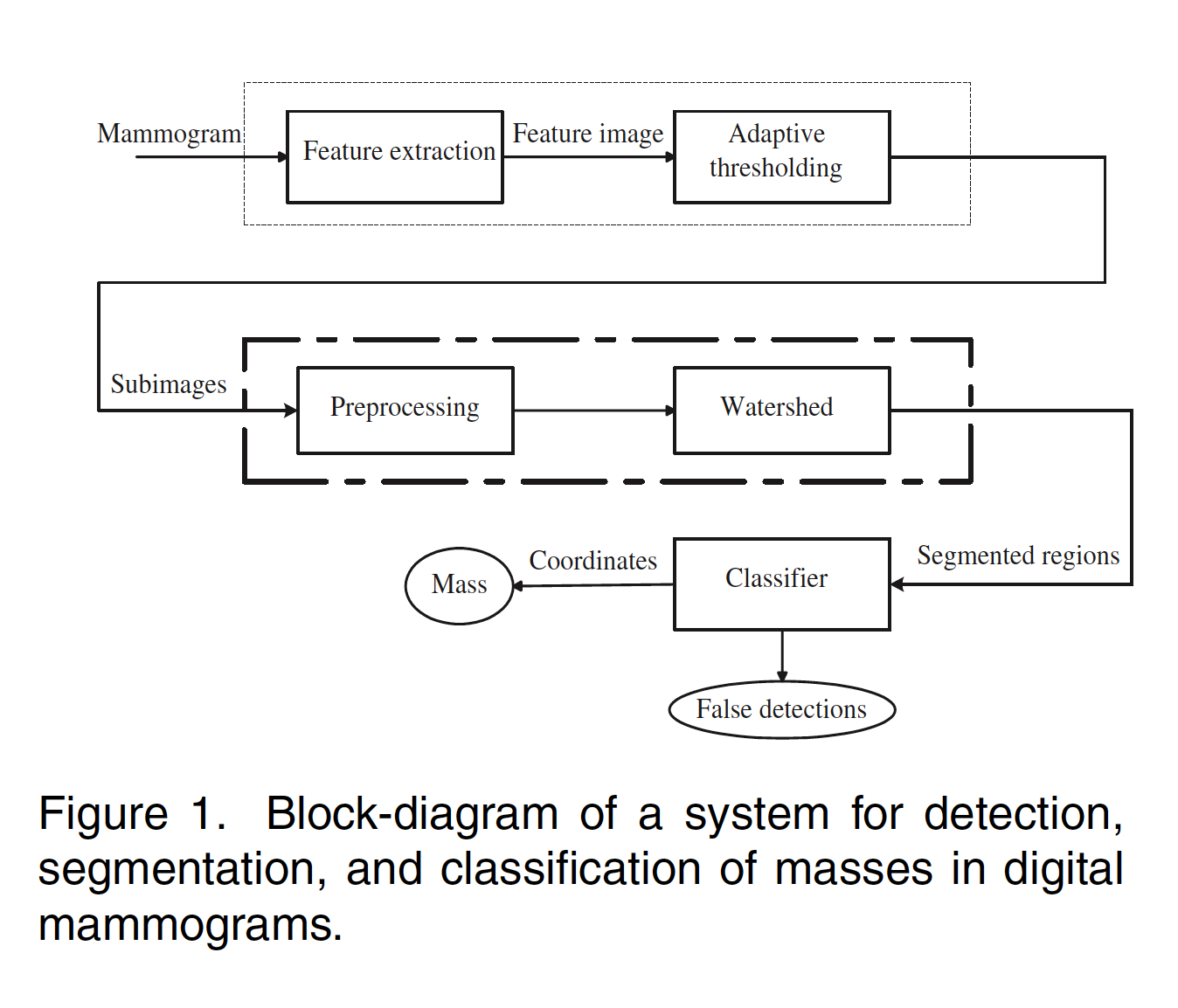
**Watershed segmentation of detected masses in digital mammograms**

Abstract

A method for segmentation of detected masses in digital mammograms is introduced. The method is based on **gray scale mathematical morphology**. In a preprocessing step, **image enhancement based on a local histogram technique is applied**, followed by a morphological smoothing operation. The **watershed transform is then applied to the gradient of the smoothed image resulting in segmented regions.** A good segmentation is important in order to be able to extract useful feature measures from the segmented regions. These feature measures can be **input to a classifier which classifies each region as either a mass or a false detection.** Initial experiments have been performed using mammograms from the **MIAS database**. Results of the experimental study indicate that our scheme **can provide useful contour extraction for mass structures.**

