An Information Tracking Approach to the Segmentation of Prostates in Ultrasound Imaging

by

Robert Sheng Xu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2010

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Outlining of the prostate boundary in ultrasound images is a very useful procedure performed and subsequently used by clinicians. The contribution of the resulting segmentation is twofold. First of all, the segmentation of the prostate glands can be used to analyze the size, geometry, and volume of the gland. Such analysis is useful as it is known that the former quantities used in conjunction with a PSA blood test can be used as an indicator of malignancy in the gland itself. The second purpose of accurate segmentation is for treatment planning purposes. In brachetherapy, commonly used to treat localized prostate cancer, the accurate location of the prostate must be found so that the radioactive seeds can be placed precisely in the malignant regions. Unfortunately, the current method of segmentation of ultrasound images is performed manually by expert radiologists. Due to the abundance of ultrasound data, the process of manual segmentation can be extremely time consuming and inefficient. A much more desirable way to perform the segmentation process is through automatic procedures, which should be able to accurately and efficiently extract the boundary of the prostate gland with minimal user intervention. This is the ultimate goal of the approach presented in this thesis.

Although the choice of ultrasound imaging modality has many advantages over other modalities such as CT and MRI in terms of its high benefit-to-cost ratio and prevalent accessibility, ultrasound imaging also has intrinsic disadvantages that makes the segmentation task at hand non-trivial. It is well known that the quality of ultrasound images is relatively poor because of its low contrast ratio and the presence of speckle noise and shadowing artifacts. All of the aforementioned qualities makes the prostate boundary poorly defined in clinical ultrasound images. In order to overcome the difficulties involved, the overall segmentation task is divided into three separate smaller tasks i.e. preprocessing, feature extraction, and inclusion of prior shape knowledge. The proper choices of each individual task is explored to achieve the best possible performance.

Finally, the proposed segmentation algorithm uses a probability distribution tracking framework to accurately and efficiently perform the task at hand. The basis for this methodology is to extract image and shape features from available manually segmented ultrasound images for which the actual prostate region is known. This information characterizing the prostates are then encoded in the form of statistical models (i.e. probability density functions). Then, during the subsequent segmentation, the proposed algorithm seeks to find a region in the new ultrasound images whose features closely mirror the learned features of known prostate regions. This is a regional based approach, as the empirical features within the segmenting region are also encoded into a probability distribution, and minimizing the distance between the empirical and model distributions becomes the main segmentation criterion of the proposed approach. Promising results were achieved using this method in a series of *in silico* and *in vivo* experiments.

Acknowledgements

I would like to thank my supervisors Dr. Magdy Salama and Dr. Oleg Michailovich for providing me with countless hours of fruitful discussions and technical expertise. Without their encouragement and support along the way, the work presented here would not have been possible. I would also like to thank my family for always being there for me in every step of the way.

Dedication

This is dedicated to my loving parents, Charlie Xu and Alice Dong. Without your unconditional love and support through all these years, I wouldn't have come this far.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

According to recent global cancer statistics, there were 679,000 new cases of prostate cancer worldwide in 2002, making it the fifth most common type of cancer in the world and the second most common in men [2]. The cost of treating this type of cancer runs into millions of dollars annually, while the emotional impact on the patients and their families is tremendous. Therefore, a desirable way to alleviate the financial and emotional strain of cancer patients lies in the form of early diagnosis. When cancer is detected in its early stages, treatment becomes much more effective and the patient's chance of survival is dramatically increased.

Currently, the two main methods of diagnosis consist of the digital rectal examination (DRE) and the prostate specific antigen (PSA) blood tests. Although PSA tests remains of the most popular screening tests, the elevated levels of PSA in itself is not definitive in the analysis of prostate cancer as other factors can also affect the levels of PSA. A more accurate screening method measures the PSA density (PSAD) of the patient. As the name suggests, PSAD measures the concentration of PSA in the patient's prostate, or the PSA value per cubic centimetre of the prostate. Elevated levels of measured PSAD compared with standard population norms suggests potential for disease. Moreover, PSAD can help doctors distinguish between benign prostatic hyperlasia (BPH) from prostate cancer. The former is a non-cancerous growth of the prostate gland, which is a condition frequently found in men over the age of 50 [1]. If a patient exhibits elevated PSAD levels, a prostate needle biopsy procedure is then performed to surgically remove pieces of tissue from different areas of the prostate to be examined for a definitive conclusion. Needless to say an accurate PSAD measurement is desired to avoid unnecessary biopsy procedures that are invasive in nature. On the other hand, an underestimate of the PSAD levels can be

even more devastating as early stages of cancer can potentially be undetected. Therefore, accurate estimates of both the PSA levels and the volume of the patient's prostate are required.

Currently, the method of measuring the prostate volume is through the use of transrectal ultrasound (TRUS) scan of the prostate gland followed by manual 2D slice segmentation performed by the radiologist. This is a highly time-consuming task, therefore an automatic segmentation procedure with minimal user input is considered desirable and necessary.

1.2 Objectives

The ultimate goal of this thesis is to find a combination of methods which will efficiently delineate the prostate boundaries in ultrasound images with high accuracy and robustness. To this end, the problem of ultrasound segmentation is broken down into three smaller problems consisting of pre-processing, feature extraction and inclusion of prior shape knowledge. Finally, we wish to integrate the solutions to each of these objectives into a unifying segmentation framework.

Due to the presence of artifacts such as speckle noise, shadowing and low contrast, traditional image processing and computer vision techniques are known to be difficult to apply in the case of ultrasound images. In order to segment the prostate gland with high accuracy, useful edge features corresponding to the boundary of the prostate must be found. Unfortunately, due to the presence of noise, classic edge detectors such as Canny and Sobel will detect spurious edges and may even fail to detect the relevant edges associated with the true prostate boundary. Therefore, the first sub-objective is to implement a pre-processing filter that is capable of reducing noise in TRUS images while preserving the relevant edge information. Hence, improving the chances of useful edge detection.

The second step to our problem involves the proper selection of features capable of discriminating between the image background and the object of interest. Image intensity is an obvious and simple choice of discriminating feature. However, again, due to the noise and shadowing artifacts, intensity alone may not be able to discriminate between the regions of background and foreground. An analysis of discriminating power of different texture and image features needs to be performed to find the proper subset of features.

Another sub-objective of this thesis is to use a proper shape descriptor that will improve the chances of accurate segmentation. Due to the many deficiencies of TRUS images, it may not be possible to detect all the necessary edge information or find a set of features that clearly discriminates the prostate region from the background region. In that case, the segmentation algorithm can still be improved through the incorporation of prior shape knowledge. Since the prostate gland can only have certain plausible configurations, the

segmentation algorithm can be restricted via shape priors. Thus, different schemes for shape prior incorporation will be analyzed and a suitable scheme will be presented for our purposes.

1.3 Thesis Overview

- Chapter 2: will provide background information regarding the basics of the ultrasound imaging modality
- Chapter 3: provides a literature review of previous methods
- Chapter 4: will discuss the proposed algorithm and the completion of the above mentioned objectives
- Chapter 5: shows the experimental results of the proposed algorithm
- Chapter 6: provides the discussions and conclusions to the overall study

Chapter 2

Ultrasound Imaging

2.1 Introduction

The history of ultrasound invention dates back to World War II, in which the study of radar and sonar equipment were the precursor to the development of ultrasound devices. Since the invention of medical ultrasound imaging in the 1950s, this technology has become an essential tool that is prevalent in clinics and hospitals worldwide. Due to the safe nature of ultrasound imaging at diagnostic power levels, and its relative low cost, the frequency and range of applications of ultrasound imaging in modern practices is only second to X-ray imaging. One of the most widespread applications of ultrasound imaging is to produce images of the unborn child for evaluation of fetal growth and detection of physical abnormalities. Nearly all of the children born after 1970 in Europe and North America have had at least one ultrasound scan [3]. Additional applications include analysis of tumours, bone structures and imaging of soft tissues such as the heart, kidneys, liver, gallbladder and prostate. Needless to say, ultrasound technology has become an important medical diagnostic tool in modern healthcare.

The basic principles of echo ultrasound is demonstrated in Fig.2.1 [3], in which a sound pulse produced by a transducer is transmitted into the human body and at different tissue boundaries echoes will be reflected and received by the same transducer to form an ultrasound image. Due to the difference in physical properties of various human tissues, which leads to changes in sound velocity, reflections occur at the boundaries where non-uniform sound velocities occur. The following sections of this chapter will provide an overview of the physical principles of ultrasound as well as the available echo imaging systems.

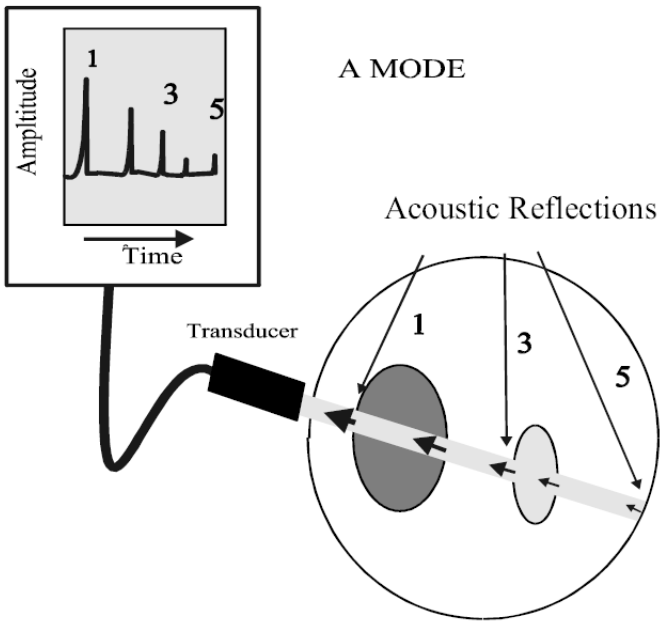

Figure 2.1: The most basic ultrasound scanning mode, A-mode, is demonstrated here. A single pulse of ultrasound is transmitted by the transducer in a line fashion into the patient. Due to the different physical properties of tissues, reflections occur, and the reflected waves that travel back along the same path of transmission is received by the transducer. Depth information is conveyed by the difference in time of the transmitted and received sound wave. Then a subsequent amplitude profile of a single line through the patient is produced [3].

2.2 Physical Principles of Ultrasound

2.2.1 Velocity of Sound

An ultrasound wave propagates in the same manner as all sound waves, which is analogous to a wave of mechanical vibrations. During the propagation, the created pressure disturbances cause local oscillatory movements from one group of atoms to the next along the direction of the sound wave's travel. As a result, this will create a series of compressions and rarefactions in the wave, this is shown in Fig.2.2 [3]. Unlike electromagnetic waves, which can propagate through a vacuum, sound waves must travel through a physical medium. Due to this phenomenon, the frequency and velocity of sound waves heavily depend on the bulk elastic properties of the physical medium and its density.

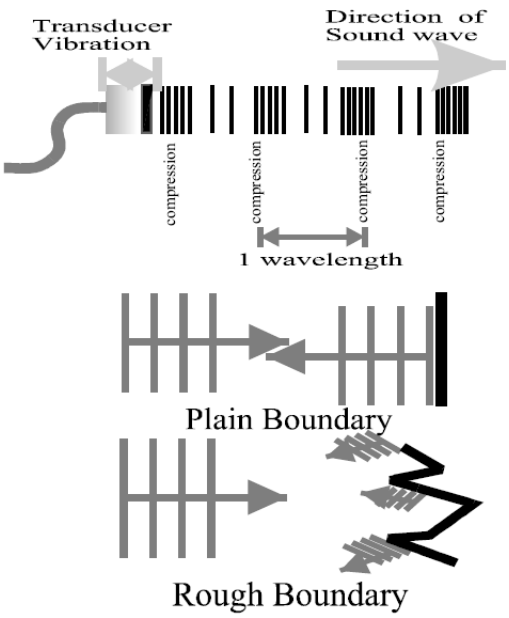

Figure 2.2: A longitudinal wave transmitted from a transducer is shown at the top, the wave shows repeating patterns of compression and rarefaction. The middle image shows ultrasound wave reflections for a smooth boundary. The bottom image shows ultrasound diffuse reflections in the case of rough boundaries, which is characteristic of human tissues.

In general, a physical mediums can support two types of sound waves, one of them being longitudinal or compression waves while the other being shear or transverse waves. In the case of solids, which have both finite shear modulus and bulk compressibility, supports both shear and compressional waves. On the other hand, in the case of liquids and gases, they have a shear modulus of zero, which means the creation of an internal stress is not required to change the shape of a volume of liquid or gas. However, liquids and gases still have a finite bulk compressibility, this allows the propagation of longitudinal waves through both physical mediums. In the majority of ultrasound applications, the velocity of longitudinal waves are of main interest as all soft tissues and body fluids are considered as liquids of varying viscosity. A precise velocity formula is given [3] as:

$$v_{sound} = \sqrt{\frac{K}{\rho}} \tag{2.1}$$

where K and ρ represents the bulk modulus (reciprocal of compressibility) and the density of the material respectively. Provided with the knowledge of speed of sound in

different tissues, the round trip time of a received echo can be calculated in order to provide depth information for an echo imaging system.

The definition of ultrasound is generally any sound wave with its frequency being above the bandwidth of normal human hearing. In adults, this range lies between 10 - 30,000 Hz. A general relationship between velocity of a wave in a physical medium and its frequency f and wavelength λ is as follows:

$$v = f\lambda \tag{2.2}$$

Since the velocity of a wave in a specific medium is defined, there is an inverse relationship between frequency of the wave and its wavelength. Ultimately, for ultrasound imaging systems, the spatial resolution of the acquired images is determined by the wavelength of the transmitted sound pulse. Hence, image resolution is regulated by the frequency of the transmitted sound pulses [3].

2.2.2 Ultrasound Fields

In sharp contrast to radar systems where the object of interest is very far way, the application of medical echo imaging systems generally involves body organs that are close to the ultrasound transducer. The ultrasound field produced by the transducer is divided into two zones, a near zone or Fresnel zone that is within a few wavelength of the transmitter and a far zone or Fraunhofer zone that begins just outside the near zone. In the Fresnel zone, the ultrasound fields have highly variable amplitude and phase. This fluctuation is attributed to significant interference between the waves that are produced at different points on the transmitter. As the waves start to propagate beyond the near zone, the amplitude of the fields becomes more smoothly decreasing, while the width of the ultrasound beam starts to spread laterally (see Fig.2.3) [3].

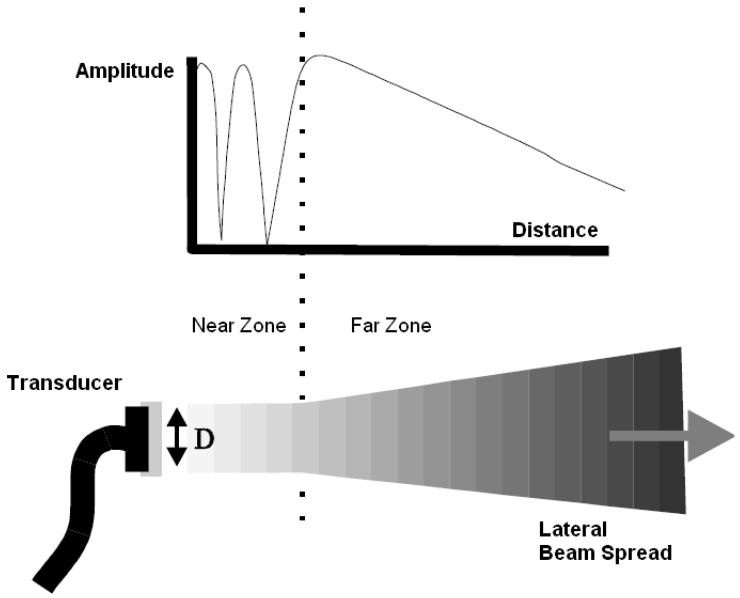

Figure 2.3: The top images shows the amplitude profile of the ultrasound fields in the near and far zone. There is significant fluctuation in the amplitude profile within the near zone caused by both constructive and destructive wave interference. Then, the amplitude profile becomes more smoothly varying as wave interference decreases in the far zone. The bottom image shows the increase in beam width as the waves propagate into the far zone, this correlates with a reduction in lateral resolution (without beam focusing correction).

2.2.3 Attenuation

As ultrasound waves travel through the human body, the energy of the wave decreases in intensity and amplitude as some of the energy is absorbed, while others are lost through scattering. Many tiny cellular structures are encountered by the ultrasound wave as the latter propagates towards the reflecting boundaries. A variety of motions is induced in these cellular structures by the ultrasound waves, but they are heavily damped by viscous friction, causing the organized ultrasound wave energy to transform into random heat energy. Moreover, due to the rough surface of human tissue and the existence of tiny cellular structures, many reflected waves are scattered in all directions, and thus many reflected signals are lost as only the waves that travel back along the transmitted path are detected by the transducer. Attenuation therefore occurs due to the combination of effects from absorption, scattering, and reflection. Moreover, these effects are dependent on frequency of the ultrasound beam as well as the physical properties of the attenuating medium. A general formula governing the transmitted intensity is given [3] as:

$$I = I_o e^{\mu X} \tag{2.3}$$

where $\mu = \mu_{scatter} + \mu_{absorption}$ and X is the distance traveled in the material. Since scattering and absorption are both depth and frequency dependent, the units of attenuation coefficients are measured in dB/cm/MHz. A typical value for attenuation coefficients in ultrasound is 1 dB/cm/MHz, therefore for a 10 MHz ultrasound signal the attenuation becomes approximately 10 dB/cm in soft tissue. As a comparison to X-rays, with its attenuation approximately being 0.8 db/cm, ultrasound is much more heavily attenuated in biological tissues. Moreover, higher frequency ultrasound waves are attenuated more rapidly than lower frequency waves, but generates higher spatial resolution images. Therefore, there exists a trade off between obtaining useful data from deeper depth and resolution [3]

2.2.4 Reflection of Ultrasound Waves

The pattern of reflected and scattered waves from the human interior is extremely complex due to the rough nature (at microscopic scales) of internal organ boundaries and the presence of many tiny cellular structures. Hence, many reflected waves are lost, as ultrasound imaging only detects scattered waves that preserve the organized wave pattern and returns along the incident beam path as an echo. Therefore, the back-scattered component represents a small fraction of the total information presented in the ultrasound fields.

