Statistical Texture Model for mass Detection in Mammography

Modelo estadístico de textura para detección de masas en mamografía

Nicolás Gallego-Ortiz [1]**, David Fernández-Mc-Cann** [2]

[1] Ms.C en Ingeniería. nicgallego@ieee.org

[2]   Ph.D. en Telecomunicaciones. Profesor Asociado. Universidad de Antioquia. Medellín, Colombia. dfernan@udea.edu.co

Key words

Biomedical Engineering, Breast-Cancer, Mathematical Models, Radiodiagnostic, Statistical Methods.

Palabras clave

Ingeniería Biomédica, cáncer de seno, métodos estadísticos, modelos matemáticos, radiodiagnóstico.

Abstract

In the context of image processing algorithms for mass detection in mammography, texture is a key feature to be used to distinguish abnormal tissue from normal tissue. Recently, a texture model based on a multivariate gaussian mixture was proposed, of which the parameters are learned in an unsupervised way from the pixel intensities of images. The model produces images that are probabilistic maps of texture normality and it was proposed as a visualization aid for diagnostic by clinical experts. In this paper, the usability of the model is studied for automatic mass detection. A segmentation strategy is proposed and evaluated using 79 mammography cases.

Resumen

En el contexto de procesamiento de imágenes para detección de masas en mamografía, la textura es una característica clave para distinguir tejido anormal del normal. Recientemente un modelo de textura basado en mezcla de gaussianas multivariadas fue propuesto. El modelo produce imágenes que son mapas probabilísticos de la normalidad de la textura y fue propuesto como una ayuda de visualización para diagnostico por médicos expertos. En este artículo la usabilidad de dicho modelo es estudiada para detección automática de masas. Una estrategia de segmentación es propuesta y evaluada en 79 casos de mamografía.

Introduction

Breast cancer is currently the second most common cause of death among women after lung cancer. Approximately one out of ten women can develop breast cancer during their lifetime according to the American Cancer Society (2012). Mammography is a powerful diagnostic tool for early detection, and Computer aided detection (CAD) systems for mammography have proven to be effective in facilitating early detection and diagnosis, a key factor improving patient prognosis and reducing mortality rates (Sampat, Markey & Bovik, 2004). There are two types of lesions associated to breast cancer: masses and micro-calcifications (BIRADS-Committee, 2003). Detection of masses is still an open issue in research and more challenging problem than the detection of micro-calcifications (Oliver et al., 2010). The work of Sampat et al.(2004) and Olivier et al. (2010) presents a comprehensive survey.

Mammography with CAD is the most widely researched problem of diagnostic radiology (Yoon et al., 2007). A large number of works addressed the problem of mass detection as a problem of texture analysis: Li et al. (1995), Mencattini et al. (2008), Haindl & Mikes (2008) to cite a few. An unsupervised scheme for mass segmentation has proven to be feasible and to have many advantages over supervised schemes due to great variability of normal tissue texture. Li et al. (1995) developed this approach based on a stochastic model for texture using Markov Random Fields (MRF).

The same idea of unsupervised texture analysis using local neighborhoods is found in Haindl et al. (2008). They proposed an automatic segmentation algorithm for fibrous glandular tissue of mammograms based on texture features from a MRF model and segmentation based on clustering in feature space with a multivariate gaussian mixture model. Grim et al. (2009) proposed the use of log-likelihood images computed from a multivariate gaussian mixture model of gray levels of local neighborhoods. They proposed log-likelihood images as a novel visualization tool for human interpretation and discuss the properties of such statistical constructs

for abnormality detection. The scope of their study did not include any automatic segmentation or classification stage.

As suggested by Grim et. al. (2009) the properties of log-likelihood images for mass detection are: (1) Masses should be mapped as regions that are darker than normal tissue on log-likelihood images. (2) Borders are enhanced by iso-countour-like lines that the model produces on the regions of transition between texture components. The properties are illustrated in a few cases and no detection strategy is proposed.

