Texture Descriptors applied to Digital Mammography

C. Mata, J. Freixenet, X. Lladó and A. Oliver

June 5, 2008

Abstract

Breast cancer is the second cause of death among women cancers. Computer Aided Detection has been demonstrated an useful tool for early diagnosis, a crucial aspect for a high survival rate. In this context, several research works have incorporated texture features in mammographic image segmentation and description such as Gray-Level co-occurrence matrices, Local Binary Patterns, and many others. This paper presents an approach for breast density classification based on segmentation and texture feature extraction techniques in order to classify digital mammograms according to their internal tissue. The aim of this work is to compare different texture descriptors on the same framework (same algorithms for segmentation and classification, as well as same images). Extensive results prove the feasibility of the proposed approach.

Introduction

Breast cancer is the leading cause of death among all cancers for middle-aged women in most developed countries [1]. Any diagnostic tool that could help to improve the sensitivity or specificity of breast cancer would be highly valued. The usefulness of mammography in the symptomatic patient is undisputed; mammography is primarily used to demonstrate the presence of breast cancer and, specifically to indicate the size and location of a tumor.

This paper reviews the most recent published techniques that use texture descriptors in computerized digital mammographic analysis. In this sense, we want to show advantages and disadvantages of different texture descriptors when applied to mammographic images. Moreover, we have selected and implemented three key methods: two of the most used texture descriptor such as Gray-Level Co-occurrence Matrices (GLCM), and Laws' measures masks, and a relatively new texture feature in the field like the Local Binary Patterns (LBP). In order to compare texture features, we follow an approach based on grouping pixels according to their appearance (fatty or dense) using the Fuzzy C-Means as the segmentation method.

The databases used in our paper are the widely known MIAS database and the digital Trueta database. Both databases have been manually classified by a set of experts according to their internal tissue in the Breast Imaging Reporting and Data System (BIRADS) [2] categories. According to this protocol, the mammograms can be classified as:

- BIRADS I: the breast is almost entirely fatty.
- BIRADS II: there is some fibroglandular tissue.
- BIRADS III: the breast is heterogeneously dense.
- BIRADS IV: the breast is extremely dense.

First row of Figure 1 shows an example mammogram of each class in increasing density order.

The rest of this paper is organized as follows. The following Section reviews the state of the art focused on texture descriptors in digital mammography. Section 2 presents the breast density segmentation approach using texture. Afterwards, the evaluation of experimental results is presented in Section 3. Finally, the paper ends with conclusions and further work.

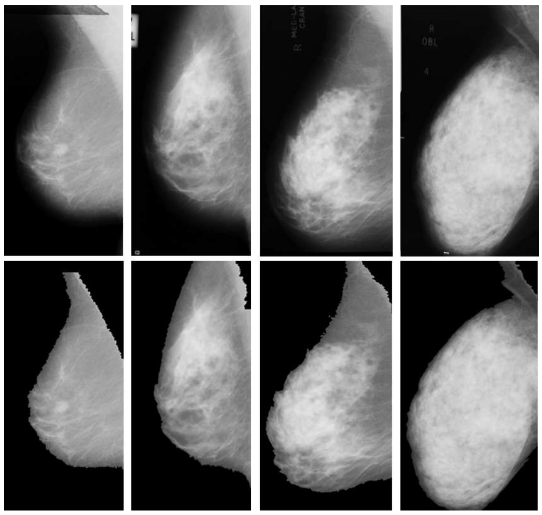

Figure 1: The first row shows a mammogram of each BIRADS class in increasing density order (increasing BIRADS), while the second row shows the result of the breast profile segmentation.

1 State of the Art

We have reviewed some of the most recent publications focused on breast density segmentation and classification, as well as sign (or abnormality) detection using texture descriptors. In this sense, microcalcifications and masses are well known signs in the field. Studies of breast cancer were aimed to improve radiologist's diagnostic performance by indicating suspicious areas. The increment of research papers, contributions and a variety of computer based methods in mammography was fundamental.