If we only consider the useful back-scattered component of the ultrasound waves and making the idealized assumption of a smooth planar internal reflection boundary between different organs, it turns out that a ratio between the transmitted (T) and the reflected (R) components can be formulated [3] as:

$$\frac{R}{T} = \frac{Z_1 - Z_2}{Z_1 + Z_2} \approx \frac{v_1 - v_2}{v_1 + v_2} \tag{2.4}$$

where Z_1 and Z_2 are defined to be the acoustic impedances of the material on either side of the the boundary respectively. The acoustic impedance Z is generally defined as $Z = \rho v_{sound}$ and v_1, v_2 represents the velocity of sound in the medium immediately beside the reflecting boundary. This idealized generalization is an upper-bound on the reflection ratio between two mediums, but it also helps explain certain limitations of the ultrasound imaging modality. Due to the large velocity mismatch between air and skin, if the transducer is not in good, airless contact with the skin, very little ultrasound waves will be transmitted into the body. Similarly, a large mismatch between soft tissue and bone structure exists, which explains why ultrasound is a bad choice for imaging the internal structure of the brain [3].

2.3 Ultrasound Imaging Systems

2.3.1 Ultrasound Generation and Detection

The most crucial component inside an ultrasound transducer is the piezoelectric crystal (see Fig. 2.4), which is responsible for the generation and detection of ultrasound waves. A typical example of the material used to construct this crystal is lead zirconate titanate (PZT). Due to the natural properties of the piezoelectric material, applied electric pulses will cause the crystal to shrink and expand, thereby creating a longitudinal sound wave with alternating compressions in the surrounding medium. Conversely, as a receiver, the crystal will generate electric signals in response to the mechanical strain caused by reflected echo pulses. These electric signals are then conducted via the attached electrodes to an RF amplification circuit.

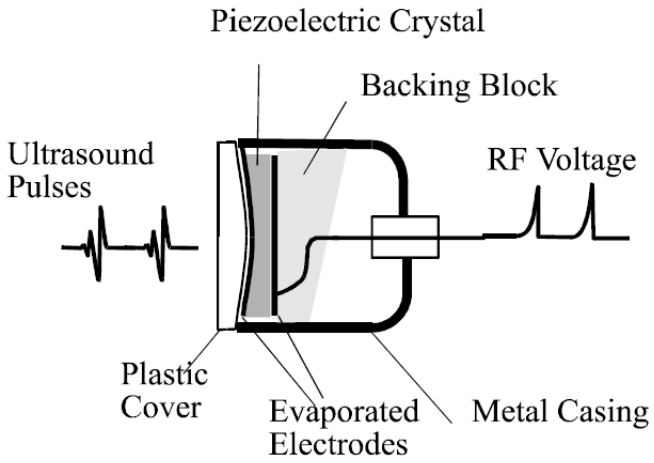

Figure 2.4: A single element transducer is shown. As a transmitter, the piezoelectric crystal is excited by the electric voltages pulses applied to the electrodes at RF frequencies. As a receiver, the RF voltages are produced in response to the returning ultrasound pulses [3].

2.3.2 A Mode Scans

A-Mode scans or amplitude scans is the most basic form of ultrasound imaging, illustrated in Fig. 2.1. In this scenario, the amplitude profile produced will represent the level of acoustic impedance mismatch and the position of the reflecting boundaries along the path of the transmitted pulse.

2.3.3 B Mode Scans

B-Mode scans or brightness mode is the most common form of ultrasound imaging, this is depicted in Fig. 2.5. It consists of multiple A-mode scans taken at different positions or angles and then combined to produce a 2D image in the field of view. In a B-scan image, the x, y coordinates encapsulates the positional information, while the intensity or brightness values represent the strength of the ultrasound reflections.

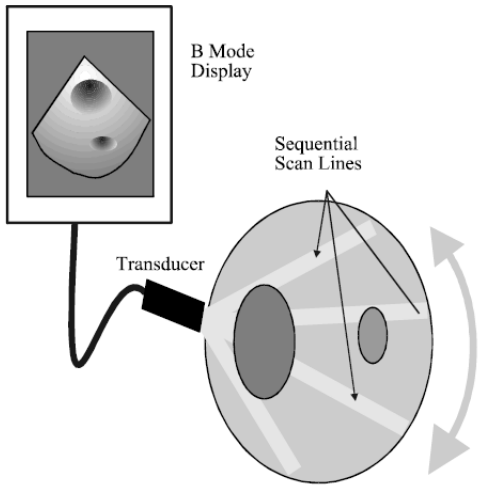

Figure 2.5: A standard B-Mode imaging scan setup [3]. Multiple A-scans are taken and combined to produce the 2D ultrasound image.

2.3.4 Ultrasound System Overview

Having focused on the production and detection of transmitted and reflected ultrasound pulses, now a brief overview of a general ultrasound imaging system is provided. As illustrated in Fig.2.6, and explained before, the transducer is responsible for emitting the excitation ultrasound pulses. Then, reflected echo pulses from inside the body are received by the transducer, the latter generates an electric signal in response and the signal is subsequently sent to a conventional RF circuit. The RF circuit amplifies the signal, and the amplified signal is then passed as an input into the swept gain amplifier, which attempts to correct for the attenuation effects inside the patient's body. This is an important step because attenuation in ultrasound is constant for certain tissues. Weaker echoes from the same tissue must be amplified accordingly to produce a more uniform appearance. Finally, the attenuation compensated signal is passed through a detector demodulator, which ex-

tracts the lower frequency amplitude envelope of the echoes. The resulting amplitude profile is displayed on the monitor of the imaging system.

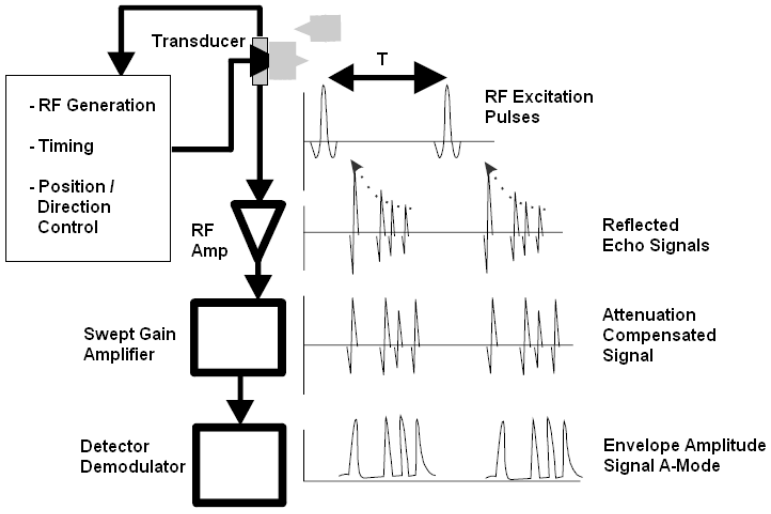

Figure 2.6: A general block diagram of a full ultrasound system is shown [3]. The above diagram illustrates the full signal processing scheme for a single line A-scan. 2D B-scans are composed of series of A-scan processes depicted above. Note that the successive RF excitation pulses are emitted with a delay of period T. During the delay period the echoes must be received and processed before emitting another excitation pulse such that there is minimal interference between the reflections of the previous pulse and the current excitation pulse.

2.3.5 Minimally Invasive Ultrasound Probes and TRUS Imaging

Although the use of transducers outside the body has provided us with a much greater insight into many anatomical structures near the skin's surface, there still exists many small but important organs that lie relatively deep inside the body. These important structures include many of the organs in the reproductive and digestive system. Since these organs can be several centimetres deep into the body, the ultrasound excitation pulses must travel through soft tissues and perhaps bone structure to reach these targets. Along the way, the pulses will be heavily attenuated due to absorption and scattering effects mentioned before, thereby producing very poor images. Fortunately, the advancements in construction of ultrasound transducers have led to significant miniaturization of the latter. Thus, clinicians can take advantage of natural body orifices to carry out B-scans from inside the body. This has led to significant improvements in image quality in terms of both

spatial resolution and contrast.

A prevalent application of ultrasound imaging performed inside the body is in transrectal ultrasound (TRUS), which is used in the diagnosis of prostates (see Fig.2.7 [1]). The prostate is a walnut sized gland located beneath the bladder and in front of the rectum in men. Due to the position of this gland, improved image quality can be achieved by performing TRUS imaging. During the latter procedure, a small lubricated transducer is inserted into the rectum, and subsequent high frequency pulses are emitted to produce a B-scan image of the prostate and its surrounding tissues as before [1].

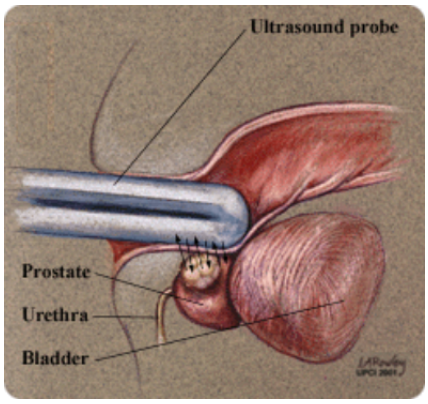

Figure 2.7: A TRUS imaging setup [1]

Chapter 3

Literature Review

3.1 Overview of Segmentation Techniques

Since the size and shape of the prostate are important quantities that are used by clinicians to identify abnormal developments within the gland, an automatic or semi-automatic segmentation technique for prostates in TRUS images is regarded as a pivotal component of the TRUS-based diagnosis of prostate cancer. The desired outcome of the segmentation procedure is to accurately partition the given TRUS image into two regions corresponding to the actual prostate and the background image. In order to solve this problem, a number of existing strategies have been proposed in literature. Given the similar properties of the existing techniques, they can be generally grouped into three main categories. These include edge based segmentation algorithms, classification based algorithms, and the more recent shape prior based segmentation algorithms, which is an extension to previous strategies.

3.2 Edge Based Segmentation Algorithms

3.2.1 Edge Guided Manual Boundary Delineation

One of the earlier algorithms proposed to enhance the segmentation procedure in TRUS images is described in the work of [4]. This algorithm still requires heavy user interaction at the final stage of the segmentation procedure, but it appears to achieve the original objective of the authors, which is to enhance segmentation accuracy via reducing intra- and inter-observer variability. The algorithm consists of 4 main stages, the first stage involves contrast enhancement and speckle reduction, the second stage performs an edge-preserving

smoothing procedure, this is followed by knowledge based false edge removal and subsequently, the detected edges are used as a visual guide to the final manual segmentation. The overall flow of the authors approach is shown in Fig.3.1.

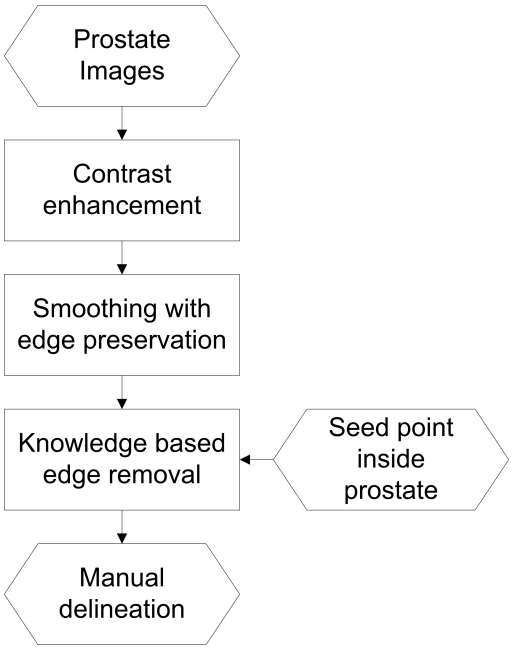

Figure 3.1: Edge Guided Delineation Algorithm

Contrast Enhancement and Speckle Reduction

The authors of [4] use a "sticks" filtering algorithm, which can reduces the number of false edges by modeling edges in ultrasound images as a line process [5]. The idea is to project the image onto a family of line segments or sticks, differing in orientation, but centered at and passing through the pixel of interest (POI) in the original image. Then, the highest value obtained under any projection is then plotted as the pixel intensity of the enhanced image at the same pixel position. Examples of a family of sticks in a 3×3 neighbourhood is depicted in Fig.3.2.

As the stick length increases, the speckles in the image are blurred while the prostate boundaries becomes more defined. However, if the stick length is chosen to be longer than the actual prostate edges, the SNR of the filtered image actually degrades because more noisy pixels would be included in the stick templates. In the authors' algorithm an intermediate stick length of 15 is chosen as a compromise between the two trade-offs.

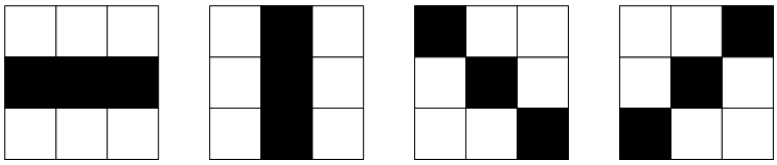

Figure 3.2: Four possible angular orientations of sticks with length of three

Smoothing with Edge Preservation

After the contrast enhancement stage is complete, the authors of [4] use an additional noise suppression technique similar to anisotropic diffusion filtering [6]. This technique attempts to model the original image $I(x, y)$ with a piecewise smooth function $S(x, y)$, which is modelled as a weak elastic membrane with breaks and creases where discontinuities exist in $I(x, y)$. Then, an energy cost function is associated with the membrane, which when minimized leads to an edge preserved smoothed image. The energy cost function is a sum of three terms, which the authors call a data mismatch penalty (D), a roughness penalty (R), and a discontinuity penalty (C). They are defined as [4]:

$$D = \sum_{x,y} \big(S(x, y) - I(x, y)\big)^2 \tag{3.1}$$

$$R = \beta^2 \sum_{x,y} \big((\nabla S(x, y))^2 (1 - l(x, y)\big)^2 \tag{3.2}$$

$$C = \alpha \sum_{x,y} l(x, y) \tag{3.3}$$

where ∇ represents the spatial gradient operator, β^2 measures elasticity of the membrane, α is the discontinuity penalty, and $l(x, y)$ is defined as [4]:

$$l(x, y) = \begin{cases} 1, & \|\nabla S(x, y)\| \geq \sqrt{\alpha/\beta} \\ 0, & \text{otherwise} \end{cases} \tag{3.4}$$

Finally, the optimal smooth image $S^*(x, y)$ is found by minimizing E as follows [4]:

$$\begin{aligned} S^*(x, y) &= \arg\min_{S,l} \{D + R + C\} \\ &= \min_{S} \big(D + \min_{l} \sum_{x,y} (R + C)\big) \end{aligned} \tag{3.5}$$

Note, that in the paper values of α and β are empirically defined to reflect optimal smoothing as evaluated by an expert observer.

Knowledge Based Edge Removal

Once a smooth approximation $S(x, y)$ is obtained from the original image $I(x, y)$, the classic Canny edge detector algorithm [7] is applied to obtain an edge map. However, due to the underlying difficulties with TRUS imaging, many false/spurious edges are present. Therefore, the authors post-process the resulting edge map by using *a priori* knowledge regarding TRUS images of the prostate. A manual seed point input is required to be performed by the user to indicate the middle position of the prostate, and then false/spurious edges are removed according to 2 conditions. First condition assumes edges of the prostate are generally formed as dark-to-light transitions from inside of the prostate to outside. Thus, a radial search extending out from the seed point removes all edges that are formed from light-to-dark transitions. Second condition removes edges that are considered to be outliers if they are sufficiently far from the manual seed point in the radial direction.

Finally, the post-processed images are used as visual guides for an expert to delineate the prostate manually. The proposed method of [4] does effectively reduce the intra- and inter-observer variability, when it comes to manual segmentations. However, the main drawback of this algorithm lies in the fact that it still requires a very time-consuming manual segmentation step at the end of the procedure.

3.2.2 Edge Guided Discrete Dynamic Contour

Another edge guided algorithm described in [8] is semi-automatic in nature. It requires the user to define 4 initial seed points that lie on the boundary of the prostate. Subsequently, an edge guided discrete dynamic contour (DDC) algorithm [9] is performed to drive the contour towards strong local edges of the prostate. Moreover, the authors have provided the option to manually manipulate the final contour and re-perform the segmentation to achieve a potentially more accurate result.

Discrete Dynamic Contour

In the authors approach a rough estimation of the prostate boundary is calculated based on the 4 initial seed points. Then, the final contour of the prostate is refined with the edge-driven DDC.

The DDC consists of a sequence of points that are connected by straight-line segments. The points represent a set of vertices, which will deform to fit strong local edges in the

image. The vertices and line segments of the DDC are illustrated in Fig.3.3 (a). As illustrated, the unit edge vector $\hat{d}_i$ points in from vertex V_i to V_{i+1}. Moreover, a local unit tangential vector $\hat{t}_i$ is defined for each vertex as follows [8],

$$\hat{t}_i = \frac{\hat{d}_i + \hat{d}_{i+1}}{\|\hat{d}_i + \hat{d}_{i+1}\|}, \tag{3.6}$$

where $\|\cdot\|$ denotes the operator of L^2 norm. A local radial unit vector $\hat{r}_i$ can also be defined from $\hat{t}_i$ by rotating $\hat{t}_i$ 90 degrees. Thus, $\hat{r}_i$ and $\hat{t}_i$ define a local coordinate system for vertex i.