The scope of this work is to test these two hypotheses on a larger population sample. Qualitative assessment is performed on log-likelihood images and a visualization strategy is proposed. As the model components and samples from both normal and abnormal tissue on the log- likelihood feature space are projected via principal component analysis.
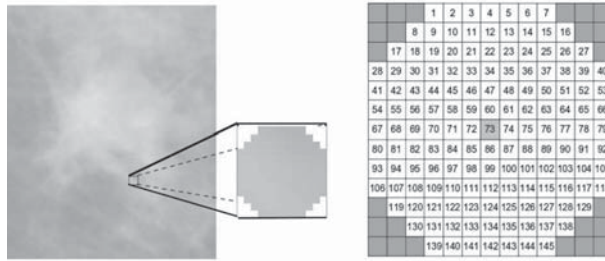

Figure 1. Sliding window for features extraction

MATERIALS AND METHODS

The images used for this study were taken from the Digital Database for Screening Mammography (DDSM), mass sets BCRP_MASS_0 and BCRP_MASS_1 (Heath et al., 2001). These sets contain 39 and 40 cases respectively. Each case is called a full field mammography case (4 images, left and right cranio-caudal CC and mediolateral oblique MLO). Each case contains a malignant mass visible in one or two images. In total there are: 79 cases, 316 images, out of which 160 have a visible mass. Ground truth mass contours are provided with the datasets. The multivariate gaussian mixture model for abnormality-enhanced visualization was proposed by Grim et al. (2009).

This model captures normal patterns in each mammographic case and enhances texture abnormalities based on gray level distribution of pixels within a local neighborhood.

Formally, let S be the set of all measurement vectors x formed by extracting the gray levels of N pixels of a local neighborhood as the center pixel iterates over each pixel of each image of case $S = \{x^{(1)}, x^{(2)},... \}, x^{(i)} \in R^N$, as illustrated in Figure 1.

Image textures can be modeled with a multivariate gaussian mixture:

$$P(x) = \sum_{m \in M} w_m F(x|\mu_m, \Sigma_m) \tag{1}$$

Where each w_m is a mixture proportion satisfying $w_m \geq 0$ and $\sum_{m \in M} w_m = 1$. Each component $F(x|\mu_m, \Sigma_m)$ is a multivariate gaussian with diagonal covariance matrix. The constraint on the covariance matrix is due to computational causes.

The model parameters $\theta = \{w, \mu, \Sigma\}$ are estimated by the EM algorithm (Grim et al., 2009). An efficient implementation may use lookup tables to save computation time and should prevent variable underflow (Gallego-Ortiz & Fernandez-Mc-Cann, 2011).

With the estimated model, log-likelihood images are computed over the original texture image at every pixel location as the logarithm of the probability in Equation (1), called likelihood when interpreted as a function of parameters θ.

$$logP(x|\theta) = log \left[\sum_{m \in M} w_m F(x|\mu_m, \Sigma_m) \right] \tag{2}$$

For model selection, the number of components $M = |M|$ and the size of the neighborhood N is assessed by visual inspection of the characteristics of log-likelihood images produced with Equation (2), computed after model fitting for various M and N combinations.

As N is typically on the order of tens to hundreds, the model and samples can be projected to a lower dimensional space for visualization and characterization purposes. This is achieved using principal component analysis (PCA) of the data points x $\in S$ and the mixture model as follows. Let $C \in R^{N \times N}$ be the covariance matrix of dataset S. The eigen-analysis of C produces

$$CU = \Lambda U \tag{3}$$

where Λ is the diagonal matrix with the eigenvalues in non-increasing order and U is the matrix formed by the corresponding eigenvectors in its columns.

By using the first $k \leq N$ eigenvectors U_{red}, we obtain the desired projections z = U_{red}^T x.