Bovis and Singh (2000) [3] studied how to detect masses in mammograms on the basis of textural features using five co-occurrence matrices statistics extracted from four spatial orientations: horizontal, left diagonal, vertical and right diagonal. A classification is performed using each feature vector and linear discriminant analysis. According to Martí et al. [4], GLCMs are frequently used in computer vision obtaining a satisfactory results as texture classifiers in different applications. Their approach uses mutual information with the purpose to calculate the amount of mutual information between images. Blot and Zwiggelaar [5, 6] proposed two approaches based in detection and enhancement of structures in images using GLCM.

In 2003, different approaches based again on GLCM as the feature descriptors were developed. Youssry et al. [7] presented a neuro-fuzzy model for fast detection of candidate circumscribed masses in mammograms using GLCM as the texture features. On the other hand, Martí et al. [8] proposed a supervised method for the segmentation of masses in mammographic images using the texture features which present an homogeneous behavior inside the selected region.

In recent studies, Lyra et al.[9] study how to identify breast tissue quality data quantification using a CAD system. In their approach, images were categorized using the BIRADS breast density index, and the texture features were derived for each sub-region from an averaged GLCM.

Law's masks is another texture descriptor commonly used in mammography. For instance, Bovis and Singh [3] used this measure to classify mammographic images according to the internal density. On the other hand, Varela et al. [10] developed an image processing algorithm for classifying the breast lesions based on quantitative measures as shape, contrast, and spiculation. Features based in Laws' texture energy were also extracted from the straighten border region.

As a summary, Gray-Level Co-occurrence matrices is the most used technique for texture description in the majority of works, showing the importance of GLCM compared to other techniques. Laws' texture is the second more used feature extraction technique. Moreover, different authors as Karahaliou et al. [11] and Oliver et al. [12] combined both texture features in their approaches. On the other hand, Local Binary Patterns is a modern and promising technique for texture description, although there are still few works applying it on digital mammography. Lladó et al. [13][14] and Oliver et al. [12][15] are recent examples.

2 Breast Density Segmentation Using Texture

Figure 2 shows a graphical scheme of the proposed framework, structured in four stages: pre-processing, feature extraction, segmentation and classification.

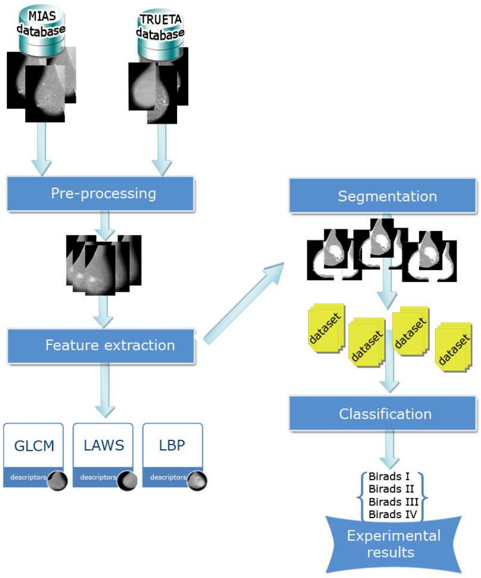

Figure 2: Developed framework for the comparison of texture descriptors.

2.1 Step 1: Pre-processing

The first step of the framework is the breast profile segmentation. The aim of this segmentation is to separate the breast from other objects in the mammogram with a minimum loss of breast tissue. In general two independent steps are performed [16]. The first one aims to segment the background and annotations from the whole breast area, while the second one involves separating the pectoral muscle (when present) from the rest of the breast area.

In this work we used a previous developed approach [16]. Firstly, an automatic thresholding algorithm is used to separate the area composed of the breast and pectoral muscle from the background of the image. Afterwards, a region growing algorithm allows to automatically locate muscle and extract it from the breast. Figure 1 shows two different examples of the breast profile segmentation. Note that this segmentation results in a minor loss of skin-line pixels, but those pixels are deemed not to be relevant for tissue estimation.

2.2 Step 2: Feature extraction

Once the pre-processing is applied, an extraction process is used in order to obtain the texture features. As already said, in this work we compare three different feature descriptors: GLCM, Laws' masks and LBP. The result of this extraction process is a feature vector per pixel describing the texture of its surrounding neighborhood. Note that, as this process is done per each pixel in the image (in fact, per each pixel in the breast), the result of this step is a collection of texture images.