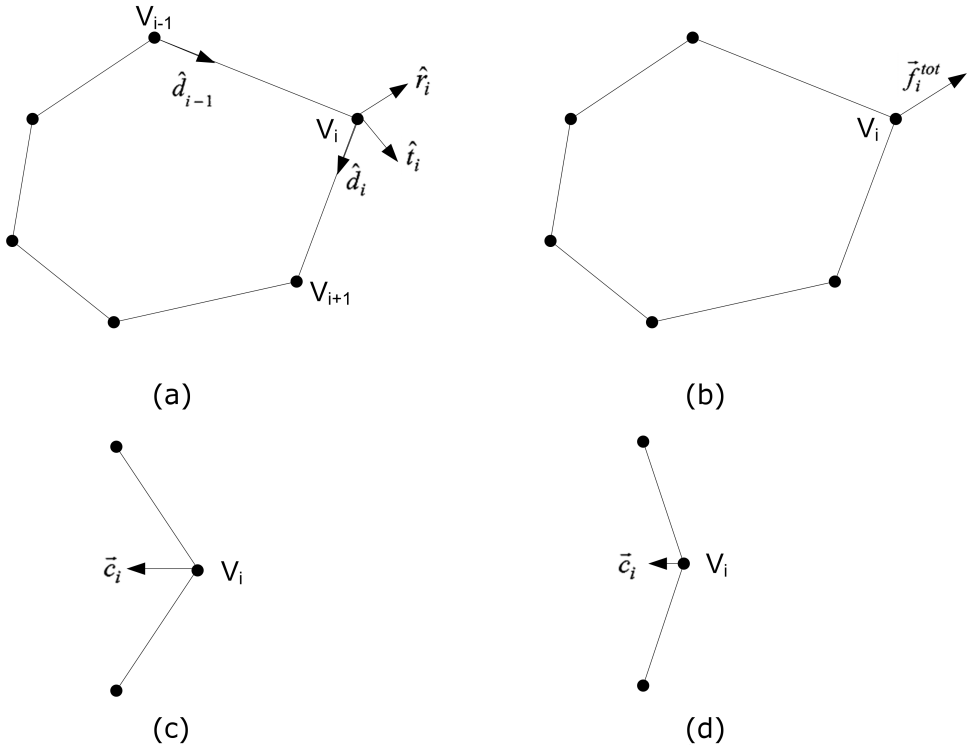

Figure 3.3: (a) Configuration of the DDC. (b) Force vector driving a DDC vertex. (c) Curvature vector $\mathbf{c}_i$ measures change between unit edge vectors along the vertices. (d) small curvature vector.

The DDC evolves according to force vectors acting on each of its vertices as shown in Fig.3.3 (b). The single force vector $\mathbf{f}_i^{tot}$ is actually a weighted combination of internal ($\mathbf{f}_i^{int}$), image ($\mathbf{f}_i^{img}$) and damping ($\mathbf{f}_i^{d}$) forces [8],

$$\mathbf{f}_i^{tot} = w_i^{int}\mathbf{f}_i^{int} + w_i^{img}\mathbf{f}_i^{img} + \mathbf{f}_i^{d}, \tag{3.7}$$

where w_i^{int} and w_i^{img} are relative weights for the internal and image forces respectively. In the TRUS prostate segmentation problem, the image forces act to drive the DDC vertices towards strong local edges. Thus, the image force can be defined at each pixel in the image as [8],

$$E(x_p, y_p) = \|\nabla(G_\sigma * I(x_p, y_p))\|, \tag{3.8}$$

$$\mathbf{f}_i^{img}(x_p, y_p) = \frac{2\nabla E(x_p, y_p)}{\max \|\nabla E(x_p, y_p)\|}, \tag{3.9}$$

where E represents the energy associated with a pixel having coordinates (x_p, y_p), G_σ is a Gaussian smoothing kernel with a standard deviation of σ, and I is the original image. The $*$ denotes the operator of convolution. The idea behind the energy E is that it is locally maximized at the pixels that lie on strong edges, and the definition of the image force $\mathbf{f}_i^{img}$ drives the DDC towards these local edges.

Note that the image forces have a limited range effect around the edges, the vertices of the DDC will not be attracted to the edges if they are outside of this range. The definition of σ controls this range, the higher the σ the larger the spatial range. However, if σ becomes too large, it may result in poor localization of the desired edge. The authors of [8] selects a value of $\sigma = 5$ as a compromise.

The image forces acting on the DDC are obtained via sampling of the force field in (3.9) at the vertex positions (x_i, y_i). Moreover only the radial component of the image force is considered. Thus, the force vector acting on vertex i is calculated as follows [8],

$$\mathbf{f}_i^{img} = (\mathbf{f}^{img}(x_i, y_i) \cdot \hat{r}_i)\hat{r}_i, \tag{3.10}$$

where "$\cdot$" denotes the vector dot product.

The internal force acts to minimize local curvature at each vertex, keeping the DDC smooth in the presence of noise. It is defined as [8],

$$\mathbf{f}_i^{int} = (\mathbf{c}_i \cdot \hat{r}_i - \frac{1}{2}(\mathbf{c}_{i-1} \cdot \hat{r}_{i-1} + \mathbf{c}_{i+1} \cdot \hat{r}_{i+1}))\hat{r}_i, \tag{3.11}$$

$$\mathbf{c}_i = \hat{d}_i - \hat{d}_{i-1}, \tag{3.12}$$

where $\hat{c}_i$ is the curvature at vertex i.

The last component of the total force vector $\mathbf{f}_i^{tot}$ is the damping force $\mathbf{f}_i^{d}$, which opposes the motion of evolution of each vertex to keep the DDC stable and prevent oscillations. It is defined as [8],

$$\mathbf{f}_i^{d} = w_i^d \mathbf{v}_i, \tag{3.13}$$

where w_i^d is a negative weighting factor and $\mathbf{v}_i$ is the velocity at vertex i during contour evolution. The authors have experimentally set the weights to be $w_i^{img} = 1.0$, $w_i^{int} = 0.3$, and $w_i^d = -0.5$. Thus, the main driving force behind the DDC evolution is the image force, which favors deformation towards edges rather than smoothing due to internal forces.

Finally, the position of each vertex of the DDC at time $t + \Delta t$ is calculated based on the following difference equations [8],

$$\mathbf{p}_i(t + \Delta t) = \mathbf{p}_i(t) + \mathbf{v}_i(t)\Delta t, \tag{3.14a}$$

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \mathbf{a}_i(t)\Delta t, \tag{3.14b}$$

$$\mathbf{a}_i(t + \Delta t) = \frac{1}{m_i}\mathbf{f}_i(t + \Delta t), \tag{3.14c}$$

where $\mathbf{p}_i = (x_i, y_i)$ is the position of vertex i, $\mathbf{a}_i$ is its acceleration, m_i is its mass, and Δt is an artificial time step. The mass of each vertex and the artificial time step are set to unity. Initial vertex positions $\mathbf{p}_i$ are defined by the user input, while the initial velocity and acceleration are set to 0. Subsequently, the position, velocity, and acceleration of each vertex is updated iteratively according to the above equations. The segmentation procedure is complete when velocity and acceleration of each vertex approaches zero.

It is important to note that because TRUS images of the prostate suffer from speckle, shadowing, and refraction artifacts, sections of the prostate boundary will be poorly defined. Therefore the choice of the initial DDC contour is highly crucial.

Initialization Procedure

The authors of [8] has devised an initialization procedure where the user needs to input four seed points for the initial DDC as shown in Fig.3.4.

A local $x - y$ coordinate system is defined from the user selected seed points. Then, suppose the DDC is represented in parametric from as [8],

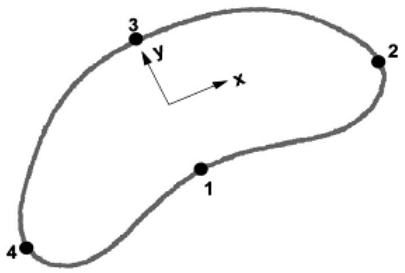

Figure 3.4: Location of 4 user selected seed points labeled (1)-(4) and the local x-y coordinate system

$$\mathbf{p}(s) = \begin{Bmatrix} x(s) \\ y(s) \end{Bmatrix}, \tag{3.15}$$

where $\mathbf{p} = (x, y)$ is the position of a vertex on the DDC, and s is a parameter which varies from 0 to 1 along each curve segment. The four seed points define four segments along the prostate boundary: 1-2, 2-3, 3-4, and 4-1. Then the authors use a cubic interpolation function to define additional points along the four segments to fully characterize the prostate boundary. They are defined as follows [8],

$$x(s) = a_3 s^3 + a_2 s^2 + a_1 s + a_0, \tag{3.16a}$$

$$y(s) = b_3 s^3 + b_2 s^2 + b_1 s + b_0, \tag{3.16b}$$

where s is a parameter which varies from 0 at a starting point of a segment to 1 at the ending point, and a_i and b_i are coefficients.

Once the coefficients are determined, vertices of the DDC are uniformly interpolated within each segment at $\Delta s = 0.1$ units. Hence, the initial contour is defined.

Deformation and Editing

After initialization, the DDC is evolved according to the force field in (3.7) that was described previously. Ladak et al. ackowledge the fact in some difficult cases, the initialization routine may not place all of the vertices in close vicinity of the actual prostate boundary. Thus, an optional editing stage is implemented such that after the segmentation process is complete, the user may drag some of the vertices closer to the observed boundary and re-perform the segmentation process again in order to achieve a better result.

3.2.3 Combined Region and Edge Based Segmentation

Due to the aforementioned difficulties with US images, using edge information alone cannot guarantee accurate segmentation results. Therefore, Yu et al. propose a region based active contour method in [10], which is followed by a secondary edge-guided optimization process. Since, the initial region based optimization process is less sensitive to noise and spurious edges in the US images, it is able to achieve a rough segmentation that is in the vicinity of the true prostate boundary, while the subsequent edge guided process further improves the segmentation accuracy.

Manual Initialization

The proposed algorithm is first initialized by manually selecting two points in the input US image, which are located just outside of the top left and bottom right corners of the prostate. Then an initial contour, an ellipse, is inscribed between the two seed points as shown in Fig. 3.5

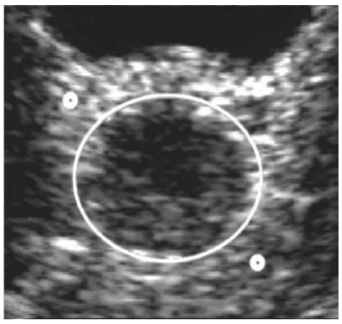

Figure 3.5: Top left and bottom right points in white are manually defined by the user, and subsequently the ellipse bounded by the seed points is defined as the initial contour.

Area Weighted Binary Flow

As mentioned before, since the prostate boundary often appears noisy discontinuous in US images, the authors propose to use a region-based active contour. This approach termed Area-weighted Mean-difference Binary Flow model (AMBF) does not depend on local image edge features, and it is derived by minimizing the following cost functional [10]

$$E(C) = -\frac{1}{2}\frac{A_u A_v}{A}(u-v)^2, \tag{3.17}$$

where the energy E is dependent on the closed curve C. Moreover, A_u and A_v are the areas of the region inside and outside of the curve C respectively. A defines the area of the image, defined to be $A = A_u + A_v$, while u and v are mean intensities defined as follows [10],

$$u = \frac{1}{A_u} \int_{insideC} G_\sigma * I(x, y) dx dy, \tag{3.18a}$$

$$v = \frac{1}{A_v} \int_{outsideC} G_\sigma * I(x, y) dx dy, \tag{3.18b}$$

where $\tilde{I} = G_\sigma * I$ represents the Gaussian-filtered version of the original image, and G_σ represents a Gaussian kernel with a standard deviation of σ. The σ is chosen empirically to be 4, which allows for sufficient noise suppression without over-smoothing the prostate edges.

Contour Representation

The authors of [10] use the level-set function to implicitly represent the contour C mentioned above. First define the image domain to be Ω, while Ω_{in}, Ω_{out}, and $\partial\Omega_{in}$ are defined to be inside C, outside C, and C respectively. Then a corresponding level-set function can be used to conveniently describes the different regions of the image domain as follows,

$$\varphi(x, y) = \begin{cases} < 0, & (x, y) \in \Omega_{in} \\ > 0, & (x, y) \in \Omega_{out} \\ = 0, & (x, y) \in \partial\Omega_{in}. \end{cases} \tag{3.19}$$

Primary Region Based Optimization

Subsequently, finding the optimal C (i.e. minimize the cost functional $E(C)$) via the AMBF model and the level-set representation amounts to finding C, which is the zero-level set of the function $\varphi(x, y; t \to \infty)$. This steady state solution is found by solving the following partial differential equation,

$$\frac{\partial \varphi}{\partial t} = \delta(\varphi) \frac{1}{A} (u - v) \Big[A_v (\tilde{I} - u) + A_u (\tilde{I} - v) - \frac{1}{2} (A_u - A_v)(u - v) \Big], \tag{3.20a}$$

$$\varphi(x, y; 0) = \varphi_0(x, y) \text{ in } \Omega, \tag{3.20b}$$

$$\delta(\varphi)\frac{\partial\varphi}{\partial\vec{n}} = 0 \ \text{ on } \partial\Omega, \tag{3.20c}$$

where Ω and $\partial\Omega$ are the image domain and its boundary respectively, $\vec{n}$ denotes the exterior normal to $\partial\Omega$, $\partial\varphi/\partial\vec{n}$ denotes the normal derivative of φ on $\partial\Omega$. Moreover, A_u, A_v, u and v can be defined in terms of the level-set function as,

$$A_u = \int_\Omega H(\varphi(x,y))dxdy, \tag{3.21a}$$

$$A_v = \int_\Omega [1 - H(\varphi(x,y))]dxdy, \tag{3.21b}$$

$$u = \frac{1}{A_u}\int_\Omega \tilde{I}(x,y)H(\varphi(x,y))dxdy, \tag{3.21c}$$

$$v = \frac{1}{A_v}\int_\Omega \tilde{I}(x,y)[1 - H(\varphi(x,y))]dxdy, \tag{3.21d}$$

where H denotes the standard heaviside function as $H(x) = (x)_+$.

The partial differential equation in 3.20 is solved numerically by discretizing both time and space coordinates and using the following update equation system,

$$\begin{aligned}\frac{\varphi_{i,j}^{n+1} - \varphi_{i,j}^{n}}{\Delta t} &= \delta(\varphi_{i,j}^{n})\frac{1}{A}[u(\varphi^n) - v(\varphi^n)]\{A_v(\varphi^n)[\tilde{I}_{i,j} - u(\varphi^n)] + \\ &\quad + A_u(\varphi^n)[\tilde{I} - v(\varphi^n)] - (1/2)\times \\ &\quad \times[A_u(\varphi^n) - A_v(\varphi^n)][u(\varphi^n) - v(\varphi^n)]\}\end{aligned} \tag{3.22}$$

and the initial level set function is set to

$$\varphi_{i,j}^0 = \sqrt{(i - x_c)^2 + a^2/b^2(j - y_c)^2} - a \tag{3.23}$$

where a, b, x_c, y_c are the two half (major/minor) axes and the center coordinates of the initial elliptical contour respectively. Also, i, j are discretized spatial coordinates and Δt is an artificial time step set to 0.1.

The authors of [10] do not allow the AMBF segmentation to proceed until convergence, instead they stop the contour evolution when the area of the region inside the contour differs from that of the initial ellipse by 20%. This is determined empirically and allows the contour to move sufficiently close to the vicinity of the true prostate boundary. From the resulting level-set function, the primary optimal contour is obtained by $C = \{(x,y) \in \Omega \mid \varphi(x,y) = 0\}$.

Secondary Edge Guided Optimization

After the initial segmentation contour is obtained from the primary optimization algorithm, the result is further optimized via a local edge guided segmentation. The latter is performed on a speckle reducing anisotropic diffusion (SRAD) [11] filtered image, which enhances the US image by reducing speckle and improving image contrast. Moreover, the authors use an edge detection algorithm they developed, instantaneous coefficient of variation (ICOV), to detect local image edge features.

To perform the secondary optimization, the authors propose to formulate the problem via an active contour model in the polar coordinate system as opposed to the traditional active contour model in the Cartesian coordinate system. The relationship between the two are shown in Fig. 3.6. $\vec{R}_c$ denotes the position of point o', which is the center of the polar coordinate system in the Cartesian system. $\vec{R}$ and $\vec{\rho}$ represents the position vectors to the point P in the Cartesian and polar coordinate systems respectively.

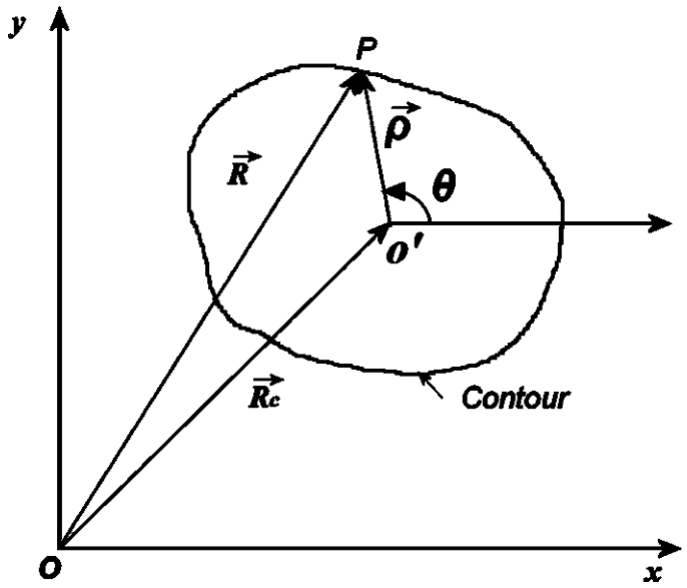

Figure 3.6: Geometry of the polar coordinate model

Next, the authors propose to find a radial segmenting curve represented by $\rho = \rho(\theta), \theta \in [0, 2\pi]$, which propagates the active contour towards local strong edges while imposing certain elasticity and rigidity restrictions. This is achieved via minimizing the energy functional defined below [10]:

$$E(\rho) = \int_0^{2\pi} \{(\alpha(\theta)/2)[\rho'(\theta)]^2 + (\beta(\theta)/2)[\rho''(\theta)]^2 + E_{ext}(\vec{R}_c + \rho(\theta)\hat{e}_\rho)\} d\theta, \tag{3.24}$$

where α and β are user defined constants restricting the elasticity and rigidity of the contour. $\hat{e}_\rho = \hat{e}_x \cos\theta + \hat{e}_y \sin\theta$ where $\hat{e}_x$ and $\hat{e}_y$ are unit vectors in the x and y directions respectively. E_{ext} is an external function that is minimized, near strong local edges. It can be for example defined as the ICOV edge detector [11],

$$E_{ext}(x,y) = -\frac{|(1/2)\|\nabla\tilde{I}_s\|^2 - (1/16)(\nabla^2\tilde{I}_s)^2|}{[\tilde{I}_s + (1/4)\nabla^2\tilde{I}_s]^2}, \tag{3.25}$$

where $\tilde{I}_s$ is the SRAD filtered version of the original intensity image I_d. The ICOV edge detector is more robust as compared with the standard Gaussian filtering and gradient edge detectors in the case of ultrasound images. Thus, the final segmentation is achieved upon completion of the secondary edge driven segmentation procedure.