The projection error, is calculated by reconstructing the original data points x from the projections z to get x' = U_{red} z. The average projection error is then given by $e = \frac{1}{|S|} \sum_{x \in S} (x - x')^T (x - x')$. The total variability is given by $v_t = \frac{1}{|S|} \sum_{x \in S} x^T x$, so that the total variability conserved is $v = 1 - \frac{e}{v_t}$.

Similarly, the model can be projected by $\mu'_m = U_{red}^T \mu$ and the component covariance matrices by the quadratic form, $\Sigma'_m = U_{red}^T \Sigma'_m U_{red}$ for all components $m = 1,...,M$.

RESULTS

Model selection was performed by visual image evaluation for several configurations. A good compromise was obtained with images down-sampled at 200 µm/pix, with M=36 components and N=69 variables. Segmentation of breast tissue was undertaken with a simplified model of M=3 and N=69.

The model was estimated for each case, i.e. data of the four images. Then log-likelihood images are computed for each image with the estimated model parameters for the case. Figure 2 shows three images with masses, original and corresponding log-likelihood images.

Three categories were defined after visual inspection of masses in log-likelihood images, according to the degree they support the mass hypothesis: Positive, Weak, Negative. For example in Figure 2, case 0016 is a 'positive' because the center of the mass is darker than the surrounding regions and there are contour lines. Cases 1468 and 1134 are 'Weak' since even if there are contour lines, the mass center is not darker than the surrounding tissue. Under weak classification fall the images that have one of the properties but not both. Finally, negative cases are those that do not have any of the expected properties. Classification counts are summarized in Table 1. The weaker hypothesis among images was the darker seed over mass region.

To investigate the geometry of the feature space, PCA was used to project data and model as discussed. For this purpose a dataset of mass and normal tissue was selected from each case using ground truth data of the database.
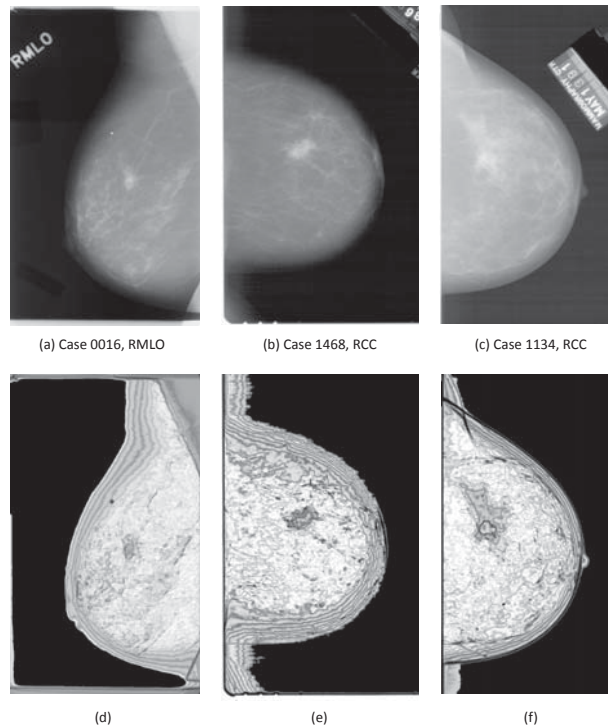

(a) Case 0016, RMLO   (b) Case 1468, RCC   (c) Case 1134, RCC

(d)   (e)   (f)

Figure 2. Original images and their log-likelihood images

Figure 3 summarizes the results of projecting the model along with the datasets to k=2 dimensions. The ratio of variability conserved in these projections are: 0.99879, 0.9985 and 0.9994 respectively.

Figure 3 depicts for each of the illustration cases: the projected estimated model as contour lines and surfaces respectively and both normal tissue feature vectors as small dots and small triangles for mass tissue feature vectors over the contour lines plot, while only mass vectors are plotted on the log-likelihood surfaces.