Let us briefly explain each descriptor. Firstly, the idea of GLCM is to calculate the co-occurrence matrix for small regions of the image and then use this matrix to find statistic values. In our approach, we have used distances: 3, 5 and 10, angles: 0, 45, 90 and 135, and the following statistics: Contrast, Correlation, Uniformity, Homogeneity, Probability, Inverse and Entropy [17]. In total, 196 descriptors were derived from GLCM. Secondly, Laws' texture consists on obtaining statistic measures from the convolution result using masks filter [18]. Thus, we can obtain three statistics such as mean, absolute mean and standard deviation. In total, 117 descriptors were derived from Laws' texture. Finally, the local binary patterns (LBP) operator is defined as a gray-scale invariant texture measure, derived from a general definition of texture in a local neighborhood. For each pixel in an image, a binary code is produced by thresholding its neighborhood (for instance, the closest 8 pixels) with the value of the center pixel. A histogram is then constructed to collect up the occurrences of different binary codes representing different types of curved edges, spots, flat areas, etc.[19]. This histogram is dealt as the feature vector. In this work we test the use of four mapping: LBP Basic, Uniform LBP, Rotation-invariation LBP and Uniform rotation-invariation LBP. In total, 36 descriptors were obtained.

2.3 Step 3: Segmentation

The segmentation is a crucial step in order to obtain a quantitative measure of the density of the breast. We use the well known Fuzzy C-Means algorithm [20], because it has a good trade off between performance and computational time.

In our implementation based on [4], the criterion function minimized by the algorithm is defined by:

$$e^2(\Xi, U) = \sum_{n=1}^{N} \sum_{t=1}^{T} u_{nt}^2 \parallel p_n - c_t \parallel^2 \tag{1}$$

where Ξ is the partition of the image, U is the membership matrix: u_{nt} represents the membership of pattern p_n to belong to cluster t, which is centered at:

$$c_t = \frac{\sum_{n=1}^{N} u_n t p_n}{\sum_{n=1}^{N} u_{nt}} \tag{2}$$

where N is the number of patterns in the whole image (i.e. the number of pixels), and T the number of clusters, which has to be known a priori. With the aim of segmenting tissue into fatty and dense classes we used in this work only two clusters. Moreover, the gray-level mean of each cluster is used to establish a ranking of cluster densities: a higher mean corresponds to a higher probability of being a cluster of dense tissue.

2.4 Step 4: Classification

The classification of mammograms according to the BI-RADS categories was performed in four different ways: k-Nearest Neighbors(KNN) classifier, Fisher discriminant classifier (Fisher), Linear discriminant analysis (LDC) and Support Vector Machine algorithm (SVM). Moreover, a feature selection method is used in order to reduce the descriptor set size. Concretely, we used single forward selection (SFS) because it is an effective approach at a reasonable computational cost [21].

3 Experimental Results

Two mammographic databases are used to evaluate the approach and compare the performance of each texture descriptor: the MIAS database and the Trueta database. The MIAS database is composed by a set of the 322 mammograms of left and right mammograms from 161 women. The spatial resolution of the images is $50\mu m \times 50\mu m$ and quantized to 8 bits with a linear optical density in the range $0 - 3.2$. On the other hand, we used 159 mammograms from Trueta database. These mammograms are acquired using a full-field digital mammograph (Siemens Mammomat Novation), and stored in DICOM format in a PACS server. The mammograms are 70 micron pixel edge.

The results in this paper were obtained using a 10-fold cross-validation methodology. Each dataset was divided in 10 different groups, nine of them were merged for training the classifiers, while the remaining group was used to test them. This procedure was repeated until all groups were used for testing. Moreover, all these methodology is repeated five times in order to obtain significant results (for obtaining standard deviations). The tables show the percentage of correct classification, i.e. the total number of correct classified mammograms divided by the total number of mammograms.

3.1 MIAS Database

Table 1 shows the results obtained using individually the three feature extraction techniques for each classifier. We can state that using the SVM classifier we obtained better results than using the others ones, with a mean value of 67% using GLCM, 62% using Laws and 68% using LBP.

On the other hand, once feature selection is applied for each dataset, we have obtained better performances. Again, SVM is the best one obtaining a mean value of 75%, 72% and 79% respectively. Hence, the best combination is obtained when using the LBP as the feature extraction technique and the SVM classifier.

