Mammograms Feature Extraction using Fuzzy Surface

Corresponding Author:
Dr. Rash B Dubey,
Professor, ECE Dept, Hindu College of Engg, Sonepat, 121003 - India

Submitting Author:
Dr. Rash B Dubey,
Professor, ECE Dept, Hindu College of Engg, Sonepat, 121003 - India

Article ID: WMC002327
Article Type: Research articles
Article URL: http://www.webmedcentral.com/article_view/2327
Subject Categories: BREAST
Keywords: Breast cancer, fuzzy surface, feature selection and features extraction.
How to cite the article: Dubey R B, Nagpal A. Mammograms Feature Extraction using Fuzzy Surface. WebmedCentral BREAST 2011;2(10):WMC002327
Mammograms Feature Extraction using Fuzzy Surface

Author(s): Dubey R B, Nagpal A

Abstract

Mammography is the most contemporary option for the premature detection of breast cancer in women. The principal feature within the breast region is the breast contour. Extraction of the breast region and delineation of the breast contour is an essential pre-processing step in the process of computer aided detection. Primarily it allows the search for abnormalities to be limited to the region of the breast without undue influence from the background of the mammogram. The methodology involves the use of fuzzy surface for selecting the features of mammograms. Feature extraction is an essential pre-processing step to pattern recognition and machine learning problems. It is often decomposed into feature construction and feature selection. It is well known that mammographic images have some degrees of fuzziness such as indistinct borders, ill-defined shapes, and different densities. Due to the nature of mammography and breast structure, fuzzy logic would be a better choice to handle the fuzziness of mammograms than traditional methods. There are many features such as shape features, texture features etc. The surface viewer is used to display the dependency of one of the outputs on any one or two of the inputs — that is, it generates and plots an output surface map for the system. A variety of samples has been tried out to generate and plot the surface maps.

Introduction

1. Introduction

Breast cancer is the most common type of cancer found in women. It is the most frequent form of cancer and one in 22 women in India is likely to suffer from breast cancer [1]. Breast cancer is the leading cause of death among women in many countries. Detecting a breast cancer at the earliest stage has the most important impact on prognosis. Mammography is the most cost effective method to detect early signs of breast cancer [2] and is the most contemporary option for the premature detection of breast cancer in women. Breast cancer is considered as one of the primary causes of women mortality. The mortality rate in asymptomatic women can be brought down with the aid of premature diagnosis. Despite the increasing number of cancers being diagnosed, the death rate has been reduced remarkably in past decade due to the screening programs. Premature detection of breast cancer increases the prospect of survival whereas delayed diagnosis frequently confronts the patient to an unrecoverable stage and results in death [3]. So far, many systems have been developed to detect the micro-calcification (MCC) in mammograms. They usually detect suspicious regions first and then techniques can be applied to the features of these regions. The existing features for detecting the MCC could be divided into several branches such as shape features, statistical texture features, wavelet features and etc [4]. Two recent advances in mammography include digital mammography and computer-aided detection. Various techniques have been designed for the detection of breast cancer, but all of them are using either genetic algorithm or CAD technology. The use of fuzzy logic is to deal with uncertainty for diagnosis risk status of breast cancer [5].

The remainder of this paper is structured in four sections. In Sections II details of the proposed methodology are presented. Implementation is described in Section III. Results discussions are drawn in Section IV.

Methods

2. Proposed Methodology

The proposed methodology is outlined in Fig. 1.

2.1 Pre-processing

The image pre-processing refers to the initial processing of raw image to correct the geometric distortions, calibrate the data radio metrically and eliminate the noise and clouds that present in the data. These operations are called pre-processing because they normally carried out before the real analysis and manipulations of the data occur in order to extract any specific information. The aim is to correct the distorted or degraded image data to create a more faithful representation of the real scene. The purpose of pre-processing is to remove noise and radiopaque artifacts contained within the mammogram and increase region homogeneity, with the objective being to improve in algorithm reliability and robustness. Mammograms often contain artifacts in the form of...
identification labels, markers, and wedges in the unexposed air-background (non-breast) region. Such artifacts are usually radiopaque in the sense that they are not transparent to radiation. One of the problems with precise segmentation of the breast region is that the existence of such artifacts often results in a non-uniform background region which may cause a segmentation algorithm to fail.