3.2.4 Summary of Edge Based Methods

Although classic edge-based segmentation algorithms perform very well in situations where the structural boundaries of the object(s) of interest are well defined, their effectiveness is greatly hampered in the case of ultrasound images. The challenges lie in the many imaging artifacts that are present in TRUS images of the prostate. These include, but are not limited to, signal dependent speckle noise, and shadowing regions in the vicinity of prostate boundaries. As such, the detected local edges are known to be unreliable features.

Therefore, a number of different strategies have been implemented to address the shortcomings of edge-based methods. As discussed above, the approach in [4] employs the use of pre- and post-processing techniques to reduce the number of erroneous edges detected by the standard Canny edge detector. Subsequently, the resulting edge map is superimposed over the original images as a visual guide to subsequent manual delineation. Although this is useful for reducing both intra- and inter-observer variances of prostate delineation, the method still requires heavy user interaction. Alternatively, the method of [8] is a semi-automatic algorithm, which uses four user defined seed points to generate an initial segmenting contour that is subsequently driven towards strong local edges. However, because of the unreliability of detected local edges, a follow-up manual editing method is implemented to improve the accuracy of the delineation. In order to improve robustness, a two step segmentation procedure was proposed in [10]. At the first stage of this method, a coarse segmentation is obtained based on texture-related features of the TRUS images, followed by forcing the segmentation boundary to converge to local image edges during the second stage. Although the latter method provides a considerable improvement, it (as well as the previously mentioned methods) still remain prone to the errors caused by the existence of spurious (noise-induced) edges and shadowing artifacts.

3.3 Classification Based Segmentation Algorithms

3.3.1 Segmentation via the Use of Neural Networks

One of the classic machine learning techniques in the form of neural networks has also been applied to the problem of segmentation. In particular, the most common type of neural network, being the layered feedforward network, is studied in the work of Prater and Richard [12]. In their study, different neural network architectures and input feature patterns were evaluated, and it was demonstrated that the proper choice of input patterns and architecture can significantly influence the accuracy of the segmentation.

Operation of Feedforward Neural Networks

A layered feedforward neural network generally consists of one input and one output layer and one or more hidden layers. A simple three layer network is shown in Fig. 3.7 below.

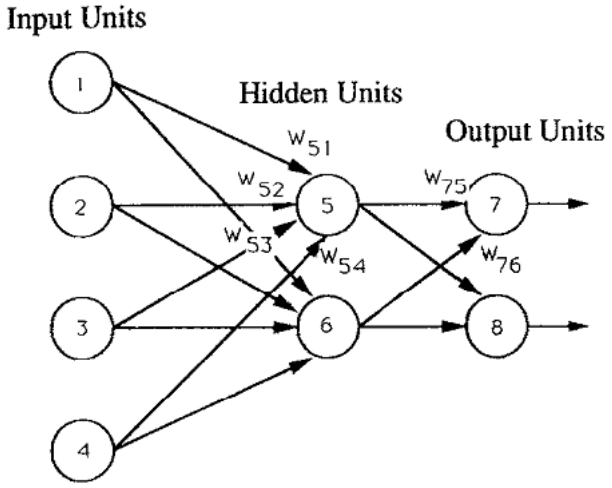

Figure 3.7: A simple three layered feedforward neural network [12]

Features or training patterns are provided to the input layer, which are subsequently passed along to the hidden layers and eventually to the output layer. Each unit in the hidden and output layer is defined as a neuron, which has as its input a weighted sum of values from the previous layer (input or hidden) and is passed through a sigmoid function to obtain a particular output. This is demonstrated in Fig. 3.8 below. Specifically, $x_1 \cdots x_N$ are the inputs from a previous layer, $w_{i1} \cdots w_{iN}$ are corresponding input weights, and S_i being the output of the neuron unit.

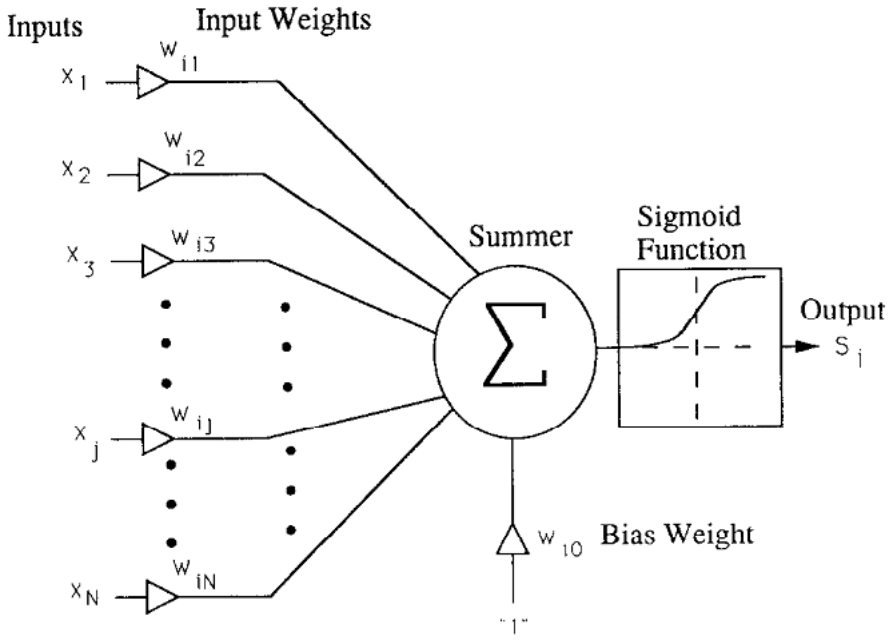

Figure 3.8: Structure of a neuron unit [12]

In particular the output of the neuron S_i can be defined as [12]

$$\mathrm{S_i} = f(\sum_{j=0}^{N} w_{ij} x_j) \tag{3.26}$$

where,

$$f(x) = \frac{1}{1 + e^{-x}} \tag{3.27}$$

Due to the combination of multiple neuron units across many layers, the network implements a nonlinear classifier, that is dependent on its architecture as well as the distribution of the weights within the network. While the network architecture is predefined by the designer, the optimal set of weights are initially unknown and randomly assigned. Subsequently, a very large set of training patterns with expected output values are supplied to the neural network. The empirical outputs from the training patterns are compared with the expected values, and the weights of the network are re-adjusted to reduce the amount of error. The previous process is performed over many iterations to obtain a set of optimal network weights that correctly classify all or most of the training patterns. In particular, the authors [12] used the method of conjugate gradient optimization [13] to obtain an optimal set of weights.

Neural Network Architectures and Results

The evaluation framework of the authors in [12] is as follows: a 40 image set of the prostate of a 29 year old man is obtained using a TRUS unit. The center image (slice 21) of the prostate is manually segmented by an expert sonographer, and the resulting pair of the acquired TRUS image and the ground truth image is used to train all the subsequent neural networks described in this section. The acquired and ground truth images are shown in Fig. 3.9 subplot A and B respectively below. Once the different networks are trained via the same training images, their performance is assessed on two TRUS image slices from the same 40 image set and adjacent to the center image. Specifically, one adjacent image closer to the apex of the prostate (slice 22) and one adjacent image closer to the prostate base (slice 20) are used for performance evaluation; the testing images are shown in Fig. 3.10 below.

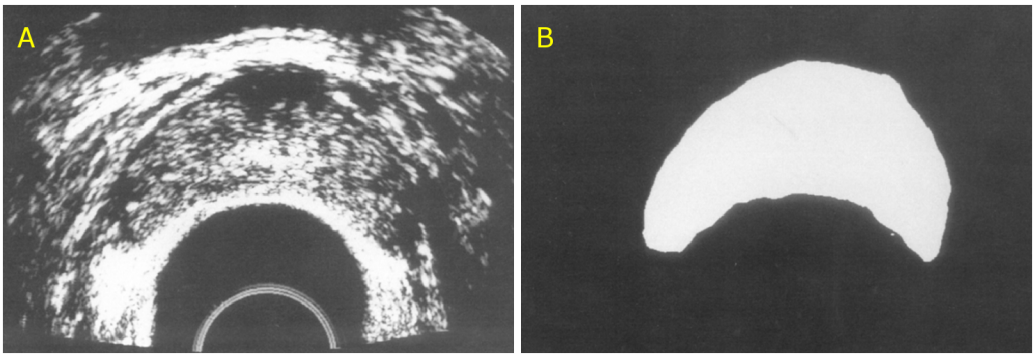

Figure 3.9: (Subplot A) Acquired center TRUS image (slice 21) of a prostate. (Subplot B) Ground truth manual segmentation of the prostate region. [12]

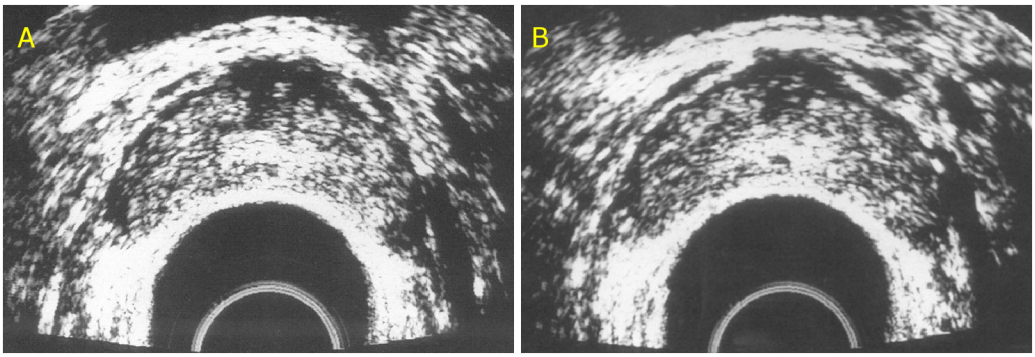

Figure 3.10: (Subplot A) Acquired TRUS image closer to the apex (slice 22). (Subplot B) Acquired TRUS image closer to the base (slice 20). [12]

A number of neural network architectures were used to segment the testing images, and their accuracy is summarized in the Table 3.1 [12] below.

Table 3.1: Segmentation Results of Various Neural Networks

Network Architecture	Number of Training Patterns	Classification Accuracy (%)		
		Training	Testing (apex)	Testing (base)
(1) 5-6-2	1334	90.0	87.3	86.9
(2) 132-30-2	3000	98.8	92.8	91.1
(3) 132-30-2	15908	99.6	92.9	90.9
(4) 61-10-2	7500	99.3	95.3	94.7

The first neural network that was evaluated by the authors had a 5-6-2 structure, which represented 5 input, 6 hidden, and 2 complementary output units. Only one of the output units would produce an "on" value indicating whether the point of interest (POI) is inside or outside of the prostate. In terms of the input units, for every POI, the extracted input patterns include the intensity mean value of a 3×3 window centred at the pixel and the maximum of 3×3 window intensity means above, below, to the left and to the right of the POI. Moreover, the number of training patterns used include 1334 pixel sites, which consists of 0.8% of the entire training image. The idea behind these input patterns lies in the fact that there exists bright regions just outside of the prostate boundary, which is attempted to be captured by the maxima of local window means in four outward directions from the POI. Although this network is relatively simplistic and utilizes very few training patterns, the classification accuracy is a respectable 87.3% and 86.9% for the two adjacent TRUS images.

The second neural network expands to the architecture of 132-30-2, which uses a significantly larger number of input patterns as well as hidden units. In particular, the network uses 132 inputs consisting of 1) the intensity mean of a 3×3 window centered at the POI; 2) 91 local window means of size 3×9 and 9×3 in the same row and column of the POI respectively; and 3) 40 intensity means of 3×3 windows forming a rectangular structure centered around the POI. The choice of input patterns is designed to capture the bright "halo" around the prostate as well as some local intensity features. The more descriptive feature set, increase in number of training patterns, and the larger network resulted in a better classification accuracy for the two testing images; the percentage of pixels correctly classified for image slice 22 and slice 20 were 92.8% and 91.1% respectively. The third network used the same 132-30-2 architecture as the second network; however, a much larger amount of training data (15908 patterns) was used to retrain the same network. The authors [12] noted that the overtraining of the network did not result in improvement in

overall performance. In fact, overtraining actually degraded the network performance for the testing image that was closer to the prostate base.

A fourth network was constructed using the 61-10-2 structure. This network reduced the number of hidden units because the authors found many units from the previous network architecture that were essentially inactive. The latter was discovered by eliminating specific hidden units in the network and evaluating the resulting difference. Moreover, although the number of input units were reduced, they were redesigned to be more descriptive as compared with the previous network. Specifically, the 61 inputs consist of: 1) the mean of a 3×3 window centred at the POI; 2) 31 means of 3×15 windows in the same row as the POI; 3) 21 means of 15×3 windows from the same column as the POI; 4) 4 means of 3×3 windows 15 pixels from the POI in the same row and column; and 5) 4 means of 3×3 windows 11 pixels diagonally away from the POI. Therefore, other than the 3×3 mean centred at the POI, the other small local 3×3 windows forms a circular ring around the specific POI. It was shown that the achieved classification accuracy for the two adjacent images were 95.3% and 94.7% respectively.

Finally, the visual segmentation results for each of the four networks described above are provided in Fig. 3.11 below. From the visual results, as well as the results from Table 3.1, it can be seen that the architecture of the neural network and the choice of input patterns can significantly influence the segmentation performance. The design process can be viewed as follows: the first simple 5-6-2 network captured the most significant visual cues used by expert sonographers, which are the bright halo regions around the prostate. The second 132-30-2 network improved upon the previous network by using more local texture information in addition to capturing the bright regions. Finally the 61-10-2 network was obtained via optimizing the second network; specifically, the authors' final network design removed inactive hidden units from the previous design and used more relevant local texture information as input patterns. Although relatively accurate segmentation results were achieved using the optimal network, significant segmentation artifacts can still be seen near the true prostate boundaries.

3.3.2 Unsupervised Texture Based Prostate Segmentation

A different classification approach was presented by Richard and Keen [14] towards the problem of TRUS segmentation of the prostate. In their work, a texture-based method was presented based on feature extraction techniques originally proposed by Laws [15]. The extracted texture measures from the TRUS images are subsequently used as input patterns for the chosen pixel classification scheme, based on the unsupervised learning techniques presented by Hsiao and Sawchuk [16]. With the combination of the above mentioned techniques, the authors [14] showed the feasibility of applying automatic texture based classification algorithms for ultrasound images of the prostate.

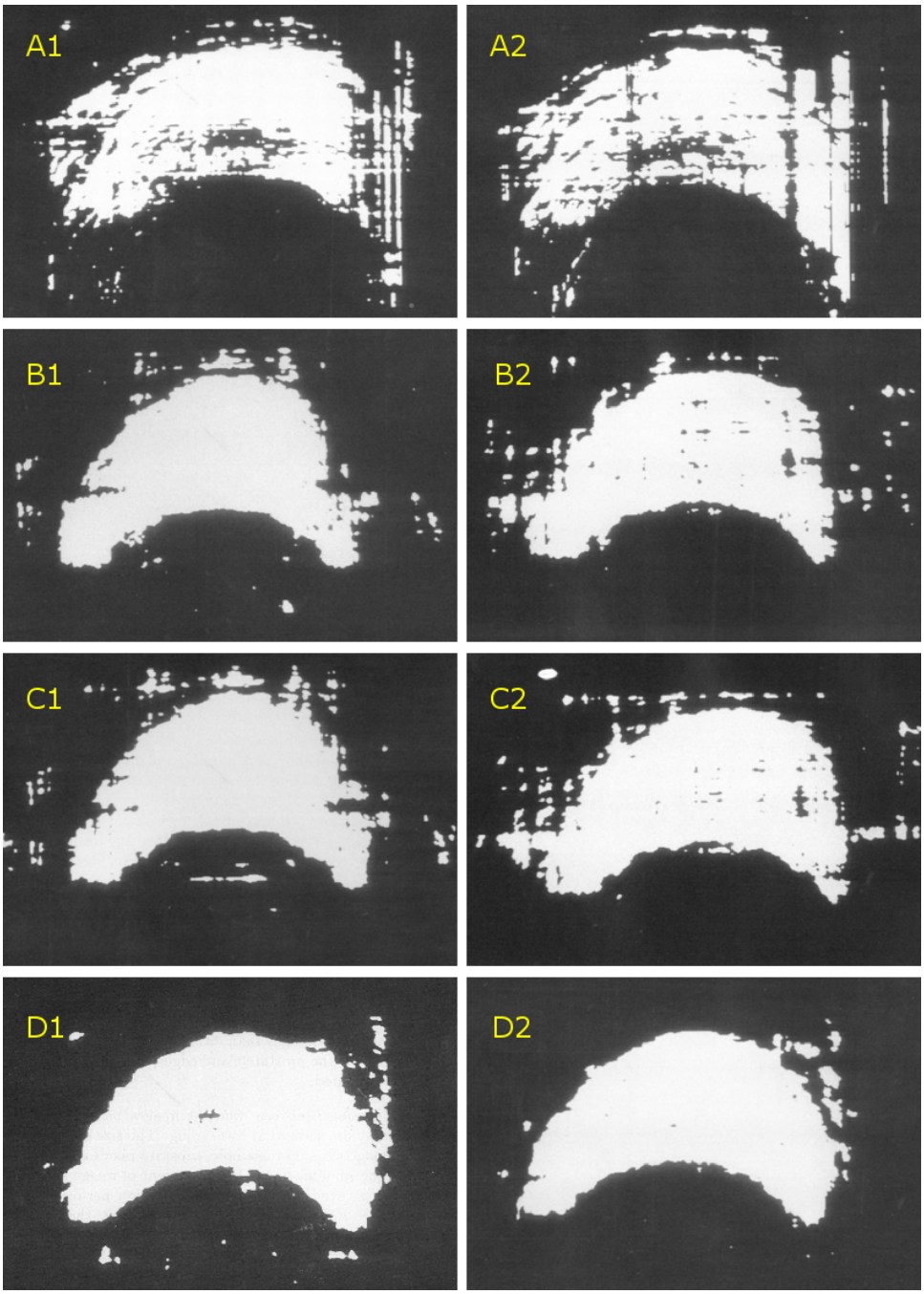

Figure 3.11: (Subplots A1-A2) Segmentation of the image slice 22 and 20 respectively using the 5-6-2 neural network. (Subplots B1-B2) Segmentation using the 132-30-2 neural network. (Subplots C1-C2) Segmentation using the 132-30-2 network with larger training set. (Subplots D1-D2) Segmentation using the 61-10-2 network architecture. [12]

Algorithm Outline

In the first stage of the algorithm, four different texture feature images are extracted from the TRUS image by convolving the 5×5 Laws feature masks [15] with the original image. The convolution masks are shown in Fig. 3.12 below. Next, an "energy" measure is calculated for each pixel in the four texture feature images and subsequently smoothed. Afterwards, a K-means clustering algorithm is applied to the feature vectors obtained from the four distinct "energy" images to determine the number of classes, as well as the mean and variance of each class. Subsequently, an initial probabilistic membership is assigned to each pixel within the TRUS image. This is followed by a time consuming iterative probabilistic relaxation algorithm [16] to update the memberships of each pixel until a predetermined number of iterations has been performed. The main objective of the probabilistic relaxation algorithm is to enforce spatial constraints between class memberships of neighbouring pixels. Finally, each pixel in the original TRUS image is assigned a class label, as determined by the highest class probability membership value. The overall algorithm is outlined in Fig. 3.13 below.