	GLCM		Feature selection	
	Mean	**Deviation**	**Mean**	**Deviation**
KNN	0.580	0.019	0.679	0.019
Fisher	0.634	0.026	0.671	0.014
LDC	0.658	0.023	0.714	0.018
SVM	0.669	0.008	0.751	0.009

	LAWS		Feature selection	
	Mean	**Deviation**	**Mean**	**Deviation**
KNN	0.550	0.019	0.673	0.019
Fisher	0.456	0.022	0.594	0.015
LDC	0.583	0.018	0.694	0.017
SVM	0.628	0.009	0.721	0.008

	LBP		Feature selection	
	Mean	**Deviation**	**Mean**	**Deviation**
KNN	0.626	0.010	0.717	0.011
Fisher	0.667	0.011	0.708	0.014
LDC	0.653	0.010	0.749	0.010
SVM	0.684	0.008	0.790	0.005

Table 1: MIAS: Percentage of correct classification with and without feature selection.

Table 2 shows one example of confusion matrix obtained when using LBP and SVM. Confusion matrices should be read as follows: rows indicate the object to recognize (the true class) and columns indicate the label the classifiers associates at this object. Note that mammograms belonging to BIRADS I are more correctly classified compared to other classes.

	B-I	B-II	B-III	B-IV
B-I	80	4	2	1
B-II	10	76	11	2
B-III	7	7	72	3
B-IV	3	6	7	28

Table 2: MIAS: Confusion matrix of the LBP description using the SVM classifier

3.2 Trueta Database

Table 3 shows again the results obtained using individually the three feature extraction techniques for each classifier, but applied to the Trueta database. Again, the SVM classifier obtained better results: 56% using GLCM, 58% using Laws and 61% using LBP. On the other hand, using the feature selection process in each dataset, the results increase to 70%, 67% and 74% respectively. Again, LBP provides the best results.

	GLCM		Feature selection	
	Mean	**Deviation**	**Mean**	**Deviation**
KNN	0.525	0.012	0.563	0.012
FISHER	0.457	0.034	0.602	0.026
LDC	0.505	0.010	0.605	0.022
SVM	0.562	0.026	0.704	0.005

	LAWS		Feature selection	
	Mean	**Deviation**	**Mean**	**Deviation**
KNN	0.494	0.021	0.533	0.029
Fisher	0.508	0.033	0.601	0.011
LDC	0.537	0.023	0.562	0.014
SVM	0.581	0.027	0.674	0.008

	LBP		Feature selection	
	Mean	**Deviation**	**Mean**	**Deviation**
KNN	0.522	0.019	0.581	0.011
Fisher	0.600	0.016	0.644	0.013
LDC	0.588	0.023	0.624	0.014
SVM	0.616	0.027	0.739	0.006

Table 3: Trueta: Percentage of correct classification with and without feature selection.

Table 4 shows one example of confusion matrix obtained when using LBP and SVM. Note that in this case, the mammograms belonging to BIRADS II are better classified. The difference between both datasets can be due to the fact the MIAS is a digitized database while the Trueta is a fully digital one.

	B-I	B-II	B-III	B-IV
B-I	50	14	6	10
B-II	1	30	0	1
B-III	3	2	36	2
B-IV	1	0	1	2

Table 4: Trueta: Confusion matrix of the LBP description using the SVM classifier

3.3 Discussion

Our initial experiments consisted on the evaluation of the proposed method using the texture features independently. According to our experiments, the best results were obtained using Local Binary Patterns and feature selection. Moreover, Support Vector Machines (SVM) provides the best results. The first three columns of Table 5 summarizes the best obtained results.

In order to take advantage of the different texture descriptors, we evaluate in the same framework two feature combinations: firstly, Laws and LBP, and secondly GLCM, Laws and LBP. The last columns of Table 5 shows the obtained results. Note that using both combinations, the results increases. However, the best increase is obtained when combining all the descriptors. This shows that GLCM, Laws and LBP provide different texture description, and it is necessary to combine them for obtaining a best description of the images. Comparing again the performance between each database, we still obtained better results using the MIAS database than using the Trueta one, with a difference of 5%.