DISCUSSION

Log-likelihood images produced by the model were less discriminating than expected for the mass hypothesis. Only 57.5% of the mass image cases have a positive or weak degree of agreement. Weak images are mostly lacking dark cores for mass regions (first hypothesis). The performance of a detection algorithm searching for dark regions and some sort of concentric activity will present a similar performance.

Table 1. Classification counts

Positives	Weaks	Negatives	Totals	
absolute count	20	72	68	160
relative fraction	0.125	0.45	0.425	1.0

Visualization of some typical cases evidences the fact that the texture representation obtained by taking the gray levels of local neighborhoods in raster order resides in regions very close to the hyper diagonal of the feature space, $x_1 = x_2 = \ldots = x_N$. The distribution of examples in the bidimensional plots in Figure 3 are concentrated along the first principal component axis u1, the horizontal axis in these plots. The amount of variability conserved in this low dimensional projection is of 99 %, which suggests that the geometry of the high dimensional space is not very different.

Note also that, in general, the representation fails to discriminate normal texture from abnormal texture because mass and normal feature vectors are closely interlaced or superimposed in the feature space. From the perspective of the model, sometimes the mass feature vectors may fall in model valleys, as evidenced in Figures 3a and 3b for Case 0016, but, in general, they fall in dense regions not always with low probability densities as in Figures 3c to 3e for cases 1468 and 1134.

The above information suggests that the local representation alone is not enough to discriminate between mass and normal tissues. The absence of dark cores for masses can be explained by the fact that local texture representations of masses and dense normal tissue are very similar. To classify them more precisely, a more sophisticated feature space should be studied, somehow including the multi-scalar aspects of texture. For that purpose we believe that a combination of both local and context features may improve the performance of an initial detection system based on texture analysis (Hupse & Karssemeijer, 2009). Conceptually this could extend this approach and is suggested as future work.
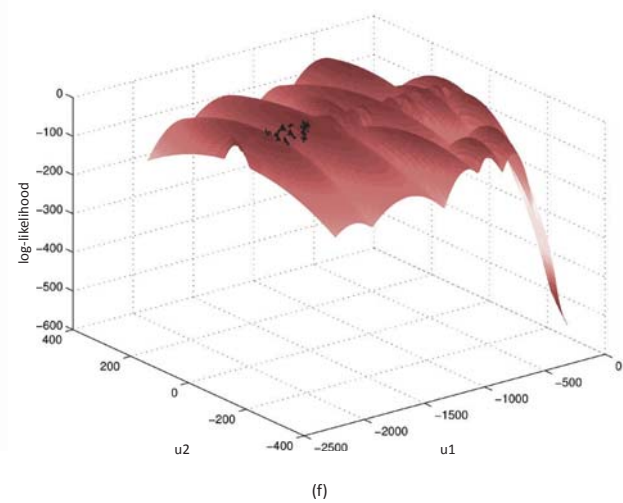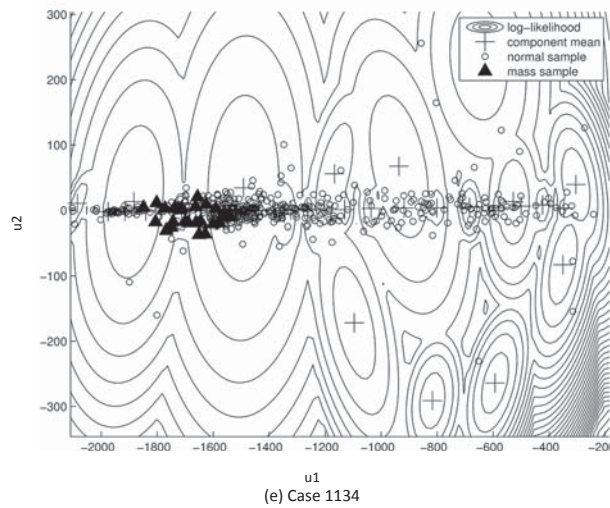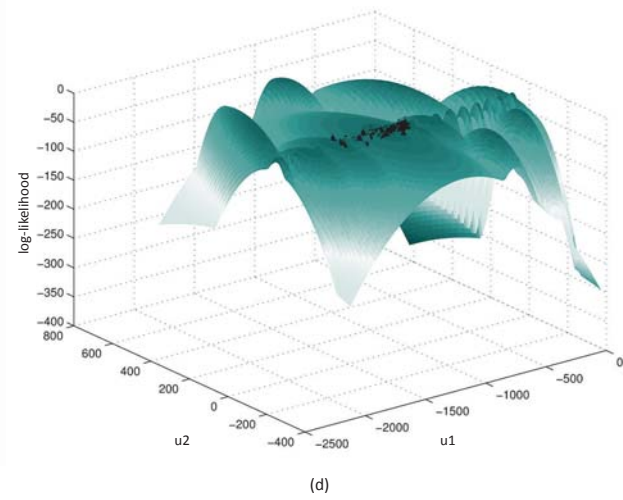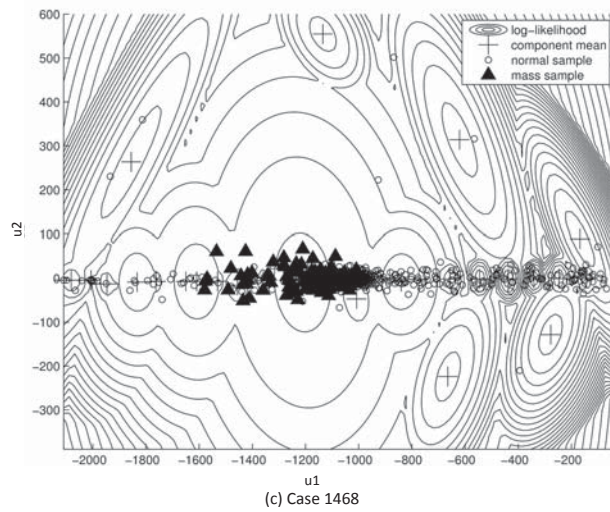
(a) Case 0016