-1	-4	-6	-4	-1
-2	-8	-12	-8	-2
0	0	0	0	0
2	8	12	8	2
1	4	6	4	1

a)

1	-4	6	-4	1
-4	16	-24	16	-4
6	-24	36	-24	6
-4	16	-24	16	-4
1	-4	6	-4	1

b)

-1	0	2	0	-1
-2	0	4	0	-2
0	0	0	0	0
2	0	-4	0	2
1	0	-2	0	1

c)

-1	0	2	0	-1
-4	0	8	0	-4
-6	0	12	0	-6
-4	0	8	0	-4
-1	0	2	0	-1

d)

Figure 3.12: a)-d) Laws' 5×5 micro-texture masks [14]

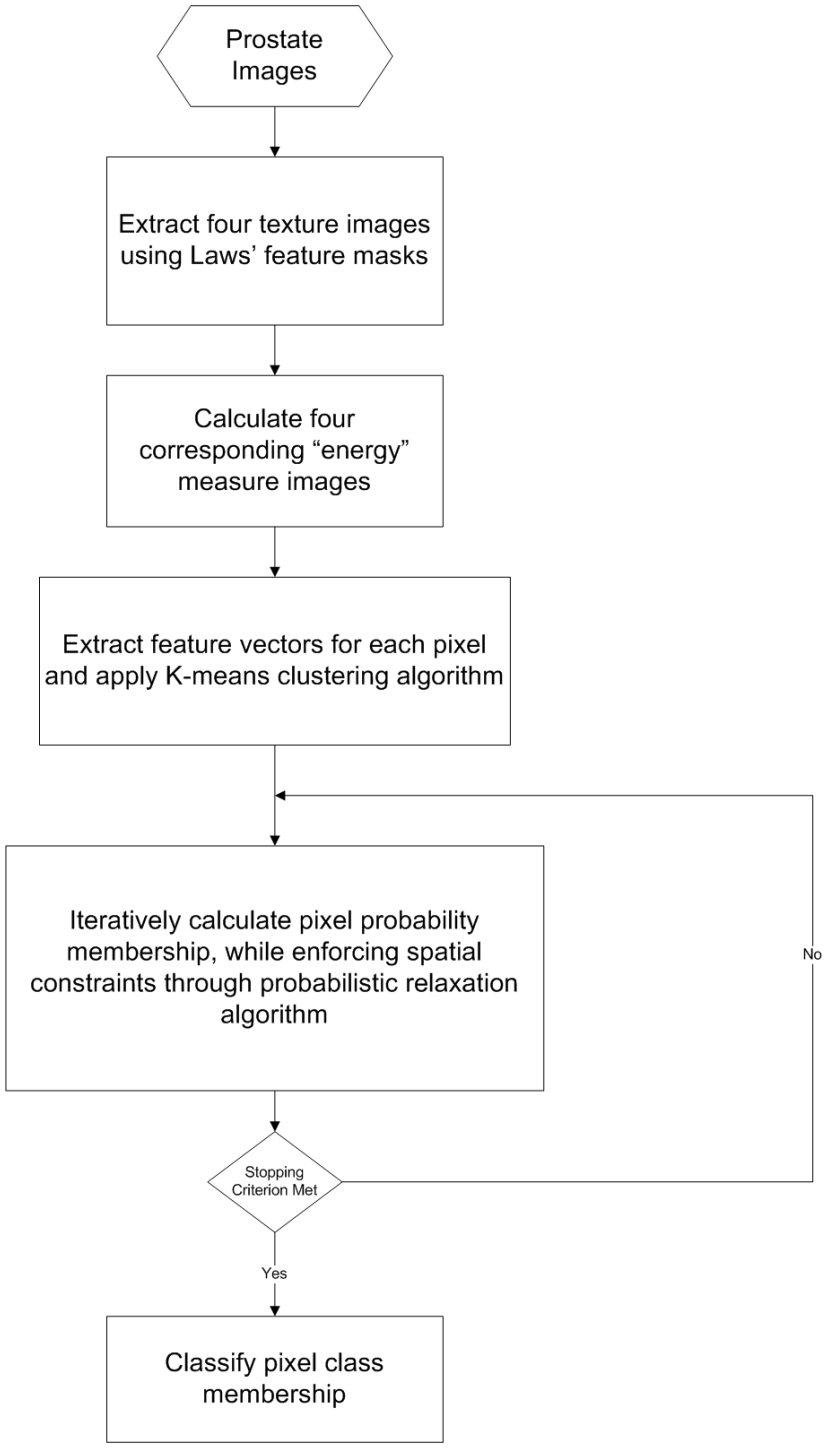

Figure 3.13: Automatic texture based classification algorithm outline

Automatic Segmentation Method

Assuming the original TRUS image is of size $M \times N$ and denoted by $f[i, j]$, and the four 5×5 Laws' convolution masks are denoted by $h[i, j]$. Then the four texture feature images are computed by convolving the original image with each of the convolution masks as follows [14]:

$$g[i, j] = h[i, j] * f[i, j] = \sum_{m=-2}^{m=2} \sum_{n=-2}^{n=2} h[m, n] f[i + m, j + n]. \tag{3.28}$$

Subsequently, texture images $h[i, j]$ are converted to $s[i, j]$, which calculates the sample mean deviation around each pixel according to [14]:

$$s[i, j] = \frac{1}{(2n + 1)^2} \sum_{k=i-n}^{i-n} \sum_{l=j-n}^{j+n} |g[k, l] - \mu|, \tag{3.29}$$

where μ is the mean of the particular texture image $g[i, j]$ given by

$$\mu = \frac{1}{MN} \sum_{ij} g[i, j]. \tag{3.30}$$

the resulting $s[i, j]$ images have been termed texture-energy measure images by the authors [14]. Then, each pixel location from the original TRUS image can now be associated with a feature vector with dimensions 1×4, consisting of feature values from the four texture-energy measure images at the same pixel location.

Subsequently, the $M \times N$ feature vectors are grouped into distinct clusters via using a slight variation of the K-means clustering algorithm. Interestingly, the authors do not limit the number of distinct clusters to two, corresponding to the prostate and the background. Instead, the number of clusters is determined by a distance heuristic. Specifically, the algorithm continues to define new cluster seed points as long as the seed point is at least k (user defined parameter) times the standard deviation away from the previously defined seed points.

Once the classes are defined from the K-means algorithm, the mean and covariance of the feature vectors belonging to each class can be obtained. Then, assuming Gaussian distribution for each class, each pixel corresponding to a feature vector can be classified using the Bayes classifier, obtaining [14]

$$P_i^0(\lambda) = P_i(\lambda \mid \mathbf{x}) \tag{3.31}$$

where $P_i^0(\lambda)$ is the initial probability membership of pixel i belonging to class λ, which is equal to the conditional probability $P_i(\lambda \mid \mathbf{x})$, with $\mathbf{x}$ being a particular feature vector.

Thus far, the above initial probability labeling is entirely based on the feature vector values associated with each pixel. This method alone may be able to produce valid segmentation results, given that the distance between feature vectors from each class is sufficiently large. However, it is well known that typical TRUS images suffer from poor image quality, which makes it difficult to visually distinguish between the prostate and background regions. Hence, in the texture-energy measure feature space, it would still be difficult to identify distinct clusters. Therefore, the authors introduce an additional spatial constraint on the segmentation via a probability relaxation algorithm outlined below.

First, an estimate of the overall probability membership of a label class λ is defined as [14]

$$\hat{P}(\lambda) = \frac{1}{n} \sum_{i=1}^{n} P_i^0(\lambda), \tag{3.32}$$

where n is the number of pixels in the overall TRUS image. Then the joint probability of a pixel with class membership λ and a neighboring pixel with class membership λ' is estimated as [14]

$$\hat{P}_{i,i+\delta}(\lambda, \lambda') = \frac{1}{n} \sum_{i=1}^{n} P_i^0(\lambda) P_{i+\delta}^0(\lambda'), \delta \in \Delta \tag{3.33}$$

where $i+\delta$ is a specific neighbor, and Δ is the 8 pixel neighborhood of the pixel of interest. Next, a compatibility coefficient r is computed as follows [14]

$$r_{i,i+\delta}(\lambda, \lambda') = \frac{\hat{P}_{i,i+\delta}(\lambda, \lambda')}{\hat{P}(\lambda)\hat{P}(\lambda')} \tag{3.34}$$

Once the compatibility coefficient is computed, the following iterative update equations are applied to re-assign the probability class memberships of each pixel as follows [14]

$$S_{i,i+\delta}^{k+1}(\lambda) = \sum_{\lambda' \in \Delta} \hat{P}_i^k(\lambda) \hat{P}_{i+\delta}^k(\lambda') r_{i,i+\delta}(\lambda, \lambda') \tag{3.35a}$$

$$q_{i,i+\delta}^{k+1}(\lambda) = \frac{S_{i,i+\delta}^{k+1}(\lambda)}{\sum_{\lambda' \in \Delta} S_{i,i+\delta}^{k+1}(\lambda')}, \delta \in \Delta \tag{3.35b}$$

$$P_i^{k+1}(\lambda) = \frac{1}{8} \sum q_{i,i+\delta}^{k+1}(\lambda), \delta \in \Delta \tag{3.35c}$$

where (3.35a) computes the likelihood that a pixel is a member of a class λ while assuming another class λ' is one of its neighbours. Equation (3.35b) normalizes the likelihood, and equation (3.35c) updates the probability membership of pixel i according to the normalized likelihood.

The above update equations are computationally expensive, and are not guaranteed to converge. Thus, the number of iterations is assigned a fixed value by the authors [14]. Finally, once the number of iterations is reached, each individual pixel is classified based upon its highest probability membership of a particular class.

In this method, Richard and Keen [14] have demonstrated an unsupervised segmentation algorithm based on texture features of the prostate in TRUS images. The proposed algorithm utilizes more image information as compared to the classic edge and intensity based methods, and also has the advantage of allowing one to forgo a prior training process. However, iterative updates of the probability relaxation algorithm is very expensive in terms of computations, and is also not guaranteed to converge to a stationary point. Furthermore, the initial clustering algorithm may produce an unexpected number of classes, since the K-means clustering algorithm is very sensitive to the parameter k, which governs the cluster distance heuristic. Nevertheless, the authors have demonstrated the feasibility of an automated texture based segmentation algorithm, and it is mentioned that significant speedups can be obtained via the use of parallel computations.

3.3.3 Feature Classification Using Support Vector Machines

A similar feature based classification approach was recently presented by Zaim et al. [17]. However, the method of feature extraction and the choice of classifier is fundamentally different from the unsupervised clustering method of [14]. Specifically, Zaim et al. proposes to extract features of the prostate and non-prostate regions via the use of multi-wavelet decomposition. Multi-wavelets have shown to possess attractive mathematical properties such as orthogonality, symmetry and having compact support. Due to these desirable properties and their ability to characterize local texture information, they have been applied to a host of object analysis and recognition problems with high levels of success [18, 19]. Therefore, the authors have chosen multi-wavelets towards the problem of feature extraction. Moreover, in order to obtain local features pertaining to the TRUS image, the authors [17] propose to apply the multi-wavelet decomposition to overlapping sub-windows inside the original image rather than the entire image itself. Finally, the obtained feature vectors are then used to train a Support Vector Machine (SVM) classifier for distinguishing between prostate and non-prostate tissues.

Multiwavelet Coefficients and Feature Extraction

The original wavelet transform converted data from the spatial domain to the frequency domain and selectively stored each frequency band information with a corresponding matching resolution scale. As a natural extension to scalar wavelets, multi-wavelets uses multiple scaling and wavelet functions to perform the notion of multi-resolution analysis. In particular, an arbitrary number of N scaling functions $\Phi(x) = [\Phi_1(x), \Phi_2(x), \ldots \Phi_N(x)]$ and N wavelet functions $\Psi(x) = [\Psi_1(x), \Psi_2(x), \ldots \Psi_N(x)]$ can be used. In the work presented by Zaim et al. [17], the number of scaling and wavelet functions used is two (i.e. N = 2) and the two-scale equations become [17]

$$\Phi(x) = \sqrt{2} \sum L(k)\Phi(2x - k) \tag{3.36a}$$

$$\Psi(x) = \sqrt{2} \sum H(k)\Psi(2x - k) \tag{3.36b}$$

where L and H are low-pass and high-pass filter coefficients respectively for a 2-channel multi-filter bank. In particular, the wavelets chosen in the authors' study are constructed based on the work of Geronimo et al. [20].

Next, in order to extract feature vectors from the original TRUS image, a multi-wavelet decomposition procedure is first applied. Specifically, an iterative low-pass and high-pass filtering of the original image using the two different scaling and wavelet functions are executed. The result of the decomposition can be explained in Fig. 3.14 below:

L_1L_1	L_2L_1	H_1L_1	H_2L_1
L_1L_2	L_2L_2	H_1L_2	H_2L_2
L_1H_1	L_2H_1	H_1H_1	H_2H_1
L_1H_2	L_2H_2	H_1H_2	H_2H_2

Figure 3.14: Subdivisions of first level multi-wavelet decomposition [17]

where subscripts denote the membership of the scaling or wavelet function. For example, the block L_2H_1 contains low-pass coefficients corresponding to the second scaling

function and high-pass coefficients corresponding to the first wavelet function, both of which are applied in the vertical direction (bottom left set of blocks). Similarly, the gray blocks represents the low-pass coefficients produced by both scaling functions and are also used for the second level multi-wavelet decomposition, if necessary. The above mentioned procedure is then repeated until a desired resolution scale is achieved.

In the approach of [17], the features of an image are extracted in "patches"; which means that an image is first divided into overlapping sub-images of size 32×32 and subsequently, the multi-wavelet decomposition is applied to each of these patches individually. The latter are defined as follows [17]

$$P_n \mid_{n=1,2,...M} \text{ and } < X_{ij} > \subset P \tag{3.37}$$

where M is the number of overlapping patches and X_{ij} are pixels at row i and column j belonging to the aforementioned patches. This subdivision of images is illustrated in Fig. 3.15.

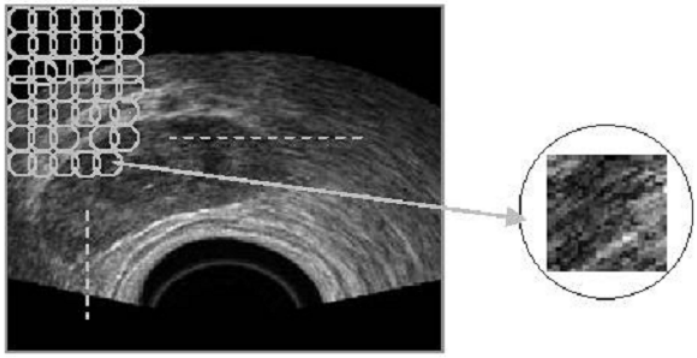

Figure 3.15: Subdivision of images into overlapping patches [17]

After applying multi-wavelet decomposition to each patch of an image, the magnitudes of the wavelet coefficients are collected for each patch as a matrix of coefficients. Then for each patch, energy (E) and entropy (Γ) measures are calculated according to [17]

$$\Gamma = \frac{\sum_{i=1}^{m}\sum_{j=1}^{m} X_{ij}^2}{N^2} \tag{3.38a}$$

$$\mathrm{E} = -\frac{\sum_{i=1}^{m}\sum_{j=1}^{m} (X_{ij}^2) log(X_{ij}^2)}{N^2} \tag{3.38b}$$

Subsequently, a feature vector V can be constructed for each patch consisting of energy and entropy values for the multi-wavelet decomposition blocks i.e.

$$V = [\Gamma_1 \mathrm{E}_1 \Gamma_2 \mathrm{E}_2 \ldots \Gamma_K \mathrm{E}_K] \tag{3.39}$$

where K is the number of wavelet decomposition blocks (e.g. $K = 16$ for the first level multi-wavelet decomposition).

Kernel Support Vector Machine (KSVM)

The authors in [17] make use of KSVMs to perform the classification task based on the extracted feature vectors from the previous section. KSVMs are typically a two-class classifier that has shown desirable properties, and has been applied to the domain of texture classification [21] with promising results.

Essentially, given a set of points in the feature space along with the correct labelling, the SVM algorithm finds an optimal hyperplane or decision boundary that maximizes the number of correct point classifications as well as maximizing the distance between the points of differing classes. However, the original SVM theory only produced a linear decision boundary in the feature space, which may be an oversimplification of the classification problem. Thus, in the case of KSVMs, an additional transformation is applied to the original input feature space, which transforms the original space into a higher dimensional space (one in which a good hyperplane decision boundary could be found). Conversely, the hyperplane in the higher dimensional space becomes a non-linear decision boundary in the original input feature space. In particular the discriminate hyperplane in the work of [17] is defined as follows

$$f(x) = \sum_{i=1}^{l} y_i a_i k(x, x_i) + b \tag{3.40}$$

where $k(x, x_i)$ is a radial basis function kernel that transforms the input feature space to a higher dimensional space and the sign of $f(x)$ indicates the membership of feature vector x.

