Capturing Microcalcification Patterns in Dense Parenchyma with Wavelet-Based Eigenimages

Nikolaos Arikidis, Spyros Skiadopoulos, Filippos Sakellaropoulos, George Panayiotakis, and Lena Costaridou

Department of Medical Physics, School of Medicine,
University of Patras, 265 00 Patras, Greece
costarid@upatras.gr, panayiot@upatras.gr

Abstract. A method is proposed based on the combination of wavelet analysis and principal component analysis (PCA). Microcalcification (MC) candidate regions are initially labeled using area and contrast criteria. Mallat's redundant dyadic wavelet transform is used to analyze the frequency content of image patterns at horizontal and vertical directions. PCA is used to efficiently encode MC patterns in wavelet-decomposed images. Feature weights are computed from the projection of each candidate MC pattern at the wavelet-based principal components. To assess the effectiveness of the proposed method, the same analysis is carried out in original images. Candidate MC patterns are classified by means of Linear Discriminant Analysis (LDA). Free-response Receiver Operating Characteristic (FROC) curves are produced for identifying MC clusters. The highest performance is obtained when PCA is applied in wavelet decomposed images achieving 80% sensitivity at 0.5 false positives per image in a dataset with 50 subtle MC clusters in dense parenchyma.

1 Background

Mammography is currently the technique with the highest sensitivity available for early detection of breast cancer on asymptomatic women [1]. Detection of early signs of disease, such as microcalcifications (MCs), with screening mammography, is a particularly demanding task for radiologists. This is mainly attributed to the low MC contrast resolution, resulting from their small size [2]. These limitations have provided the basis for the development of Computer-Aided Detection (CAD) systems with high performance characteristics [3], [4], representing one of the most successful paradigms in medical image analysis. However, performance of such systems in case of dense tissue is challenged by the high correlation between fibroglandular tissue patches and MCs, resulting in increased false positive (FP) rate [5], [6].

One approach in CAD systems for MC detection is the use of the wavelet transform framework to analyze MCs based on their high frequency content. However, a large component of the power in a mammogram, at high spatial frequencies, is also noise, mainly originating from the inhomogeneous background of dense tissue structures, resulting in poor MC signal-to-noise ratio (SNR) [7]. Netch *et al*. [8], based on the circularly symmetric Gaussian model achieved 84% sensitivity with 1 average FP per image, using a Laplacian kernel to detect MCs as local maxima

Susan M. Astley et al. (Eds.): IWDM 2006, LNCS 4046, pp. 541 – 548, 2006.

at different frequency bands. Strickland *et al*. [9] have shown that the average 2D gray level profile of MCs is well described by a circularly symmetric Gaussian function. Since the optimum detector of Gaussian functions is the Laplacian of Gaussian, a wavelet filter close to the Laplacian of Gaussian was used to detect significant peak responses for objects of similar shape and size as the Gaussian filter. Soft or hard thresholding was used to set the low amplitude wavelet coefficients to zero, obtaining 70% sensitivity with 1 FP per image in a varying subtlety of MC clusters. Yoshida *et al*. [10] used an undecimated wavelet transform for MC detection achieving 78% sensitivity with 0.5 average FPs per image in a dataset with subtle MCs. Drexl *et al.* [11] used the continuous wavelet transform and features based on the evolution of the wavelet coefficients across scales. At 0.5 FP per image the sensitivity achieved was approximately 85%. Qian *et al*. [12] used a tree structured wavelet transform for multiresolution decomposition and a non-linear filter for suppressing image noise, achieving 94% sensitivity with 1.6 average FP per image. The aforementioned methods have been tested in image datasets of various types of breast parenchyma including dense tissue, with the exception of Lado *et al*. [13], who has worked on dense parenchyma, yielding 2.2 FPs per image with 73% sensitivity.

The aim of this study is to efficiently encode MC patterns analyzed by combining Principal Component Analysis (PCA) and wavelet decomposition. The capability of a feature vector based on this analysis is demonstrated in a detection task of subtle MC clusters embedded in dense parenchyma. To assess the effectiveness of the proposed MC cluster encoding method, the same analysis is carried out in original images.

2 Method

2.1 Labeling of Candidate MC Regions

MCs are very small structures visible as bright spots in the mammogram because their mass attenuation coefficient is higher than any other structure in the breast. However, due to the growth of MCs, there is no absolute lower bound to their contrast. Very small MCs may have low contrast relative to their background, which is sometimes close to structure noise originating mainly from fibroglandular tissue patches.

MC candidate regions are initially labeled using contrast and area criteria. In this study, pixel contrast is measured on a local basis, exploiting wavelet analysis [14]. Specifically, pixel contrast is defined as the difference between the pixel foreground and background maps normalized by their sum. The foreground pixel map corresponds to gray level values of the original image. The background pixel map corresponds to the gray level values of a low-pass filtered image. A contrast threshold of 0.5% is selected to preserve subtle MCs in dense parenchyma and an area threshold of 1.2 mm^2 to eliminate image components, which are likely to be macrocalcifications or line structures. Breast border identification is obtained with an edge detection technique based on the magnitudes of the derivative of a Gaussian operator.

2.2 Extraction of Feature Vector

A common approach of computerized MC detection methods is based on a two-stage process utilizing image feature extraction and subsequent classification to reduce FPs.