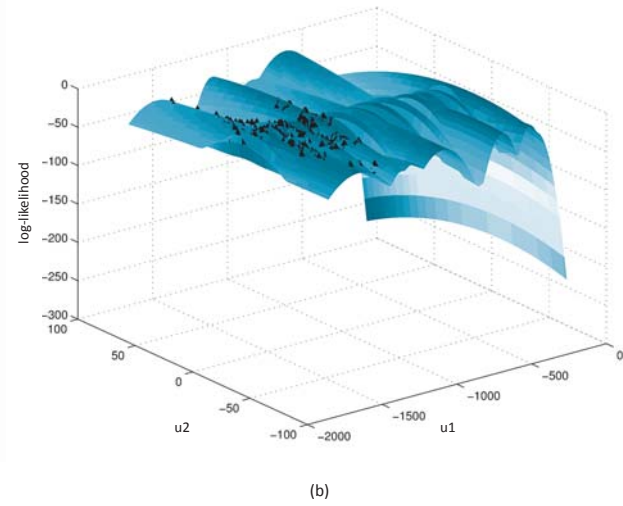

(b)
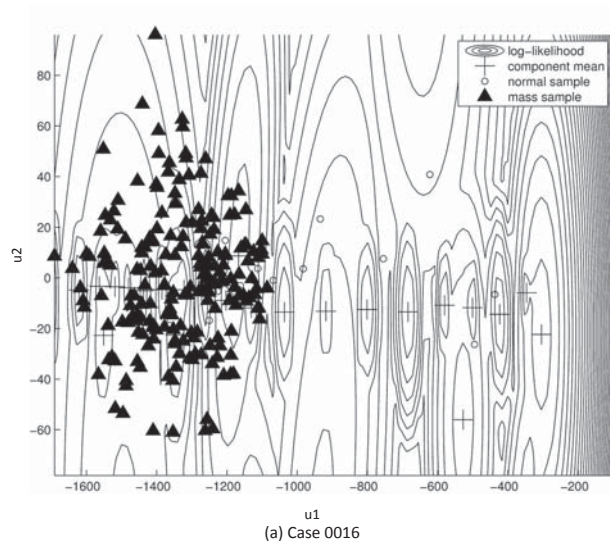

(c) Case 1468

(d)