Algorithm Validation Results

The authors [17] obtained a total of 10 TRUS images for the study and used a three-fold cross validation technique to evaluate their proposed algorithm. Specifically, the dataset were split randomly three times and 80% of the patches falling inside the prostate region and 80% of the patches falling outside of the prostate region are used for training the

KSVM. For the training stage, a patch is considered to be inside or outside of the prostate region if over 75% of the patch area is within the desired region, otherwise it is not used as a training patch. The rest of the data set, the 20%, from the TRUS images are used for validation and comparison purposes. The results of the authors' [17] proposed approach is summarized in Table 3.2 below.

Table 3.2: Segmentation Results of Multi-wavelet KSVM

Validation Set	Classification Accuracy (%) 1st-Level Decomposition	2nd-Level Decomposition
(1)	76	79
(2)	72	81
(3)	78	86
Mean	75.3	82.0

Unsurprisingly, the classification accuracy for using the second level multi-wavelet decomposition (82%) is higher than the first level decomposition features (75.3%). Through the use of multi-wavelet analysis on a local window scale, the authors [17] were able to capture locally the fine details associated with the prostate. This is in contrast to the coarse structure of the prostate that would have been detected by applying the multi-wavelet decomposition to the entire TRUS image. However, one obvious limitation of this approach lies in the fact that the patches are 32×32 pixels in size, and each individual patch is labeled as either prostate or non-prostate regions. Therefore, segmentation artifacts may occur for the patches that are located directly on the edge of the prostate boundary.

3.3.4 Summary of Classification Based Methods

As an alternative to the classical edge based segmentation algorithms mentioned in Section 3.2, another widespread approach to the segmentation problem takes advantage of the tools of supervised [12, 17] and unsupervised [14] machine learning. Specifically, the method described in [12] employs an optimized feed-forward neural network to carry out the segmentation, which attempts to capture the visual cues used by experts to classify image pixels as either inside or outside of the prostate region. In particular, the chosen features correspond to the local pixel intensity means surrounding the POI. In a more recent approach [17], the image features used by a KSVM classifier are defined via subdividing the original TRUS image into small patches and subsequently projecting them onto a basis of multi-wavelets. The idea behind their approach lies in the attempt to extract complex non-linear features at different scales and orientations within local patches

of a TRUS image. This rich source of texture information can then be used to train a robust KSVM classifier. On the side of unsupervised learning, a probabilistic clustering procedure detailed in [14] segments the prostate region based on the features obtained using a specially designed filter bank [15]. Despite the strong theoretical foundation of the classification-based segmentation algorithms, they remain prone to over-segmentation because of their disregard for the morphological features of the prostate gland. Specifically, the observed visual segmentation results often contain patches of pixels far away from the true prostate region but are incorrectly labeled as part of the prostate. Moreover, it is well known that the trained classifiers are highly sensitive to the type and quantity of input features that are provided to the classifier, which can significantly influence the accuracy of segmentation.

3.4 Shape Prior Based Segmentation Algorithms

3.4.1 Using Deformable Superellipses for Prostate Segmentation

Due to the inherent difficulties associated with the poor quality of TRUS images, classic segmentation algorithms relying on gradient and texture information alone often encountered considerable difficulties when applied in the context of TRUS prostate images. These techniques either fail completely, or require significant post-processing to eliminate false positive regions in the final segmentation result. To address these difficulties, shape prior restrictions have been recently incorporated into segmentation algorithms to further constrain the extracted prostate boundaries to be smooth and continuous. One of the prominent approaches in this recent research area is presented by Gong et al. [22]. In their work, the authors have modelled the prostate shape using deformable superellipses based on manual prostate contours that were delineated by five experts. Subsequently, the shape prior model constraints are combined with imagery data as a basis for their Bayesian segmentation framework.

Prostate Shape Modelling Using Superellipses

It is commonly known that in TRUS images, the prostate boundaries generally form smooth, closed, and nearly convex contours. Therefore, the authors [22] have chosen to model the outer shape of the prostate using deformable superellipses. They are selected for the fact that they can be adjusted to fit a large variety of natural shapes, which conforms well to the standard prostate boundaries. Specifically, a centered superellipse can be defined as [22]

$$\begin{cases} x(\theta) = a_x cos^{\epsilon}\theta \\ y(\theta) = a_y sin^{\epsilon}\theta \qquad \text{and} - \pi \leq \theta \leq \pi \end{cases} \tag{3.41}$$

where a_x and a_y denote the lengths of the semi axes and ϵ specifies the "squareness" of the ellipse. Moreover, global deformations can be applied to these superellipses to cause tapering and bending. The detailed mathematical formulation of such deformations are presented in the work of Gong et al. [22]. In particular, a superellipse can be fully characterized by a vector $\mathbf{P} = \{lx, ly, r, sy, sq, xy, t, b\}$, where the first four parameters, $\mathbf{P}_p = \{lx, ly, r, sy\}$ represents the pose parameters denoting translations in the x and y directions, rotation, and scaling respectively. The last four parameters, $\mathbf{P}_s = \{sq, xy, t, b\}$ represents the deformation parameters, denoting squareness, aspect ratio, tapering, and bending respectively. Examples of such deformations are illustrated in Fig. 3.16 below.

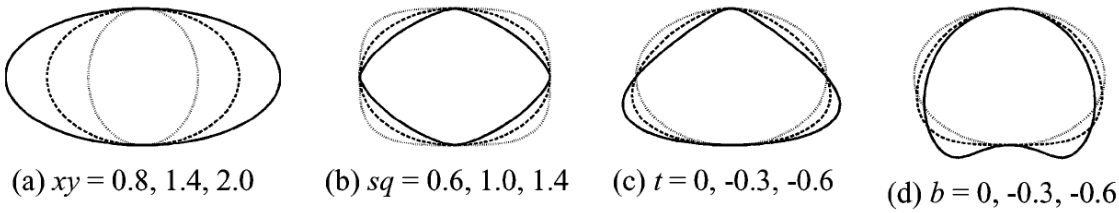

Figure 3.16: Different deformations applied to superellipses [22]

Given a set of manual delineations, an optimal set of superellipse parameters $\mathbf{P}$ can be obtained for each contour. Subsequently, the authors of [22] were able to study statistical characteristics of each parameter and construct a statistical shape model.

Bayesian Segmentation Framework

In the segmentation framework of [22], the ultimate goal is to obtain a set of optimal parameters $\mathbf{P}$, which closely characterizes the prostate boundary of an unsegmented TRUS image. The authors define this in a maximum *a posterior* estimation framework, which naturally incorporates prior knowledge regarding the statistical properties of the model superellipse.

Specifically, the authors model the shape parameters $\mathbf{P_s}$ prior using a multivariate Gaussian distribution $\Pr(\mathbf{P_s})$ and the pose parameter $\mathbf{P_p}$ prior using a uniform distribution $\mathrm{U}(\mathbf{P_p})$. Assuming shape and pose are independent, and using the edge strength map E of a enhanced TRUS image as a likelihood, the *a posterior* probability of the deformed boundary given E can be defined as [22]

$$\begin{aligned} \Pr(\mathbf{P} \mid E) &= \frac{\Pr(\mathbf{P})\Pr(E \mid \mathbf{P})}{\Pr(E)} \\ &= \frac{\mathrm{U}(\mathbf{P}_p)\Pr(\mathbf{P}_s)\Pr(E \mid \mathbf{P})}{\Pr(E)} \end{aligned} \tag{3.42}$$

The maximization of this *a posterior* probability incorporates a prior bias of likely shapes, which also coincides with the strong edges within the image. The prior knowledge is obtained via learning the optimal superellipse parameters for a set of manually delineated training images.

Segmentation Results

In the clinical study of the work presented in [22], 125 TRUS images were obtained from 16 patients, which were subsequently delineated by 5 experts. The average of the manual delineations is treated as the ground truth. In their experimental study, TRUS images from 15 of the patients were used for training purposes. Specifically, prior knowledge in equation 3.42 regarding the statistical properties of the superellipse parameters $\mathbf{P}$ were obtained from ground truth images of the 15 patients. Then, the TRUS images for the last patient were used for evaluation. This is repeated for all 16 patients, resulting in 16 such leave-one-out tests.

In order to obtain the edge strength map E of a given TRUS image, a robust edge detection algorithm [4] is employed, which is also used in the Edge Guided Manual Boundary Delineation method described in section 3.2.1. In particular, the original TRUS image is filtered using the sticks algorithm to reduce speckle noise and selectively enhance the image contrast. Then, the enhanced image is anisotropically smoothed with the edge-preserving weak membrane algorithm. Lastly, a knowledge-based filtering method is applied along with the standard Canny edge detector to obtain a robust edge map of the original image. Examples of such edge maps are illustrated in Fig. 3.17 below.

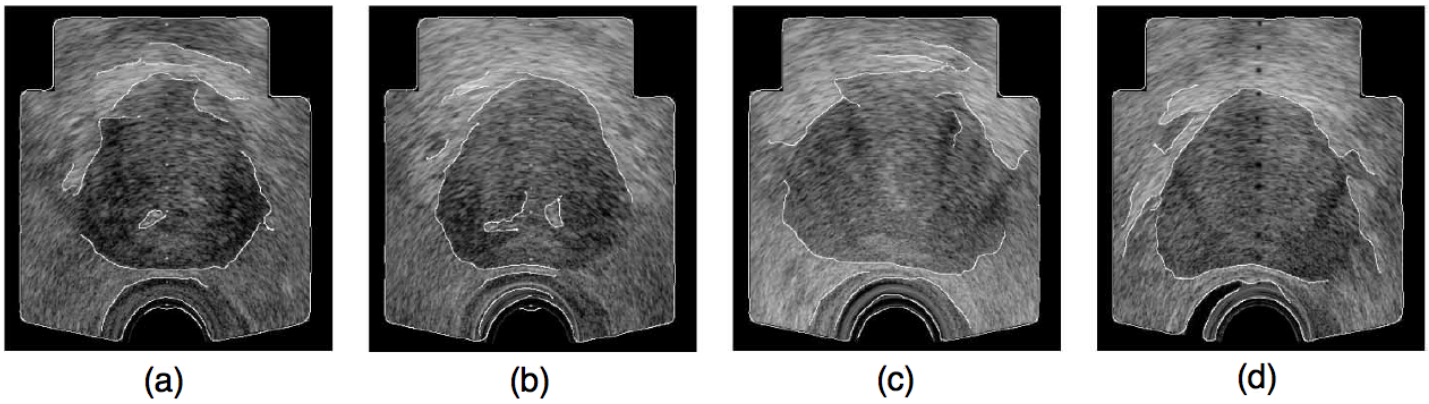

Figure 3.17: Extraction of the edge strength maps using a robust algorithm [22]

Now, given an unsegmented TRUS image, a user is required to first select four points on the prostate boundary to provide an estimate of scale or size of the prostate. Then, the scaled mean shape (obtained from the training set) is used as the initial starting point for the segmentation process. After initialization, the superellipse parameters are optimized to maximize the *a posterior* probability given in equation 3.42. Examples of TRUS segmentation results obtained by their algorithm is shown below in FIg. 3.18

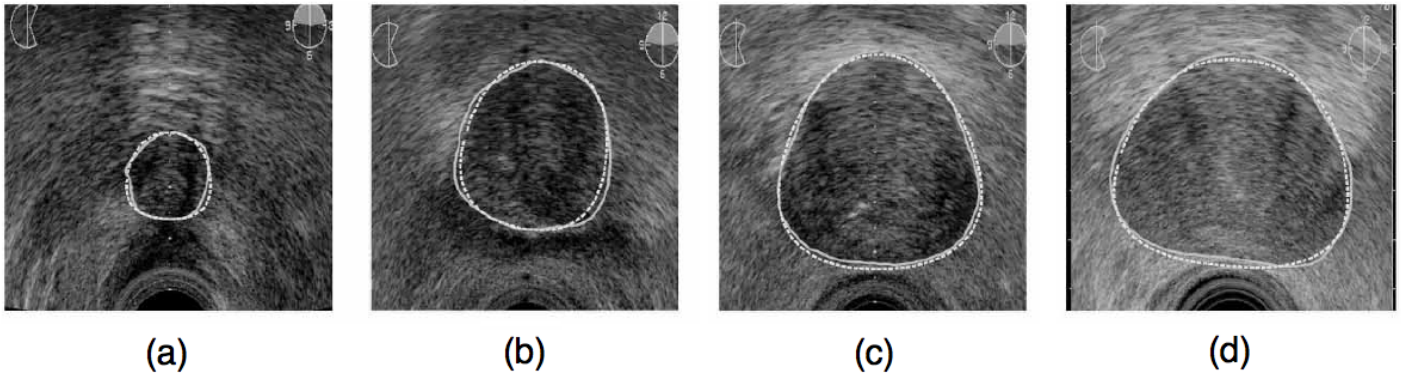

Figure 3.18: Examples of proposed segmentation method results. The solid contours are ground truth obtained via averaging the manual delineations of the 5 experts. The dotted contours are obtained via the Bayesian segmentation scheme. [22]

The observed segmentation results are very promising and the authors reported from the testing results that the mean distance between the algorithm generated segmentations

to the ground truth was 1.36 ± 0.58 mm, which is significantly smaller than the average inter-observer distance (1.82 ± 1.44 mm). Hence, a fast and promising semi-automatic segmentation method was reported by Gong et al. [22]. However, the authors did note that the definition of the proposed deformable superellipses can only accurately represent symmetric prostate shapes. This is confirmed from visual inspection of the results presented in Fig. 3.18, where the computer-generated contours closely fit the prostates (fairly symmetric) within the image. On the other hand, in the case of non-symmetric prostates, the performance of the proposed method would suffer.

3.4.2 Prostate Segmentation Using Statistical Shape Model

A hierarchical segmentation strategy is presented in the work of [23], where the prostate boundary is characterized by a set of multi-scale and multi-orientation Gabor coefficients. These features are learned from a set of training images, and the subsequent segmentation is driven towards similar regions using a hierarchical deformation strategy. Specifically, the segmentation process initially uses coarse scale Gabor features, and the subsequent deformation stages uses finer scale Gabor features. In addition to the use of local neighbourhood features, an additional restraint is introduced in the form of shape priors. In particular, the authors [23] obtain a set of manual prostate contours, and calculate a set of *attribute vectors* corresponding to local geometric characteristics of the prostate shapes. Then, the mean shape model or attribute vector is incorporated into the segmentation framework to further constrain the segmenting curve.

Algorithm Outline

The algorithm presented by Shen et al. [23] consists of three main components. These include 1) calculation of the attribute vector associated with the prostate shape model, 2) extraction of multi-scale Gabor features, and 3) hierarchical deformation segmentation using the previously obtained information. This is summarized in Fig. 3.19. In particular, the shape statistics of the prostate can be calculated from a set of training samples that were manually outlined, and the image features in the prostate image are represented by a set of rotationally-invariant Gabor features at multiple scales. Finally, the segmenting contour is deformed under forces from image features, as well as internal and external statistical constraints. These will be described in more detail in the following sections.

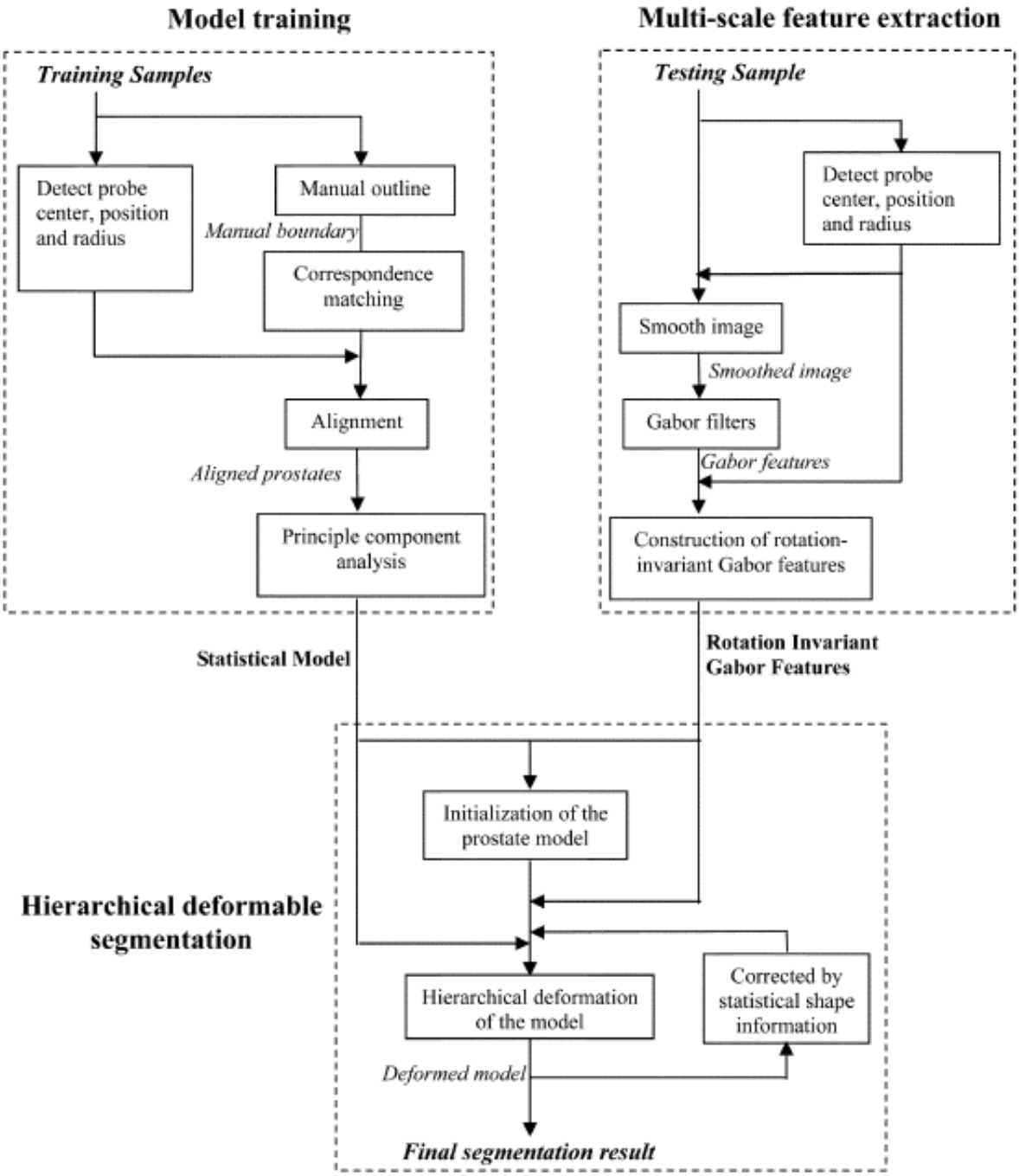

Figure 3.19: Algorithm outline of the hierarchical segmentation strategy presented in [23]

Shape Model

In order to provide a mathematical description of the geometric characteristics of the prostate shapes, the manual contours from the training images are first sampled to obtain a set of equidistant points. That is, contour C is represented by a set of ordered points such that $C = \{V_i, i = 1, 2, \ldots, M\}$. For the method reported in [23], each pair of consecutive points is approximately 2 pixels apart. Next, the authors define an affine-invariant attribute vector for each point V_i in the prostate shapes. The attribute vector values consist of the set of the triangular areas formed by point V_i and the neighbouring points $\{V_{i+vs}, V_{i-vs}\}$. This is illustrated pictorially in Fig. 3.20 below.

