, E. (1996). Bayesian approach to best basis selection. IEEE Int. Conf. on Acoust., Speech, Signal Proc.- ICASSP-96, Atlanta pp. 2634-2637.

[5] Chipman, H., Kolaczyk, E. and McCulloch, R. (1997). Adaptive Bayesian wavelet shrinkage. Journal of the American Statistical Association, vol. 92.

[6] Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE vol. 77, pg. 257-286.

[7] Choi, H. and Baraniuk, R. G. (2001). Multiscale image segmentation using wavelet-domain hidden Markov models. IEEE Trans. Image Process., vol. 10, no. 9, pp 1309-1321.

[8] Graja, S. and Boucher, J.-C. (2005). Hidden Markov Tree Model Applied to ECG Delineation. IEEE Transactions on Instrumentation and Measurement, vol. 54 num. 6.

[9] Spence, C., Parra L. C. and Sajda, P. (2006). Varying Complexity in Tree-Structured Image Distribution Models. IEEE Transactions on Image Processing, vol. 15 num. 2.