Fig.1: Flowchart.

2.2 Fuzzy logic
Fuzzy logic provides a means of calculating intermediate values between absolute true and absolute false with resulting values ranging between 0.0 and 1.0. It seeks to handle the concepts of partial truth by creating values representing what is between total truth and total false. Fuzzy logic differs from Boolean logic in that it is permissive of natural language queries and is more like human thinking; it is based on degrees of truth [13-15].

2.2.1 Fuzzy rule base
Fuzzy rule-based approach to modelling is based on verbally formulated rules overlapped throughout the parameter space. They use numerical interpolation to handle complex non-linear relationships. Fuzzy rules are linguistic IF-THEN- constructions that have the general form "IF A THEN B" where A and B are propositions containing linguistic variables. A is called the premise and B is the consequence of the rule. In effect, the use of linguistic variables and fuzzy IF-THEN- rules exploits the tolerance for imprecision and uncertainty. In this respect, fuzzy logic mimics the crucial ability of the human mind to summarize data and focus on decision-relevant information [13-15].

2.2.2 Fuzzy surface
Surface viewer is a read-only editor. The rule viewer and the surface viewer are strictly read-only tools. The rule viewer is used as a diagnostic; it can show which rules are active, or how individual membership function shapes are influencing the results. The surface viewer is used to display the dependency of one of the outputs on any one or two of the inputs that it generates and plots an output surface map for the system.

Upon opening the surface viewer, a two-dimensional curve represents the mapping from service quality to tip amount. Since this is a one-input one-output case, we can see the entire mapping in one plot. Two-input one-output systems also work well, as they generate three-dimensional plots that MATLAB can adeptly manage. When we move beyond three dimensions overall, we start to encounter trouble displaying the results. Accordingly, the surface viewer is equipped with pop-up menus that select any two inputs and any one output for plotting. Just below the pop-up menus are two text input fields that determine how many x-axis and y-axis gridlines one want to include. This allows keeping the calculation time reasonable for complex problems. To change the x-axis or y-axis grid after the surface is in view, simply change the appropriate text field, and click on either X-grids or Y-grids, according to which text field one want to change, to redraw the plot [13-15].

2.3 Feature selection
Feature selection is an important part of any classification scheme. The success of a classification scheme largely depends on the features selected and the extent of their role in the model. The objective of performing feature selection is three fold: (a) improving the prediction performance of the predictors, (b) providing faster and more cost effective predictors and (c) providing a better understanding of the processes that generated the data. There are many benefits of variable and feature selection: it facilitates data visualization and understanding, reduces the storage requirements, reduces training times and improves prediction performance.

There are various features of mammograms such as texture, shape features etc., out of all these, there are further various categories of these features. The features which are selected for the fuzzy surface implementation are contour, lines and irregular boundary etc.

A mammogram contains two distinctive regions: the exposed breast region and the unexposed air-background (non-breast) region. The principal feature on a mammogram is the breast contour, otherwise known as the skin-air interface, or breast boundary. The breast contour can be obtained by partitioning the mammogram into breast and non-breast regions. The extracted breast contour should adequately model the soft-tissue/air interface and preserve the nipple in profile.

The largest single feature on a mammogram is the skin-air interface, or breast contour. Extraction of the breast contour is useful for a number of reasons. Foremost it allows the search for abnormalities to be limited to the region of the breast without undue influence from the background of the mammogram. Segmentation of the breast-region from the background is made difficult by the tapering nature of the breast, such that the breast contour lies in between the soft tissue and the non-breast region. The precise segmentation of the breast region in mammograms is an essential preprocessing step in the computer-aided analysis of mammograms for a number of reasons [4-5,7-9].