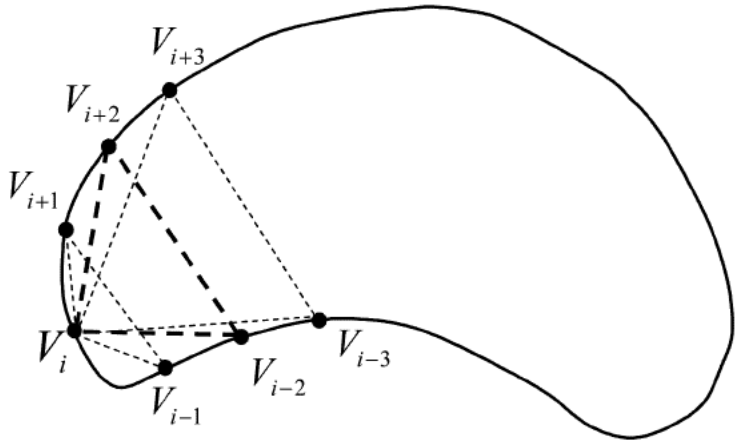

Figure 3.20: The area of a triangle formed by three points $(V_{i-vs}, V_i, V_{i+vs})$ is the vsth element of the attribute vector for point V_i [23]

The set of triangular areas for each point V_i provides information regarding the local geometric characteristics around the particular point. Subsequently, the areas from different neighbours (varying distance away) are concatenated into an attribute vector $A(V_i)$, which can be made affine-invariant by a normalization procedure [23], i.e.,

$$\hat{A}(V_i) = \frac{A(V_i)}{\sum_i \|A(V_i)\| \Big/ \sum_i 1.0} \tag{3.43}$$

where $\|A(V_i)\|$ is the magnitude of the vector $A(V_i)$, and $\hat{A}(V_i)$ is the affine-invariant attribute vector for the ith point V_i.

Multi-scale Feature Representation

A bank of Gabor filters were used to capture image features at different scales [24]. The mother function of the Gabor filter is defined in the spatial domain as [23]

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right], \tag{3.44}$$

with the corresponding Fourier transform as [23]

$$G(u,v) = \exp\left\{-\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right\}, \tag{3.45}$$

where $\sigma_u = 1/2\pi\sigma_x$ and $\sigma_v = 1/2\pi\sigma_y$, and W being a shift parameter along the u axis in the frequency domain.

Given the mother function, a Gabor filter bank can be obtained via dilation and rotation of the original mother function. Specifically, assuming one would like to have a set of Gabor filters with K orientations and S scales. Then, one can set the basic rotation and scaling factors to be, $\psi = \pi/K$ and $a = (U_h/U_l)^{1/S-1}$ respectively; with U_h and U_l being the parameters determining the upper and lower frequency bands respectively. Then a bank of Gabor filters, $g_{s,k}$ can be defined as [23]

$$g_{s,k}(x,y) = a^s g(a^s(x\,\cos(k\psi)+y\,\sin(k\psi)), a^s(-x\,\sin(k\psi)+y\,\cos(k\psi))), \tag{3.46}$$

where $s = 0,1,\ldots,S$ and $k = 0,1,\ldots,K$. Hence, for each pixel in the TRUS image, a feature vector $\{F_{s,k}(x,y\}$ can be calculated, where $F_{s,k}(x,y)$ represents the (s,k)th Gabor feature obtained via the filter $g_{s,k}$ at point (x,y).

The authors [23] chose the Gabor filter bank as the method of choice for extracting the image features because of three main reasons. First, the Gabor filter is a complex filter with real and imaginary components. The real part of the filter behaves as a smoothing filter, which attenuates noise in the TRUS image, while the imaginary part of the filter acts as a standard edge detector. Second of all, because of the multi-resolution structure of the Gabor filters, it allows a coarse-to-fine hierarchical implementation of the segmentation procedure. Lastly, the multi-orientation structure of the Gabor filters allows the extraction of both edge strength and edge direction in the TRUS images of prostates.

Moreover, the authors have devised a procedure of transforming the set of regular Gabor features, $\{F_{s,k}\}$ to the space of rotationally-invariant Gabor features $\{\tilde{F}_{s,k}\}$ through resampling and interpolation, [23], i.e.,

$$\tilde{F}_{s,k}(x,y) = (1-\beta)F_{s,k_0+k}(x,y) + \beta F_{s,(k_0+1)+k}(x,y), \qquad k = 0,\ldots,K-1 \tag{3.47}$$

where $k_0 = \lfloor \theta_{(x,y)}/\psi \rfloor$ and $\beta = (\theta_{(x,y)}/\psi) - k_0$. The resulting rotationally-invariant Gabor feature images at each scale s and orientation k $\{\tilde{F}_{s,k}\}$ can be further separated in to its real $\{\tilde{F}^{real}_{s,k}\}$ and imaginary $\{\tilde{F}^{img}_{s,k}\}$ parts respectively. An illustration of Gabor feature images at different scales and orientations is shown in Fig. 3.21.

Segmentation Framework

After defining the model and feature representations of the TRUS prostate images, the goal of the authors' [23] approach is to minimize an energy function associated with the deformable prostate model (segmenting contour). In particular, the energy function consists of two terms, including an external and an internal energy term. The external energy term is associated with the image features produced by the Gabor filter bank, which drives the segmenting curve towards the dark-to-light transitions of the TRUS prostate image. On the other hand, the internal energy term corresponds to the similarity between the attribute vectors of the deformable model and the training shape model. Therefore, it constrains the evolution of the segmenting contour to minimally deviate from the learned prostate shape, which is the mean of 10 manually delineated training shapes.

The overall energy cost function is defined as [23]

$$E = \sum_{i=1}^{M} E_i = \sum_{i=1}^{M} (w_{ext} E_i^{ext} + w_{int} E_i^{int}) \tag{3.48}$$

where E_i is the overall energy associated with the ith point on the deformable model, and w_{ext} and w_{int} are two user defined weighting parameters for the external and internal energy terms respectively.

The external energy term E_i^{ext} can be further separated into the real and imaginary components [23] as follows

$$E_i^{ext} = w_{img} E_i^{img} + w_{real} E_i^{real} \tag{3.49}$$

In particular, the imaginary energy term E_i^{img} would drive the deformable model towards the dark-to-light transitions inside the TRUS image, which is captured by the rotation-invariant Gabor features $\{\tilde{F}^{img}_{s,k}\}$. These large intensity transitions typically correspond to the boundaries of the prostate in a TRUS image. Moreover, minimizing the real energy term E_i^{real} satisfies a second criterion, which states that for each model point, the image

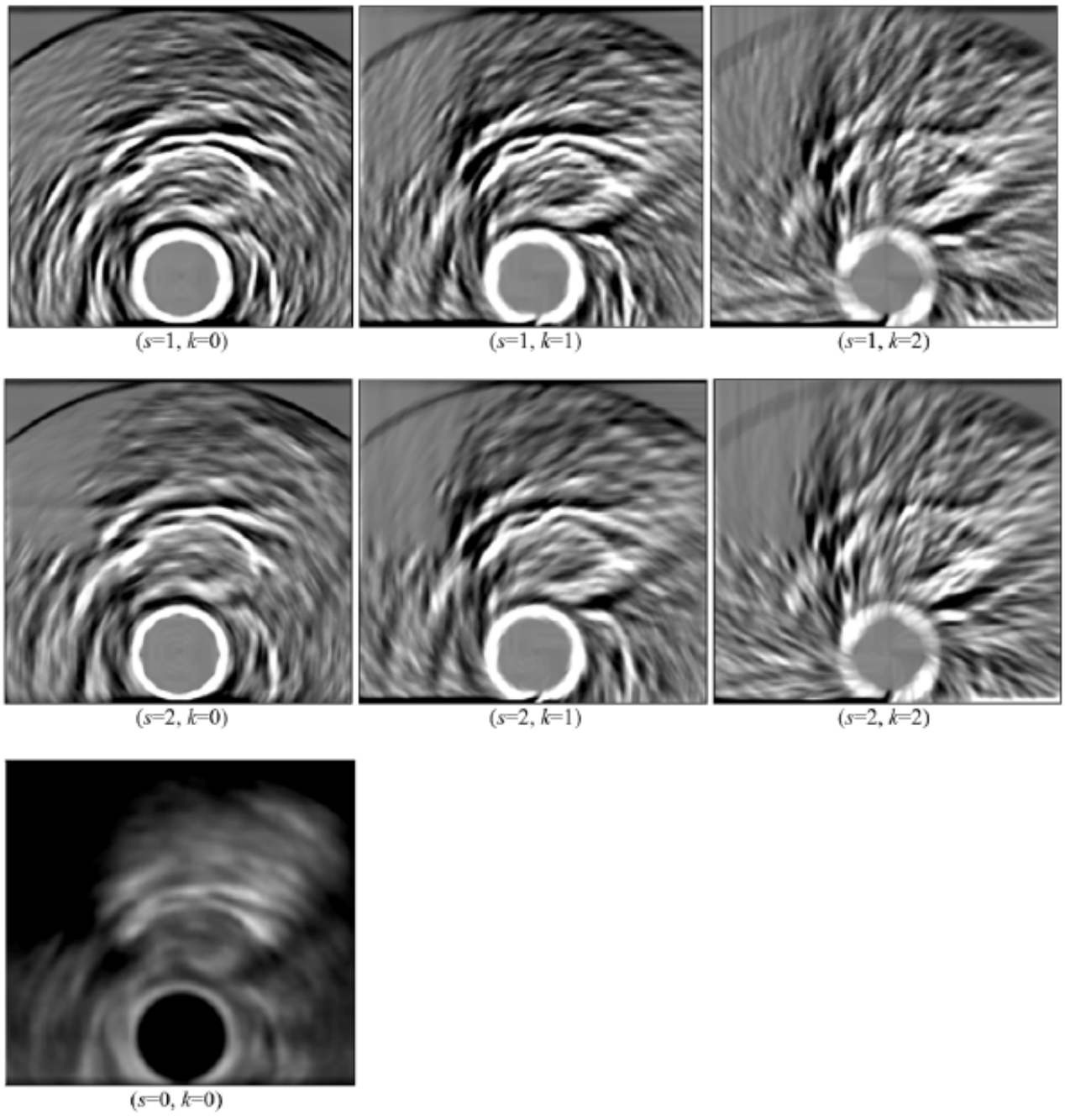

Figure 3.21: The imaginary Gabor feature images for varying scales and orientations are shown in the top two rows, and the bottom left image shows the real Gabor feature image corresponding to scale $s = 0$ and orientation $k = 0$ respectively [23]

features in its nearby neighbourhood should match well with those in the training model point.

Both external energy terms are defined [23] below

$$E_i^{img} = \sum_{s=0}^{S-1} w_s \left(\sum_{k=0}^{K-1} -\tilde{F}_{s,k}^{img}(x_i, y_i)(\vec{n}_k \cdot \vec{n}_i) \right) \tag{3.50a}$$

$$E_i^{real} = \sum_{\forall (x,y) \in N(x_i,y_i)} \sum_{k=0}^{K-1} \left(\tilde{F}_{0,k}^{real}(x, y) - \overline{\tilde{F}_{0,k}^{real}}(x', y') \right)^2 \tag{3.50b}$$

where $\vec{n}_k$ denotes the direction of the Gabor filter at position (x_i, y_i) and $\vec{n}_i$ denotes the normal direction of the current segmenting contour at point i in equation 3.50a. As for the E_i^{real} term in equation 3.50b, the similarity is calculated between the neighbourhood around point (x_i, y_i) in the deformable model space and the corresponding neighbourhood in the training model space (x', y'). $\overline{\tilde{F}_{0,k}^{real}}(x', y')$ is the average feature value calculated from the training samples at the position (x', y') in the training model space. Note that since the algorithm requires comparison between neighbourhoods in the deformable model space and the training model space, an additional optimization procedure is required to register the two models in question.

The internal energy term E_i^{int} in equation 3.48 is used to enforce the prior shape knowledge obtained from a set of training shapes. This further restricts the segmentation when images features from the external energy term alone is not sufficient in producing an accurate segmentation. As mentioned earlier, the authors associate an attribute vector with each point on the deformable model contour, which provides local geometric information of the underlying contour shape. This information is also used to preserve the shape of the deformable model during the segmentation process. Therefore, the internal energy E_i^{int} is defined [23] as

$$E_i^{int} = \| \hat{A}^{Def}(V_i) - \hat{A}^{Mdl}(V_i) \|^2 \tag{3.51}$$

where $\hat{A}^{Def}(V_i)$ and $\hat{A}^{Mdl}(V_i)$ are the attribute vectors corresponding to the deformable model and the training model shape respectively at the point V_i.

In the end, the final segmentation is achieved via minimizing the total energy function 3.48 using a greedy algorithm. Additionally, the segmentation is carried in a coarse-to-fine hierarchical structure, in which the deformable model is first evolved using the Gabor features from the coarse scales. As the number of iterations increases, the main focus of the Gabor features shifts to the finer scales. This hierarchical segmentation is implemented by dynamically adjusting the weighting parameters w_s of equation 3.50a.

Segmentation Results

The prostate model in the work presented by Shen et al. [23] is based on the average of 10 training prostate shape models that were obtained via manual delineation. The training shapes and the mean shape model are depicted in Fig. 3.22

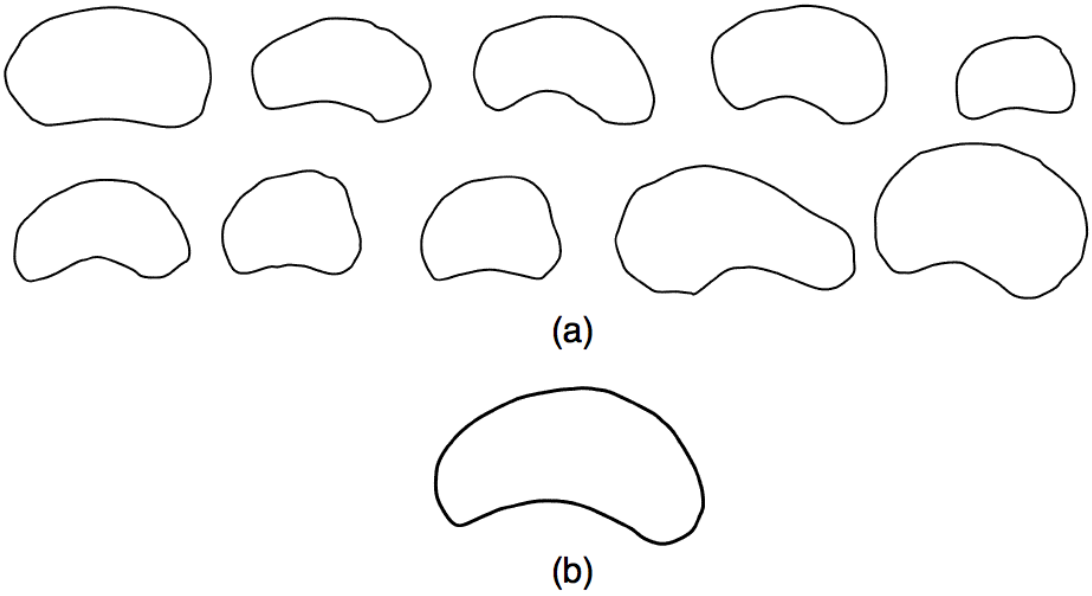

Figure 3.22: (a) Ten manually delineated training shape models. (b) Average shape model [23]

The authors validated their algorithm by comparing the computer-generated segmentations to the ground truth manual segmentations of 8 TRUS images, shown in Fig. 3.23. The segmentations results are very promising from visual inspection, and the reported mean area accuracy is 96.02%.