Table 6 and Table 7 shows the confusion matrices obtained when using all the features, the SFS for extracting the best ones, and SVM for classifying them, and using, respectively, the MIAS and the Trueta databases. Comparing the results with the obtained when testing individually the features, note that the class BIRADS I has decrease its performance. However, mammograms belonging to BIRADS II and III are more correctly classified. On the other hand, for the Trueta database, the behavior is the opposite: there is a decrease in the other classes in order to obtain a significant increase in the BIRADS I class. Again, this shows the difficulty to deal with digitized and digital mammograms.

	B-I	**B-II**	**B-III**	**B-IV**
B-I	75	1	10	1
B-II	10	83	7	3
B-III	4	2	84	5
B-IV	1	4	4	28

Table 6: MIAS: Confusion matrix of the second combination (GLCM, LAWS and LBP) using the SVM classifier

	B-I	**B-II**	**B-III**	**B-IV**
B-I	70	2	4	4
B-II	2	20	5	5
B-III	1	2	32	8
B-IV	0	0	1	3

Table 7: Trueta: Confusion matrix of the second combination (GLCM, LAWS and LBP) using the SVM classifier

Finally, Figure 3 shows a visual example of segmentation results for the four types of BIRADS: (a) BIRADS I, (b) BIRADS II, (c) BIRADS III and (d) BIRADS IV respectively. These images corresponds to the MIAS database by using the LBP texture descriptors. Three regions are described in each image (black, white and gray), where the white and gray region are considered as the two clusters. The black region is not considered as a segmentation part because correspond to the background and therefore it is not included in the clustering. Moreover, ordering by the gray-level mean of each cluster the dense tissue is denoted to brightness.

4 Conclusions and further work

A comparison of three different feature extraction methods for segmenting the breast density has been presented

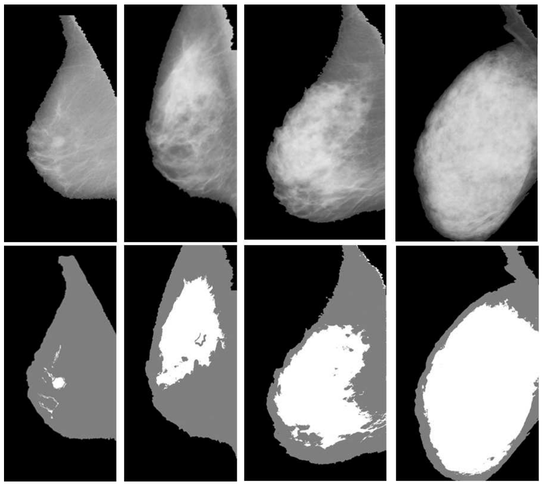

Figure 3: Segmented mammograms using LBP features. Note that the dense cluster increase accordingly to the density of the breast.

in this paper: Gray-level Co-occurrence matrix (GLCM), Laws'texture measures and Local Binary Patterns. Thus, we developed an approach based on grouping the pixels according to their appearance (fatty or dense) using a Fuzzy C-Means for breast density segmentation. The performance of the approach is analyzed using classifiers in order to classify breast tissue according to BIRADS categories. The classifiers used were: k-Nearest Neighbors classifier, Fisher discriminant classifier, Linear discriminant classifier and Support Vector Machine algorithm. Once the mammogram is classified according to its density class, results obtained using MIAS and the Trueta databases, demonstrate the feasibility of our proposal.

As further work, several research work could be considered. For instance, instead of using fuzzy C-means there are some other segmentation methods proposed in the related literature, and it will be a good idea to test them. In the same line, we have used four classifiers, but we have not combined them. As we have combined different features descriptors sets, we can also try to combine classifiers in order to improve our results.

References

[1] Eurostat, "Health statistics atlas on mortality in the european union." *Office for Official Publications of the European Union*, 2002.

[2] A. C. of Radiology, "Illustrated breast imaging reporting and data system birads," *American College of Radiology, 3rd edition*, 1998.

[3] K. Bovis and S. Singh, "Detection of masses in mammograms using texture features," *15th International Conference on Pattern Recognition (ICPR'00)*, vol. 2, p. 2267, 2000.