Introduction

A mass is defined to be a region of non-normal breast tissue and appears slightly brighter than normal tissue on a mammogram. Masses have **a range of sizes and shapes, and can be either benign or malignant.** Some masses are associated with spicules, i.e. long thin strands of tissue emanating from the central mass, while other masses can be well circumscribed. Computer based segmentation of mass structures is a challenging task requiring sophisticated techniques. In Figure 1 we show a block-diagram of a computer based system for detection, segmentation, and classification of mass structures. The upper part of the diagram represents the detection part of the system. Details concerning the detection method can be found in [1]. Of course, other detection schemes can also be **applied. Due to the fact that masses often have characteristic similar to the characteristic of normal dense breast tissue, a scheme for detection of masses will generate false positives (FPs). Thus, the outputs from the detection system are subimages containing either a mass structure or a FP.** The subimages are inputs to the segmentation part of the system, shown in the middle part of the block-diagram. The purpose of the segmentation **step is to provide useful contour extraction, i.e. to extract a contour which is close to the true object boundary.** If this is achieved, different feature measures can be extracted from the segmented regions. Finally**, these feature measures can be input to a classifier which classifies each region as either a mass or a FP.** In this paper we present a method for the segmentation of detected masses in digital mammograms, i.e we focus on the middle part of the block-diagram shown in Figure 1. Thus, we assume that the inputs to our proposed scheme are subimages containing a mass structure. Our method is based on the use of the **morphological watershed transform in combination with a preprocessing technique.**



**Materials and methods**

In the present work we use mammograms from theMIAS database [2]. The spatial resolution of the mammograms is 50μmÅ~50μm. Each pixel is represented by 8 bits.

**The watershed algorithm**

A main application of the watershed algorithm is to **segment features (objects) of interest in a gray scale image**. Briefly, the principle of the basic algorithm can be described as follows: **Consider the spatial variation in the signal intensity of a gray scale image as a three-dimensional topographic surface.** The lighter the gray tone of a pixel, the higher the altitude on the topographic surface. **By flooding this surface from each of its local minima, catchment basins having borders corresponding to the borders of objects in the image are obtained**. These borders constitute the **”watersheds” of the image**. As a mass region appears slightly brighter than normal tissue on a mammogram it represents a peak on the topographic surface. Consequently, if we invert the signal intensities of the original mammogram a mass is represented by **a local minima and can as such potentially be segmented by the watershed algorithm.**

**Watershed from markers**

The basic watershed algorithm works well if each local minima corresponds to each object of interest. **Then the watershed lines represent the boundaries of these objects.**

