Using Wavelet-Based Features to Identify Masses in Dense Breast Parenchyma

Filippos Sakellaropoulos, Spyros Skiadopoulos, Anna Karahaliou, Lena Costaridou, and George Panayiotakis

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

Abstract. Automated detection of masses on mammograms is challenged by the presence of dense breast parenchyma. The aim of this study is to investigate the feasibility of wavelet-based feature analysis in identifying spiculated and circumscribed masses in dense breast parenchyma. The method includes an edge detection step for breast border identification and employs Gaussian mixture modeling for dense parenchyma labeling. Subsequently, wavelet decomposition is performed and intensity as well as orientation features are extracted from approximation and detail subimages, respectively. Logistic regression analysis (LRA) is employed to differentiate spiculated and circumscribed masses from normal dense parenchyma. The proposed method is tested in 90 dense mammograms containing spiculated masses (30), circumscribed masses (30) and normal parenchyma (30). Free-response receiver operating characteristic (FROC) analysis is used to evaluate the performance of the method, achieving 83.3% sensitivity at 1.5 and 1.8 false positives per image for identifying spiculated and circumscribed masses, respectively.

1 Background

Computer-Aided Detection (CAD) is one of the promising approaches for improving mass detection sensitivity in mammography [1]. Various image features in combination with classification methods have been proposed for automated mass detection [2]. Kegelmeyer *et al.* [3] have introduced edge orientation features based on local edge orientation histogram analysis as well as Laws' texture energy measures to identify spiculated mass containing areas. Karssemeijer *et al.* [4], [5] detected spiculated masses employing orientation features based on three directional second-order Gaussian derivatives. Wei *et al.* [6], [7] proposed multiresolution texture analysis extracted from spatial Gray Level Dependence Matrices (GLDM) for differentiation of masses from normal tissue. Liu *et al.* [8] extended mass edge orientation analysis with a multiresolution scheme for the detection of spiculated masses. Petrick *et al* [9] and Kobatake *et al.* [10] have utilized a combination of boundary (morphological) and texture features (GLDM analysis) to identify and segment the extent of masses, respectively. Zwiggelaar *et al.* [11] introduced area patterns using principal component and factor analyses for differentiating areas

Susan M. Astley et al. (Eds.): IWDM 2006, LNCS 4046, pp. 557 – 564, 2006.

containing masses from normal tissue. Chang *et al.* [12] and Baydush *et al.* [13] applied knowledge-based approaches for discriminating masses from normal tissue.

The performance of the aforementioned mass detection methods is characterized by high sensitivity (84-96%) and is challenged by the high number of false positive detections per image (1.0-4.4), especially in case of dense parenchyma (3.7-8.4) [14], [15], [16].

The aim of this study is to investigate the feasibility of wavelet-based features in identifying spiculated and circumscribed masses in dense breast parenchyma. A set of intensity and gradient-orientation multiresolution features are investigated, in combination with Logistic Regression Analysis (LRA) as classification scheme for differentiating masses from dense breast parenchyma.

2 Method

The steps of the proposed mass identification method are provided in Fig. 1. For each mammogram, the breast border is identified using an edge detection technique based on magnitudes calculated from the derivative of a Gaussian operator. Gaussian mixture modeling is then employed for segmenting the three breast components (uncompressed fat, fat and dense parenchyma) and labeling the dense parenchyma [17]. Following, wavelet decomposition is performed and multiresolution features are extracted for each pixel of dense parenchyma. These features are used as inputs in a trained logistic regression classifier (Fig. 2). Probability images are generated and

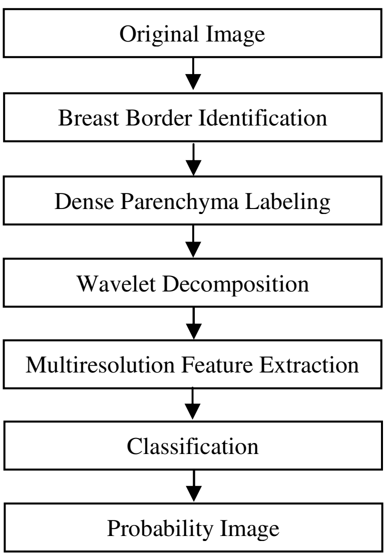

Fig. 1. Flow chart of the proposed method for breast mass identification

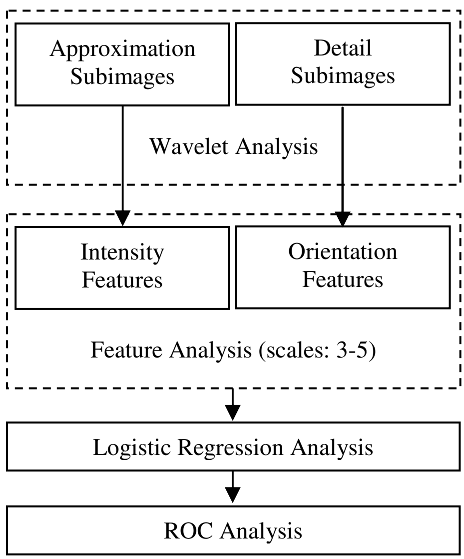

Fig. 2. Steps for classifier training in differentiating breast masses from normal dense parenchyma