[4] R. Martí, R. Zwiggelaar, and C. Rubin, "A novel similarity measure to evaluate image correspondence," *15th Interna-*

	GLCM	LAWS	LBP	1st Combination	2nd Combination
MIAS	75%	72%	79%	80%	83%
TRUETA	70%	67%	74%	76%	78%

Table 5: Description of the percentages of correct classified instances using SVM with GLCM, Laws, LBP and two combinations: Laws+LBP and GLCM+Laws+LBP

tional Conference on Pattern Recognition (ICPR'00), vol. 3, p. 3171, 2000.

[5] L. Blot and R. Zwiggelaar, "Extracting background texture in mammographic images: a co-occurrence matrices based approach," *Proceedings of the 5th International Workshop on Digital Mammography, Toronto(Canada)*, pp. 142–148, 2000.

[6] L. Blot, R. Zwiggelaar, and C. Boggis, "Enhancement of abnormal structures in mammographic images," *Proceedings of Medical Image Understanding and Analysis*, pp. 125–128, 2000.

[7] N. Youssry, F. Abou-Chadi, and A. El-Sayad, "Early detection of masses in digitized mammograms using texture features and neuro-fuzzy model," *4th Annual IEEE Conf on Information Technology Applications in Biomedicine*, 2003.

[8] J. Martí, J. Freixenet, X. Muñoz, and A. Oliver, "Active region segmentation of mammographic masses based on texture, contour and shape features," *Springer-Verlag Berlin Heidelberg*, vol. LNCS 2652, pp. 478–485, 2003.

[9] M. Lyra, S. Lyra, B. Kostakis, S. Drosos, and C. Georgosopoulos, "Digital mammography texture analysis by computer assisted image processing," *IEEE International Workshop on Imaging Chania Greece September 2*, pp. 223–227, 2008.

[10] C. Varela, N. Karssemeijer, and P. Tahoces, "Classification of breast tumors on digital mammograms using laws texture features," *Springer-Verlag Berlin Heidelberg*, vol. LNCS 2208, pp. 1391–1392, 2001.

[11] A. Karahaliou, I. Boniatis, P. Sakellaropoulos, S. Skiadopoulos, G. Panayiotakis, and L. Costaridou, "Can texture of tissue surrounding microcalcifications in mammography be used for breast cancer diagnosis," *Nuclear Instruments and Methods in Physics Research A 580*, pp. 1071–1074, 2007.

[12] A. Oliver, X. Lladó, R. Martí, J. Freixenet, and R. Zwiggelaar, "Classifying mammograms using texture information," *Medical Image Understanding and Analysis UK*, pp. 223–227, Aberystwyth Wales UK. July 2007.

[13] X. Lladó, A. Oliver, J. Martí, and J. Freixenet, "Dealing with false positive reduction in mammographic mass detection," *In Medical Image Understanding and Analysis*, pp. 81–85, 2007.

[14] X. Lladó, A. Oliver, J. Freixenet, R. Martí, and J. Martí, "A textural approach for mass false positive reduction in mammography," *To appear in Computerized Medical Imaging and Graphics. ELSEVIER*, 2009.

[15] A. Oliver, X. Lladó, R. Martí, J. Freixenet, and J. Martí, "False positive reduction in mammographic mass detection using local binary patterns," *Int ConfMed ImageComputComput Assist Interv*, vol. 4478, pp. 286–293, 2007.

[16] D. Raba, A. Oliver, J. Martí, M. Peracaula, and J. Espunya, "Breast segmentation with pectoral muscle suppression on digital mammograms," *In Lecture Notes in Computer Science*, vol. 3523, pp. 471–478, 2005.

[17] R. Haralick and K. Shanmugam, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics SMC-3 (6)*, vol. 6, pp. 610–621, 1973.

[18] K. Laws, "Rapid texture identification," *Image Processing for Missile Guidance*, vol. In SPIE Vol. 238, pp. 376–380, 1980.

[19] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." *IEEE Trans. Pattern Anal. Machine Intell. 24(7)*, pp. 971–987, 2002.

[20] J. Bezdek, "Pattern recognition with fuzzy objective function algorithms," *Plenum Press*, 1981.

[21] M. Dash and H. Liu, "Feature selection for classification," *Intelligent Data Analysis*, pp. 131–156, 1997.