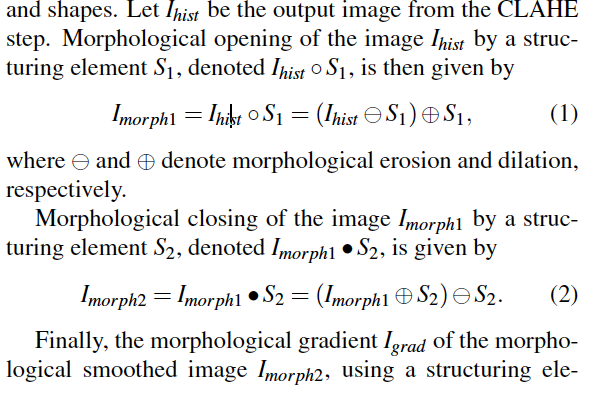
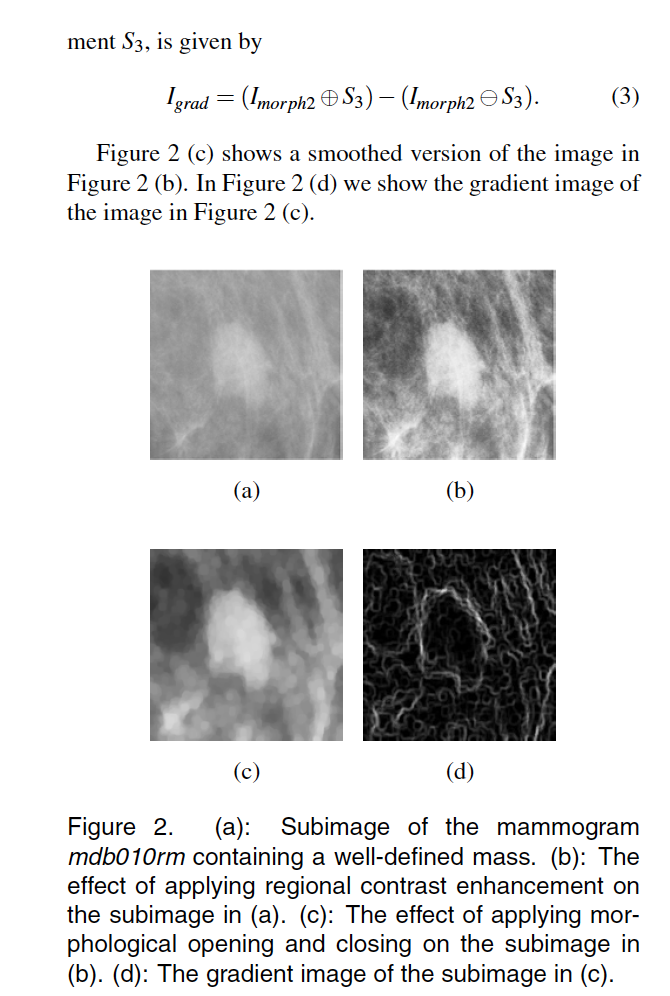
However, for our application there are many more minima in the image than there are objects of interest. As a consequence, the image becomes over-segmented. In order to avoid over-segmentation the basic watershed algorithm can be constrained by an **appropriate marking function** [3]: An internal marker has to be defined for each mass, together with a single external marker. For convenience, **each marker is associated with a color**. As for the basic algorithm, **the watershed from markers can be described as a flooding simulation process**. However, now the topography is flooded from below by letting colored water rise from the marker associated with its color. **Catchment basins forming at minima without markers are flooded by overflow from neighboring catchment basins.** If the rising waters of distinct colors are about to merge, then a dam is built to prevent the merging. **In the present work the inputs to the watershed algorithm are subimages containing a mass structure. In each subimage the center of the object of interest is positioned in the center of the subimage. Thus, an internal marker is automatically placed in the center of the subimage, while an external marker is automatically placed in a frame enclosing the subimage.**

**Preprocessing**

A good preprocessing scheme, aiming at suppressing the normal breast tissue while enhancing the signals from masses, is crucial for the performance of the subsequent segmentation step. In short, our preprocessing scheme consists **of three steps**: **Local contrast enhancement** of the input subimage containing an object of interest, **morphological smoothing** of the contrast enhanced image, **and computation of the morphological gradient of the smoothed image.** **Histogram equalisation techniques** can be useful in visualizing masses and increasing the sensitivity of subsequent processing. In this research regional contrast enhancement is performed by using the **Contrast-Limited Adaptive Histogram Equalisation** (CLAHE) technique [4]. Unlike simple global histogram equalisation techniques, CLAHE operates on small data windows (tiles), instead of the entire image. The contrast of each tile is enhanced, so that the histogram of the output region approximately matches a specified histogram. The neighboring tiles are then combined using bilinear interpolation in order to eliminate artificially

induced boundaries. Figure 2 (a) shows a subimage from the mammogram mdb010rm containing a welldefined mass. The effect of applying regional contrast enhancement

on this subimage is shown in Figure 2 (b). Generally, the watershed transform is not applied to the original image, but to its morphological gradient. This produces watersheds at the points of gray value discontinuity, as is commonly desired in image segmentation. However, as can be observed in Figure 2 (b) structures of normal tissue surrounding the mass may give rise to high gradient values. Thus, prior to computing the gradient image a morphological smoothing is performed on the contrast enhanced image. Morphological smoothing can be achieved by performing a morphological opening followed by a closing. The net effect of these two operations is to remove or attenuate both bright and dark structures of certain sizes and shapes.

**Based on an experimental study we found that our proposed method yields useful detection of mass boundaries. For well-defined masses the extracted contour is close to the true boundary. In addition, for the most spiculated masses, our method captures the more ”random” contour characteristic of these masses. A main future project is to evaluate our scheme using contours of masses drawn by a radiologist as the ”gold standard”.**