(e) Case 1134

(f)

Figure 3. PCA projections of labeled feature vectors and estimated model

CONCLUSIONS

This study focused on the usability of a gaussian mixture model based on local neighborhood intensities for automatic mass detection. It revealed that the mixture model is flexible enough to capture the complexity of normal tissue and to adapt, in an unsupervised fashion, to the distribution of each mammography case. However, the selected texture representation is not discriminating enough since in most cases normal and mass examples overlap.

The proposed framework provides a means to visualize results from different texture representations by using projection techniques. These visualizations provide insight into the feature space geometry and allow us to monitor the fitness of the model to the underlying distribution. Further investigation is suggested in relation to these ideas.

Note: This project was financed by CODI (Comité para el Desarrollo de la Investigación de la Universidad de Antioquia).

REFERENCES

American Cancer Society. (2012). *Cancer facts & figures 2012*. Atlanta: American Cancer Society

BIRADS-Committee, A. (2003). Breast imaging and reporting data system-mammography (fourth ed.). Reston VA.

Gallego-Ortiz, N., & Fernandez-Mc-Cann, D. (2011). Efficient implementation of the em algorithm for mammographic image texture analysis with multivariate gaussian mixtures. In Proceedings of the ieee workshop on statistical signal processing (ssp2011)

Grim, J., Somol, P., Haindl, M., & Danes, J. (2009, april). Computer-aided evaluation of screening mammograms based on local texture models. *Image Processing, IEEE Transactions on*, *18* (4), 765–773.

Haindl, M., & Mikes, S. (2008, dec.). Unsupervised mammograms segmentation. In *19th International Conference on Pattern Recognition.* Conferencia llevada a cabo en ICPR, Tampa, Florida, USA.

Heath, M., Bowyer, K., Kopans, D., Moore, R., & Kegelmeyer, W. P. (2001, june). The digital database for screening mammography. In Yale M. (Ed.), *Proceedings of the fifth international workshop on digital mammography* Toronto, Canada.

Heath, M., Bowyer, K., Kopans, D., Moore, R., Kegelmeyer, W.P. The Digital Database for Screening Mammography. In *Proceedings of the Fifth International Workshop on Digital Mammography*, 2000; Medical Physics Publishing: Madison, WI, USA, pp. 212–218.

Hupse, R., & Karssemeijer, N. (2009). Use of normal tissue context in computer-aided detection of masses in mammograms. *Medical Imaging, IEEE Transactions on*, *28* (12), 2033 -2041.

Li, H. D., Kallergi, M., Clarke, L. P., Jain, V. K., & Clark, R. A. (1995). Markov random field for tumor detection in digital mammography. *IEEE Transactions on Medical Imaging*, *14* (3), 565–576.

Mencattini, A., Salmeri, M., Lojacono, R., Frigerio, M., & Caselli, F. (2008). Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing. *IEEE Transactions on Instrumentation and Measurement*, *57* (7), 1422–1430.

Oliver, A., Freixenet, J., Martí, J., Pérez, E., Pont, J., Denton, E. R., & Zwiggelaar, R. (2010). A review of automatic mass detection and segmentation in mammographic images. *Medical Image Analysis*, *14* (2), 87–110.

Sampat, M. P., Markey, M. K., & Bovik, A. C. (2004). *Handbook of image and video processing*. In (pp. 1159–1217). Elsevier Academic Press.

Yoon, H. J., Zheng, B., Sahiner, B., & Chakraborty, D. P. (2007). Evaluating computer-aided detection algorithms. *Med Phys*, *34*(6), 2024–38. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17654906