Department of Computer Engineering
Boğaziçi University

CMPE 537: Computer Vision

Term Project Final Report
Prenatal Diagnosis Using Fetal Ultrasound in CAD Systems

Nuri Murat Arar 2010700009
Fatma GÜNEY 2010700090
Beyza Ermiş 2010700057

Instructor: Lale Akarun

20/01/2011
**TABLE OF CONTENTS**

Introduction 3

Related Works 4

Offline System 7
  - Image Annotation 7
  - ASM Building 8

Online System 8
  - Image Enhancement 8
  - Head Boundary Detection (Segmentation) 8
  - ASM Fit 10
  - Boundary Extraction 11
  - Feature Extraction 11
    - Hu Moments 12
    - Fourier 12
  - Classification and Results 14
    - K-Nearest Neighbor 14
    - Multivariate Parametric Classification 14

References 15
INTRODUCTION

In this project, it is aimed to construct a system which reduces the work load of human experts, health specialists and doctors in prenatal diagnose processes. We intend to work in fetal ultrasound images in order to detect and diagnose prenatal disorders since early detection of these disorders might be very helpful for the mother’s and the child’s health. More specifically, this system is planned to use for spina bifida pathology detection and diagnose. Spina bifida is one of the common birth defects called neural tube defects (NTD). It occurs when the neural plate folds incorrectly. A person with spina bifida experiences some lesions such as loss of bladder control and paralysis. It occurs in 7/10000 live births in USA. [1] Therefore, it is a very important pathology and is needed to be diagnosed in prenatal period.

The prominent symptom of spina bifida is an anomaly in the shape of skull as shown in Figure-1. Therefore, examination of fetal ultrasound images which are called as “transcerebellar head images” is necessary for the presence of lemon sign to detect the possibility of spina bifida.

![Figure 1: left) Healthy skull  right)Lemon Sign](image)

The presence of lemon sign can be determined by the localization of fetal head, extraction of the skull boundaries and extraction and examination of features indicating the lemon sign.

First, approximate head boundary is found as the first step of segmentation.

Then, active appearance model (ASM) is used to find exact skull boundary.

Fast Fourier transform is performed on the skull boundary of fetal head and the features of the every skull is extracted.
Finally, KNN classification method and multivariate parametric classification method are used to separate the healthy and suspicious skulls.

**Figure 2: System Overview**

**RELATED WORKS**

1. **Region-Based Feature Extraction of Prostate Ultrasound Images [8]:** This thesis tries to automate the cancerous region detection process. The aim is to measure of some characteristics (e.g. darkness, texture). This thesis proposes a new feature extraction method: **Spatial** location, symmetry, and other geometric measurements of the regions-of-interest, in addition to grieve and texture. They use a semi-automatic fuzzy inferencing system (FIS) to relate all the features and mimic radiologists’ knowledge. They use Fourier transform to decompose prostate images into pure frequencies. In addition to Fourier transform, they also use Gabor filters for feature extraction. However, the biggest difference with our work is their segmentation methods. They use some region segmentation methods: -**Graph-theory-based** method by constructing Minimum Spanning Tree (MST). -**Thresholding** on histogram.
2. Recognizing and classifying leaves [7]: In this work, they also utilize from the Fourier descriptors. Two types of leaves are to be recognized and classified, at the end.
3. Multi-level Shape Recognition based on Wavelet-Transform [2]: In this paper they propose a new approach to shape recognition based on the wavelet transform modulus maxima. And apply it to the problem of content-based indexing and retrieval of fish contours. The description scheme and the similarity measure developed take into consideration the way our visual system perceives objects and compares them.
In our project, the system consists of offline and online processes. In offline system, we selected thirty images and annotated them for building an Active Appearance Model. After offline system, our system is ready to perform the rest of the classical object recognition processes such as enhancement, segmentation, ASM fit, boundary extraction, feature extraction and classification.
OFFLINE SYSTEM PROCESS

1. IMAGE ANNOTATION FOR ASM

For annotating the images, we used Timothy Cootes’s ASM build tool [6]. We marked 17 base points on the boundary of the head skeleton manually. Then, the tool fills each of the two consecutive base points with 9 points. Therefore, a head boundary is represented by 161 points in each image. We selected thirty images of different shape, size and intensity in order to make ASM robust to these kinds of variations. We tried annotating with different base points such as 9 and 33 as well as 17 but we achieved the best model with 17 and used them in the final implementation.
2. ASM BUILDING

Active Appearance Model building process is performed only once at the beginning. It is a simple but a very significant process since ASM fitting process is directly depend on it. Therefore, it affects the accuracy of our overall system very much. To build our general ASM, we used the annotated dataset consisting of 30 images. Points of correspondence are placed on each image so the dataset can be readily used for building statistical models of shape. As a result of this process, we have a reference model used by ASM fitting step [5].