2.4 Graphical User Interface (GUI)
A GUI allows a computer user to move from
application to application. A good GUI makes an application easy, practical and efficient to use and the marketplace success of today’s software programs depends on good GUI design.

In computing a GUI is a type of user interface that allows users to interact with electronic devices with images rather than text commands. GUIs can be used in computers, hand-held devices such as MP3 players, portable media players or gaming devices, household appliances and office equipment. A GUI represents the information and actions available to a user through graphical icons and visual indicators such as secondary notation, as opposed to text-based interfaces, typed command labels or text navigation. The actions are usually performed through direct manipulation of the graphical elements. The term GUI is historically restricted to the scope of two-dimensional display screens with display resolutions able to describe generic information. GUIDE stores GUIs in two files, which are generated for the first time when GUI is saved or run, the two files are namely; fig file which contains a complete description of the GUI figure Layout and the components of the GUI and m file which contains the code that controls the GUI.

2.5 Edge detection

Edge detection is a fundamental tool in image processing and computer vision, particularly in the areas of feature detection and feature extraction, which aim at identifying points in a digital image at which the image brightness changes sharply or, more formally, has discontinuities. The purpose of detecting sharp changes in image brightness is to capture important events and changes in properties of the world. It can be shown that under rather general assumptions for an image formation model, discontinuities in image brightness are likely to correspond such as discontinuities in depth, discontinuities in surface orientation, changes in material properties and variations in scene illumination.

There are many methods for edge detection. Canny edge detection is us here. John Canny considered the mathematical problem of deriving an optimal smoothing filter given the criteria of detection, localization and minimizing multiple responses to a single edge. He showed that the optimal filter given these assumptions is a sum of four exponential terms. He also showed that this filter can be well approximated by first-order derivatives of Gaussians. Canny also introduced the notion of non-maximum suppression, which means that given the pre-smoothing filters, edge points are defined as points where the gradient magnitude assumes a local maximum in the gradient direction [6].

2.5.1 Thresholding and linking

Once we have computed a measure of edge strength (typically the gradient magnitude), the next stage is to apply a threshold, to decide whether edges are present or not at an image point. The lower the threshold, the more edges will be detected, and the result will be increasingly susceptible to noise and detecting edges of irrelevant features in the image. Conversely a high threshold may miss subtle edges, or result in fragmented edges.

If the edge thresholding is applied to just the gradient magnitude image, the resulting edges will in general be thick and some type of edge thinning post-processing is necessary. For edges detected with non-maximum suppression however, the edge curves are thin by definition and the edge pixels can be linked into edge polygon by an edge linking (edge tracking) procedure. On a discrete grid, the non-maximum suppression stage can be implemented by estimating the gradient direction using first-order derivatives, then rounding off the gradient direction to multiples of 45 degrees, and finally comparing the values of the gradient magnitude in the estimated gradient direction. A commonly used approach to handle the problem of appropriate thresholds for thresholding is by using thresholding with hysteresis. This method uses multiple thresholds to find edges. We begin by using the upper threshold to find the start of an edge. Once we have a start point, we then trace the path of the edge through the image pixel by pixel, marking an edge whenever we are above the lower threshold. We stop marking our edge only when the value falls below our lower threshold. This approach makes the assumption that edges are likely to be in continuous curves, and allows us to follow a faint section of an edge we have previously seen, without meaning that every noisy pixel in the image is marked down as an edge. Still, however, we have the problem of choosing appropriate thresholding parameters, and suitable thresholding values may vary over the image.

2.5.2 Edge thinning

Edge thinning is a technique used to remove the unwanted spurious points on the edge of an image. This technique is employed after the image has been filtered for noise (using median, Gaussian filter etc.), the edge operator has been applied to detect the edges and after the edges have been smoothed using an appropriate threshold value. This removes all the unwanted points and if applied carefully, results in one pixel thick edge elements. Sharp and thin edges lead to greater efficiency in object recognition.

