Understanding Hessian-Based Density Scoring

Jakob Raundahl, Marco Loog, and Mads Nielsen

IT University of Copenhagen

Abstract. Numerous studies have investigated the relation between mammographic density and breast cancer risk. These studies indicate that women with high breast density have a four to six fold risk increase. An investigation of whether or not this relation is causal is important for, e.g., hormone replacement therapy (HRT), which has been shown to actually increase the density.

No gold standard for automatic assessment of mammographic density exists. Manual methods such as Wolfe patterns and BI-RADS are helpful for communication of diagnostic sensitivity, but they are both time consuming and crude. For serial, temporal analysis it is necessary to be able to detect more subtle changes.

In previous work, a method for measuring the effect of HRT w.r.t. changes in biological density in the breast is described. The method provides structural information orthogonal to intensity-based methods. Hessian-based features and a clustering of these is employed to divide a mammogram into four structurally different areas. Subsequently, based on the relative size of the areas, a density score is determined.

We have previously shown that this method can separate patients receiving HRT from patients receiving placebo. In this work, the focus is on deeper understanding of the methodology using tests on sets of artificial images of regular elongated structures.

1 Introduction

Numerous studies have investigated the relation between mammographic density and breast cancer risk, and women with high breast density appear to have a four to six fold increase in breast cancer risk, e.g. [9, 2, 1, 7]. Therefore *density* is an important feature embedded in a mammogram. Currently, however, the density is not used to asses risk in the standard clinical screening procedures.

The specific purpose of this work is to investigate the nature of the actual structural changes in the breast tissue caused by hormone replacement therapy (HRT) detected by our clustering technique. This Hessian-based method has been validated in a previous experiment, using two sets of mammograms of 50 patients from a double blind, placebo controlled HRT experiment [8]. The method was able to significantly separate the HRT patients from placebo patients ($p = 0.0002$) [5].

The method is interesting seen both from a practical, image analysis perspective, but also from a medical point of view, where is might provide insight into important anatomical changes relating to density alterations. In order to get this insight, we have to have an in depth understanding of the method used, which is

Susan M. Astley et al. (Eds.): IWDM 2006, LNCS 4046, pp. 447–452, 2006.

the focus of the study presented. To achieve understanding we do tests on sets of canonical images of regular stripe-like patterns of different frequency.

2 Methods

Detecting HRT Using Hessian-Based Pixel Classification

The breast tissue is manually segmented. Within this region of interest (ROI), for every pixel, features based on eigenvalues of Hessian at three scales are determined. The Hessian at scale s is defined by

$$H_s(I) = \begin{bmatrix} \frac{\partial_s^2 I}{\partial_s x^2} & \frac{\partial_s^2 I}{\partial_s x \partial_s y} \\ \frac{\partial_s^2 I}{\partial_s y \partial_s x} & \frac{\partial_s^2 I}{\partial_s y^2} \end{bmatrix}$$

where ∂_s denotes the Gaussian derivative at scale s [4]. The scales used are 1, 2 and 4 mm. The features used are given by the quotient:

$$q_s = \frac{|e_1| - |e_2|}{|e_1| + |e_2| + \epsilon}$$

where e_1 and e_2 are eigenvalues of the Hessian at scale s, $e_1 > e_2$ and ϵ is a number much smaller than 1 used to avoid instabilities associated with near zero division. This quotient relates to the elongatedness in an image at a certain location (x, y) at scale s. It is close to zero if image structure is "round" and closer to 1 or +1 for more elongated structures. It is invariant to rotation of the image and locally linear scaling of the intensities.

In a training phase, a large collection of randomly chosen pixels from the different images in the data set are used to generate a representative collection of features. These features are divided into four clusters using k-means clustering [6]. The means are stored and used for nearest mean classification [6].

In the testing phase this nearest mean classifier (NMC) is used to score each mammogram as follows:

- Extract Hessian-based features
- Classify each pixel in one of four classes using the NMC
- Determine relative areas of the classes
- Compute the score from these areas

The score is based on a linear combination of the relative areas of the classes in the breast tissue. The combination is determined using a linear classifier given the HRT group and the placebo group. We assume Gaussian distributions with equal covariance and use the resulting linear Fisher discriminant [6] to separate the placebo and the HRT groups. In the HRT experiment we found that using only two of the classes gave good results and adding information about the other two did not improve the separation significantly.