ONLINE SYSTEM PROCESS

1. IMAGE ENHANCEMENT
2. HEAD BOUNDARY DETECTION (SEGMENTATION)

To extract the boundary from rest of the image, we need to do segmentation. It was one of the most challenging parts of the project. First of all, we try to use 'jseg' algorithm which separate the segmentation process into two independently processed stages, color quantization and spatial segmentation. In the first stage, colors in the image are quantized to several representing classes that can be used to differentiate regions in the image. This quantization is performed in the color space alone without considering the spatial distributions. Afterwards, image pixel colors are replaced by their corresponding color class labels, thus forming a class-map of the image. However the results are not good enough to continue for working, because 'jseg' is a color-based segmentation method. Our images only include pixels close white or close black, so this segmentation method is not suitable for us. Then, we try 'normalized cut' which propose a new graph-theoretic criterion for measuring the goodness of an image partition segmentation method, however it also did not work well. After that, we tried Edison algorithm which is based on edge detection first, and it did not work well because the ultrasound images have not clear edges and segmentation by edges could not be a logical way for that kind of images.
Finally, we decided to use 'ellipse fit' for segmentation. Because, we need only the head boundaries to detect lemon sign in the skull boundary, hence we could separate the suspicious skulls from the healthy skulls.

Ellipse fit algorithm firstly selecting three random points that are not col-linear and by using these parameters, computes the ellipse’s center and ellipse's own parameters that are within the image borders. Then, it fits the best ellipse using random points of LSimage (longest skeleton segments). This algorithm re-organize the given skeleton image by removing patches that touch the sides of the image (they are likely to represent incomplete structures) such that the longest contour along the skeleton of an object is preserved while the others are discarded. In addition, if cyclic contours are present, they are preserved if their size is considerably large [1]. As a result, we get nearly expected results from this algorithm:

![Segmentation result of ellipse fit](image)

This algorithm gives the approximate boundary of the head in the images, but we need better ellipses for classification. So, we used active appearance model to get better segmentation results.

3. ASM FIT

This step is the most crucial part of our project, when we had a good fit, then the result becomes very successful, otherwise, we got an ambiguous result. For ASM fitting
process, providing an initial location for the reference shape makes the process more efficient. In our case, the initial shape for the detected head boundary is obtained by the ellipse fit process. Hence, we give a full pre-built ASM model and an initial location of the head boundary as a reasonable approximation to ASM fitting. ASM fitting is the process of adjusting model parameters so that it produces a fit shape that best represent the head boundary by converging through iterations [5]. The output of this process is the best representing 161 points on the head boundary.

Figure 9: Convergence of ASM Fit Process; upper left is the beginning iteration, bottom right is the converged ASM Fit Result
4. BOUNDARY EXTRACTION

In boundary extraction process, our main aim is finding the exact head boundary pixel locations to be used to construct our feature vectors in feature extraction step. We specifically do not try to obtain the head boundary as a ring in given input image because for feature extraction using fast Fourier transform it is better to represent the head boundary in terms of set of x and y coordinates as stated in Zhang and Lou’s work [8].

After ASM fit process, we have 161 points on the detected head boundary. This information may be sufficient for some of the cases; however, there may be a problem with ASM fit because the input image may not be suitable in terms of intensity, blur, noise, size and boundary pixel continuity or ASM fit itself may not converge well because of the implementation. In order to overcome these problems, we made a final check for the head boundary pixels. In this process, each of 161 candidate points is checked within 3x3 window whether it has the highest pixel intensity within the window or not. If not, we kept the pixel coordinates of the pixel having highest intensity. Hence, the output of the boundary extraction step is the set of the coordinates of the most likely 161 head boundary pixels.

Figure 10: Extracted head boundaries; left is spina bifida suspicious, right is healthy.
5. FEATURE EXTRACTION

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. Analysis with a large number of variables generally requires a large amount of computation power or a classification algorithm which overfit the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. In this work, we use input images that fit the training samples and generalize to new samples. So, we need to make feature extraction of head boundaries to differentiate the healthy and suspicious skulls and detect the suspicious ones. We used two different feature extraction algorithms: Hu moments and Fourier descriptors.

5.1. HU MOMENTS

Image moments are useful to describe objects after segmentation. Simple properties of the image which are found via image moments include area (or total intensity), its centroid, and information about its orientation. Common region based methods use moment descriptors to describe shape. Region moment representations interpret a normalized gray level image function as a probability density of a 2D random variable. The first seven invariant moments, derived from the second and third order normalized central moments, are given by Hu.

Area is given by the 0th moment

\[ A = \sum_{x} \sum_{y} I(x,y) \]

Center of mass is given by the first moment

\[ \bar{x} = \frac{\sum_{x} \sum_{y} xI(x,y)}{\sum_{x} \sum_{y} I(x,y)} \]
\[ \bar{y} = \frac{\sum_{x} \sum_{y} yI(x,y)}{\sum_{x} \sum_{y} I(x,y)} \]

Figure 11: Representation of image moment
In the project, we find the first, second and eight Hu moments of every skull boundary after segmentation and use these values as features for classification. However, we cannot get the expected results by using the Hu moments, because moments combine information across an entire object rather than providing information just at a single boundary point, they capture some of the global properties missing from many pure contour-based representations: overall orientation, elongation, etc. We need to detect the lemon shape on the skull boundary, but Hu moments use these features in addition to the shape features, so we cannot get accurate results.