Implementation
3. Implementation
The algorithm works as follows:
• First of all input image is taken from database and then region of interest is selected.
• Boundaries are made with the help of edge detection algorithms. The classification measures are taken as the input of a fuzzy decision making process with two inputs and one output, purpose is to calculate the degree of membership to which the pixel belongs to the four types (contour, region, etc).
• Fuzzy logic is applied to the classification measures. The input space of the linguistic variable is comprised of the three fuzzy sets (low, med, high), and is comprised of two fuzzy sets labelled (low, high). On the basis of the low, medium and high logic, groups can be distinguished.
• Once an approximation of the breast region has been derived from the fuzzy segmentation, regions can be classified during the processing. Using fuzzy logic concept, the features used for fuzzy rule are contour, lines and irregular boundary of an image, the range is specified there are 3-input ranges one for finding the contour and other one for lines and last one is for defining irregular boundary.
• The rule base is defined and fuzzy surface is viewed and the m-file is linked with fis file of MATLAB. In fis file inputs are selected as; Input range 1: 30-50, Input range 2: 50-80 and Input range 3: 80-100.
• In output, there will be three membership function defined one for contour, second one for irregular boundary and third one for lines, these are the features are selected for fuzzy rule base and these will be grouped on the basis of range and then classified using segmentation as Output 1: Contour: Range 30-50, Output 2: Irregular boundary: 50-80 and Output 3: lines: 80-100.
• The various other steps involved in making fuzzy rule file are describes as under:

Fig 2: Fis Editor file in MATLAB
Fig 3: Fuzzy rule base
- As the rules are defined within the range specified, after this rules can be viewed in MATLAB. The output can be checked for different combinations and then feature selection can be done. The rules are viewed, can be checked for different configuration by changing the input values, the output can be checked. Once the rules have been defined the surface can be viewed. The Fuzzy rule viewer and Fuzzy surface viewer are shown in Fig 4 and Fig 5.

Fig 4: Fuzzy rule viewer
Fig 5: Fuzzy surface viewer
- As the surface is viewed, the contour obtained is shown in Fig 6. After linking the fis file with m-file, the GUI will appear as shown in Fig 7, in the screen of GUI image can be browse and then it can be view using the fuzzy technique. The features contour, irregular boundary and lines are calculated. The output with contour and lines are found maximum at 229.

Fig 6: Fuzzy surface contour
Fig 7: Screen of GUI Implementation in MATLAB
Fig 8: GUI Implementation of image (mdb001.jpg)

References

Illustrations

Illustration 1

Figures

![Flowchart]

Select the Region

Input the Image

Make boundaries with the help of Edge Detection

Fuzzy decision process making with two inputs and

Set Fuzzy sets and group for feature Selection

Basis of low, high groups can be distinguished

Classification can be derived on the basis of fuzzy logic
Fig 8: GUI Implementation of image (mdb001.jpg)

(a) Image mdb075

(b) Image mdb075

(a) Image mdb092

(b) Image mdb092
Fig 9: GUI Implementation of various images.
Disclaimer

This article has been downloaded from WebmedCentral. With our unique author driven post publication peer review, contents posted on this web portal do not undergo any prepublication peer or editorial review. It is completely the responsibility of the authors to ensure not only scientific and ethical standards of the manuscript but also its grammatical accuracy. Authors must ensure that they obtain all the necessary permissions before submitting any information that requires obtaining a consent or approval from a third party. Authors should also ensure not to submit any information which they do not have the copyright of or of which they have transferred the copyrights to a third party.

Contents on WebmedCentral are purely for biomedical researchers and scientists. They are not meant to cater to the needs of an individual patient. The web portal or any content(s) therein is neither designed to support, nor replace, the relationship that exists between a patient/site visitor and his/her physician. Your use of the WebmedCentral site and its contents is entirely at your own risk. We do not take any responsibility for any harm that you may suffer or inflict on a third person by following the contents of this website.