5.2. FOURIER DESCRIPTOR

Shape is one of the most important features in Content Based Image Retrieval. Many shape representations and retrieval methods exist. However, most of those methods either do not well represent shape or are difficult to do normalization (making matching hard). Among them, methods based Fourier descriptors achieve both well representation and well normalization. Fourier descriptors are concise and description of (object) contours where contours are represented by vectors and contour processing is used for characterizing and recognizing the shapes of object.

A DFT represents a contour as:

\[ \tilde{U} = \begin{pmatrix} x_0 + iy_0 \\ x_1 + iy_1 \\ \vdots \\ x_N + iy_N \end{pmatrix} \]

**1st Step**
Define a complex vector using coordinates \((x,y)\).

**2nd Step**
Apply the 1D DFT

\[ \tilde{F}_\mu = FFT[\tilde{U}] = \sum_{k=0}^{N-1} \tilde{U}_k \exp\left(-\frac{2\pi i}{N} k \mu \right) \]

\((x_N,y_N)\): Coordinates of the Nth point along the circumference

First step: Defines a complex vector using coordinates \((x,y)\)    Second step: Apply the 1D DFT
In this project, we use Fourier Descriptors derived from centroid distance based shape signature. First, we find the center of mass of every head boundary. Then, we calculate the distances between the center of mass of head boundary and each point on the ASM fitted. Since the sizes of the skulls are different in every image, firstly a normalization is required using the first coefficient of the Fourier Descriptor. After that, we give the result of normalized values to the fast Fourier transform as input and we get the discrete Fourier transform of the input vector [8]].

Figure 11: Healthy skull and its DFT
Fourier series, capture the more general shape properties while the later terms capture finer detail. However, unlike Fourier series, it is difficult to obtain higher order invariant moments and relate them to shape. Comparing with region based shape representation (moments), contour based shape representation (Fourier transform) is more popular. Contour based shape representation only exploits shape boundary information, these representation methods can be classified into global shape descriptors, shape signatures and spectral descriptors. With Fourier descriptors, global shape features are captured by the first few low frequency terms, while higher frequency terms capture finer features of the shape. Apparently, Fourier descriptors do not only overcome the weak discrimination ability of the moment descriptors and the global descriptors but also overcome the noise sensitivity in the shape signature representations. Other advantages of FD method include easy normalization and information preserving.

6. CLASSIFICATION AND RESULTS

Classification is a problematic process for spina bifida disease detection. Even we tried to classify by looking at our eyes, we cannot exactly sure whether it is suspicious or not. Therefore, determining a training set and a test set is very painful and ambiguous. For classification, we performed two supervised algorithm; one is K-Nearest Neighbor and the other one is Multivariate Parametric Classification. We selected 82 images as training set which consists of 31 images having spina bifida disease (suspicious) and 51 healthy images. We selected 20 images as test set which consists of 7 images having spina bifida disease (suspicious) and 13 healthy images.

6.1. K-NEAREST NEIGHBOR

We used a simple supervised classification algorithm, k nearest neighbor (KNN), to classify each image whether it has spina bifida disease or not. KNN takes a test image set, the training set, training set labels and k, number of neighbors. It simply calculates the distance between a test image feature vector and each of the training image feature vector. Then, it takes the labels of the k training images having minimum k distances. Then, it labels the test image with the label of the most frequent label. Our classification results are as follows:
### 6.2. MULTIVARIATE PARAMETRIC CLASSIFICATION (GAUSSIAN)

This classification algorithm simply fits two Gaussians each having different mean and variance for spina bifida images and healthy images. Then, it checks the correlation between these two Gaussians and determines which of the features of the test images belongs to which class. The only problem of this method is that our training image set is not distributed uniformly because we do not have sufficient number of spina bifida images. Therefore, if the algorithm confuses, it tends to classify the test image as healthy which is not very reliable. Our classification results are as follows:

<table>
<thead>
<tr>
<th>K = 1</th>
<th>Correct Classifications</th>
<th>Correct Spina Bifida Classification</th>
<th>Correct Healthy Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>13 / 20</td>
<td>4/5</td>
<td>9/15</td>
<td></td>
</tr>
<tr>
<td>K = 3</td>
<td>13 / 20</td>
<td>3/5</td>
<td>10/15</td>
</tr>
<tr>
<td>K = 5</td>
<td>14 / 20</td>
<td>2/5</td>
<td>12/15</td>
</tr>
</tbody>
</table>

### REFERENCES


[3] Adam C. Hodge, Aaron Fenster, Donal B. Downey and Hanif M. Ladak, Prostate boundary segmentation from ultrasound images using 2D active shape models: Optimisation and extension to 3D

[4] Faouzi Alaya Cheikh, Azhar Quddus and Moncef Gabbouj, Multi-level Shape Recognition based on Wavelet-Transform, Tampere University of Technology (TUT)


[8] Dengsheng Zhang and Guojun Lu, A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures, Gippsland School of Computing and Information Technology