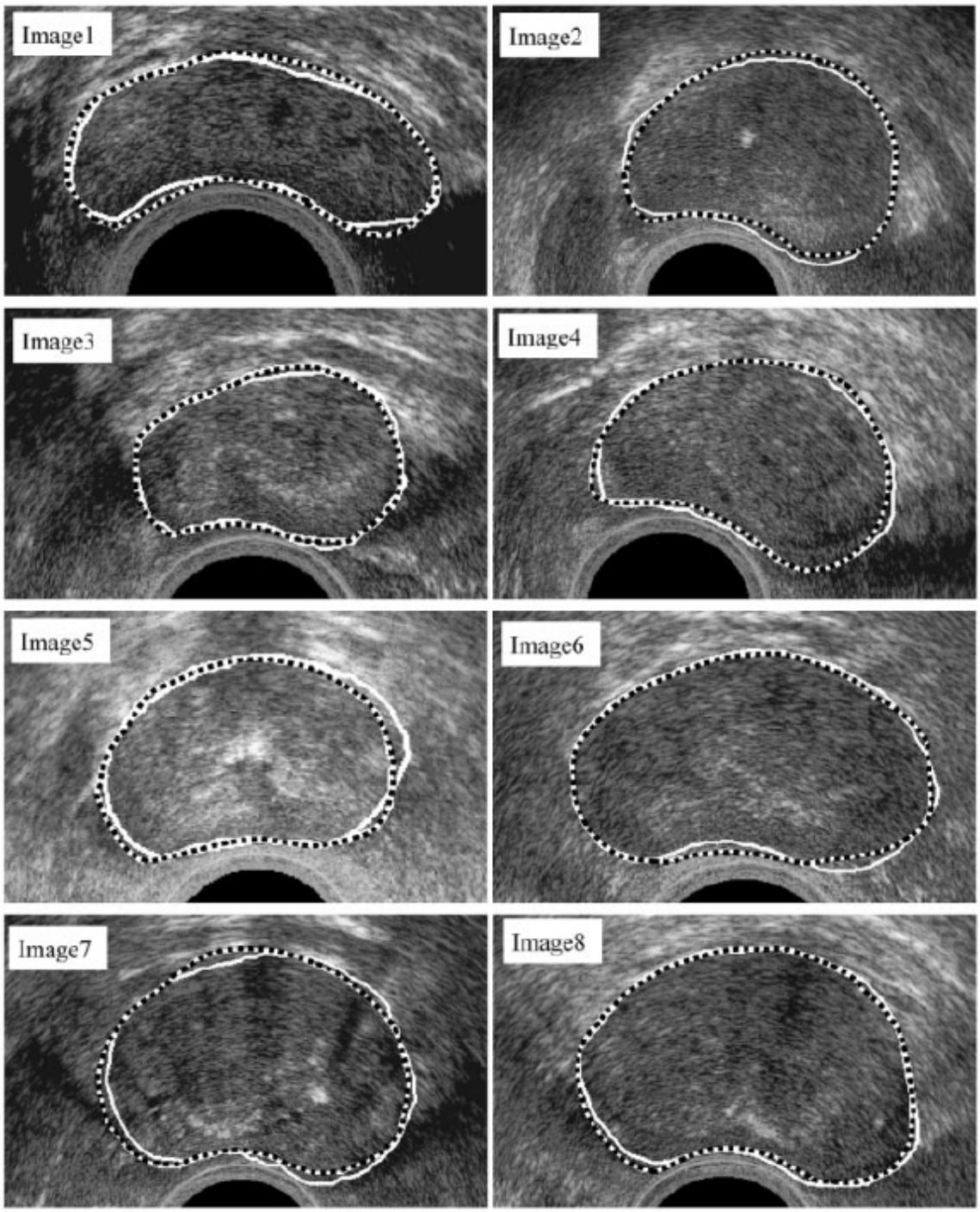

Figure 3.23: Comparison of manual and algorithm segmentations. Dashed contours are the manual segmentations and the solid contours are algorithm generated segmentations. [23]

3.4.3 Summary of Shape Prior Based Methods

The use of shape prior based segmentation techniques towards the problem of prostate segmentation in TRUS images is a fairly recent innovation. Information provided by these *shape priors* are often used to regularize the segmentation procedure, which has naturally increased the accuracy of segmentation results. Generally, the results obtained by these algorithms incorporating shape priors are smooth and retain expected prostate configurations. This is in stark contrast to the irregular prostate segmentations that are generated by many of the classification based algorithms, which often required some form of post-processing cleanup procedure. Hence, shape prior based methods is a new paradigm in segmentation procedure that is indeed very promising.

Although the trend in TRUS segmentation problems has shifted towards these shape prior based methods, they are not without their own challenges. Specifically, the question of how to define such shape models in an accurate and robust manner, as well as how much influence the prior shape information should have in the overall segmentation framework are still actively being explored.

One of the well known shape prior based segmentation methods is described in the work of [22], which models the prostate boundary as a deformable super-ellipse. Subsequently, a maximum-a-posteriori estimation framework is used to find a contour (i.e. a boundary of the prostate) that closely matches the prior shape model on one hand, and coincides with strong image gradients on the other. An obvious limitation of the above approach, as noted by the authors [22], stems from the restrictive nature of the prior model used, which can only approximate symmetric prostate shapes. A similarly restrictive model is used in the work of [23], which uses a set of attribute vectors associated with a single mean shape as the prior information. The authors of [23] uses a hierarchical shape deformation procedure to move the initial segmentation towards the location of the true boundaries of the prostates based on a set of Gabor features as well as the aforementioned shape prior. However, this algorithm is hindered by a prerequisite to register the TRUS image to the prior shape model before any geometric similarities can be measured. Additionally, the use of a greedy minimization algorithm can also lead to relatively large computations times. Overall, the shape models used in the above mentioned methods can be fairly restrictive, and in the case where a prostate's shape is dissimilar to the mean training shape model, a reduction in the segmentation accuracy may occur.

Chapter 4

Proposed Algorithm

4.1 Overview of Proposed Algorithm

In the proposed algorithm presented in this thesis, a different approach to the problem of segmentation of TRUS images is introduced. In particular, the segmentation approach should satisfy a number of essential criteria, which consist of generality, adaptability, and consistency. To this end, the algorithm should be general and adaptable in terms of the information that is used to drive the segmentation process. In some ways, the proposed method is similar to the neural network approach, because it uses information from a wide range of sources, including an arbitrary number of edge-related, texture-related, and morphological features. However, the proposed algorithm uses an active contour framework, which poses an inherent restriction on the evolution of the segmentation region. Moreover, the method takes advantage of the concept of distribution tracking [25], which encodes prior knowledge regarding the object of interest in the form of statistical distributions. The latter methodology is used to achieve consistency, as it can be applied to handle both photometric and morphological features of the prostate, thereby providing a unified framework for the prostate segmentation problem.

Finally, it is important to note that the adaptation of using the distribution tracking approach towards morphological features of the prostate introduces a novel type of statistical shape prior, which is defined as a form of "weak" shape priors. It is *weak* in the sense that it is a type of shape prior that is restrictive just enough to regularize the process of segmentation, but is still let to be dominated by information contained in the acquired images. This methodology was chosen because it appeared that previous shape prior definitions in literature were often very restrictive in nature. Although, this is desirable in the case of segmenting rigid objects under occlusion, it is not well justified in the case of TRUS imaging. Since the prostates are rarely occluded, and their boundary deformations vary

from patient to patient, the use of highly restrictive shape priors may actually introduce a sizeable bias into the estimated shapes. Hence, only the proposed "weak" shape priors are incorporated into the segmentation algorithm presented in this thesis.

Specifically, this "weak" shape prior information is encapsulated in the form of a probability density function (pdf) of an intrinsic geometric parameter of the true prostate boundary. Subsequently, a segmentation boundary is forced to evolve into a configuration whose empirical pdf of the same geometric parameter closely matches the model pdf as assessed by the Bhattacharyya distance measure [26, 27]. In this sense, the proposed algorithm is a natural extension to the work of [25, 28] to tracking of both morphological and photometric features of the object of interest (in this case prostates). Moreover, it is proven experimentally that the proposed methodology can achieve promising results for segmentation of TRUS images under the condition of low image contrast and poorly observable prostate boundaries.

4.2 Preprocessing Technique

As discussed in the previous sections, the quality of TRUS images of the prostate are often poor as a result of imaging artifacts such as low contrast and SNR, speckle noise, and shadowing. All of these characteristics contribute to the high level of difficulty in the segmentation task. Thus, it is not uncommon to observe the use of pre-processing techniques in past literature to first reduce the amount of speckle noise in a TRUS image before proceeding with the main segmentation procedure. Examples of such filters include the "sticks" filter used in the work of [4], as well as the Lee and Frost filters mentioned in the works of [29, 30]. The first filter was already discussed in section 3.2.1. The latter de-speckle filters (Lee and Frost) are both window based methods that adaptively adjusts the strength of smoothing depending on the coefficient of variation inside the particular kernel. Although both Lee and Frost filters are intended to be feature and edge-preserving in nature, certain limitations corresponding to their design do exist. In particular, the filters are sensitive to the size and shape of the filter window, and they only inhibit smoothing near edges rather than providing edge enhancement.

In the proposed algorithm of this thesis, a smoothing filter based on the methodology of anisotropic diffusion was selected. Specifically, the speckle reducing anisotropic diffusion (SRAD) filter described in the work of Yu and Acton [11] was chosen for pre-processing the TRUS images of the prostate. The SRAD filtering algorithm involves a partial differential equation (PDE) approach to speckle reduction, which allows the generation of an image scale space without bias due to filter window size and shape. It differs from past adaptive de-speckle filters in that it not only enhances edges by inhibiting diffusion across them, but it also allows diffusion to continue in the parallel direction along either side of an edge.

Moreover, SRAD is adaptive and does not require the explicit definition of a hard threshold to alter its smoothing behaviour in homogeneous and feature rich regions.

4.2.1 Speckle Reducing Anisotropic Diffusion

The SRAD framework is defined in [11] as follows. Given an intensity image $I_0(x, y)$ over the image support Ω, the output image $I(x, y, t)$ is evolved according to the following PDE:

$$\begin{cases} \partial I(x,y,t)/\partial t = \mathrm{div}[c(q)\nabla I(x,y,t)] \\ I(x,y,0) = I_0(x,y), (\partial I(x,y,t)/\partial \vec{n}) \mid_{\partial\Omega} = 0 \end{cases} \tag{4.1}$$

where $\partial\Omega$ denotes the border of Ω, and $\vec{n}$ is the outer normal to $\partial\Omega$, and $c(q)$ the diffusion coefficient can be defined [11] as

$$c(q) = \exp\{-[q^2(x,y,t) - q_0^2(t)]/[q_0^2(t)(1+q_0^2(t))]\}. \tag{4.2}$$

where $q(x, y, t)$ is the *instantaneous coefficient of variation* (ICOV), which is defined [11] as

$$q(x,y,t) = \sqrt{\frac{(1/2)(|\nabla I|/I)^2 - (1/16)(\nabla^2 I/I)^2}{[1+(1/4)(\nabla^2 I/I)]^2}}. \tag{4.3}$$

Finally, $q_0(t)$ is the speckle scale function, which is a parameter that controls the amount of smoothing applied to the image by SRAD. It is estimated [11] as

$$q_0(t) = \frac{\sqrt{\mathrm{var}[z(t)]}}{\overline{z(t)}} \tag{4.4}$$

where $\mathrm{var}[z(t)]$ and $\overline{z(t)}$ are the intensity variance and mean over a homogeneous region of the image at time t, respectively.

The ICOV $q(x, y, t)$ acts as a more robust edge detector for speckled images. In particular, it combines a normalized gradient magnitude operator and a normalized Laplacian operator in its formulation. The authors [11] report that high relative gradient magnitude and low relative Laplacian tend to indicate an edge in both bright and dark regions of a speckled image. Subsequently, both $q(x, y, t)$ and the speckle scale function $q_0(t)$ are input parameters for equation 4.2, which controls the amount of anisotropic diffusion for the TRUS image at time t. Typical SRAD filtering results are depicted in Fig. 4.1. As observed from visual inspection, the amount of speckle pattern is significantly reduced. Furthermore, edges belonging to the true boundary of the prostate are fairly well preserved. Nevertheless, certain areas of the true prostate boundary are still only partially observable due to the presence of shadowing artifacts. The latter still causes the texture features near the shadow regions to be difficult classify as either inside or outside of the true prostate region.

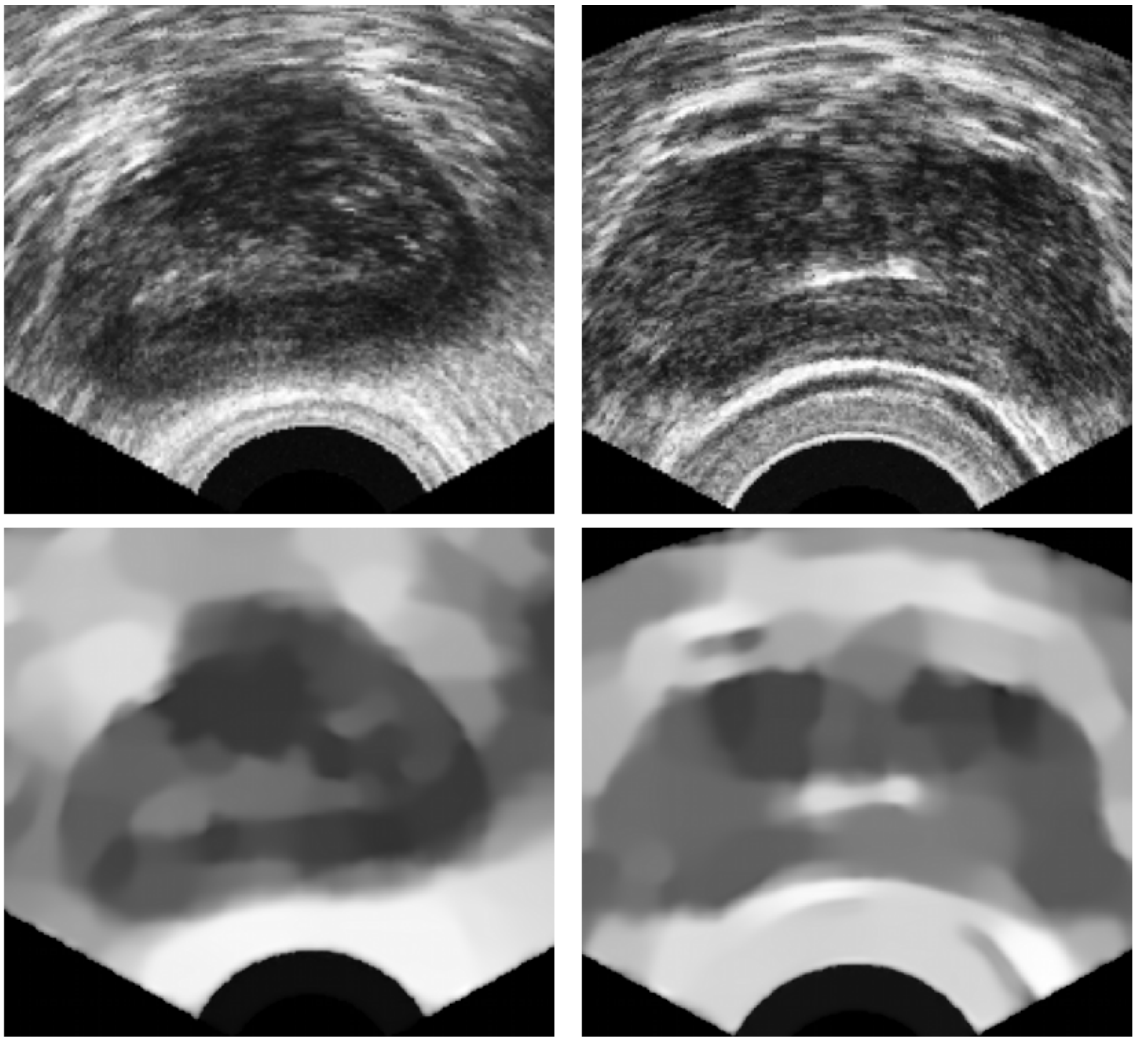

Figure 4.1: Top row: images of the original TRUS images of prostates. Bottom row: SRAD filtered images.

4.3 Information Tracking Approach

In this section, the main framework of the information tracking approach towards the problem of segmentation is presented. Specifically, the method of encoding both photometric and morphological prior information is described. This is followed by defining the segmentation problem as an optimization task, in which a region within the TRUS image is identified to be maximally similar to the previously defined statistical prior information in terms of the Bhattacharyya distance. The proposed algorithm is described in more detail below.

4.3.1 Tracking of Texture Related Features

Let u be an ultrasound image which we consider to be a scalar-valued functioned defined over a subset Ω of $\mathbb{R}^2$. In general, u can be transformed into a *vector-valued* image $I : \Omega \rightarrow \mathbb{R}^d$ of its associated (local) features, with d equal to the number of features. The transition from u to I can be formally described by a map $\mathcal{M} : u \rightarrow I$, which is equal to identity in the case when the only features of interest are the grey-levels of u. In general, however, at each $\mathbf{x} \in \Omega$, $I(\mathbf{x})$ may consist of, for example, the local statistics of $I(\mathbf{x})$, its associated Laws' texture features [15], multi-resolution Gabor features, or any combination thereof. The choice of features is arbitrary, and selected based on the strength of discrimination.

Now, let Ω_t (with the subscript t standing for "target") be a subset of Ω over which the object of interest (i.e. prostate) is supported. The segmentation algorithm proposed is then based on the following two assumptions. First, for each $\mathbf{x} \in \Omega_t$, the d components of $I(\mathbf{x})$ are assumed to be independent random variables. Second, it is assumed that one is provided with reliable estimates of the *probability density functions* (pdf) of the image features pertaining to the prostate region Ω_t. The second assumption is justified by the existence of manually delineated TRUS images, which are provided by an expert radiologist. Given the manually delineated images, one can extract the values of the features of interest within the boundaries of the prostate ground truth regions. These values in turn can be used to estimate their corresponding pdf's by means of the standard kernel density estimation procedure [31].

Let $P_t(\mathbf{z})$, where $\mathbf{z} \in \mathbb{R}^d$, be the joint pdf of the image features associated with the prostate region Ω_t. Due to the assumption of statistical independence of the image features, $P_t(\mathbf{z})$ can be factorized as $P_t(\mathbf{z}) = \prod_{k=1}^{d} p_t(z_k)$, with $p_t(z_k)$ being the pdf of the k-th image feature $z_k \in \mathbb{R}$. Given Ω_t, the pdf $p_t(z_k)$ can be approximated according to

$$p_t(z_k) = \mathcal{E}\left\{ \frac{\int_{\Omega_t} K(z_k - I_k^{tr}(\mathbf{x}))\, d\mathbf{x}}{\int_{\Omega_t} d\mathbf{x}} \right\}, \tag{4.5}$$

where K denotes a (positive valued, normalized) *kernel* function , $I_k^{tr}(\mathbf{x})$ is the value of the k-th component of a training image I^{tr} at pixel location $\mathbf{x}$, and the $\mathcal{E}$ stands for the averaging operator taken over the training set. These estimates of $p_t(z_k)$ are computed at the training stage before the actual segmentation is initialized. The computation of these "target" (model) distributions is illustrated visually in Fig. 4.2.

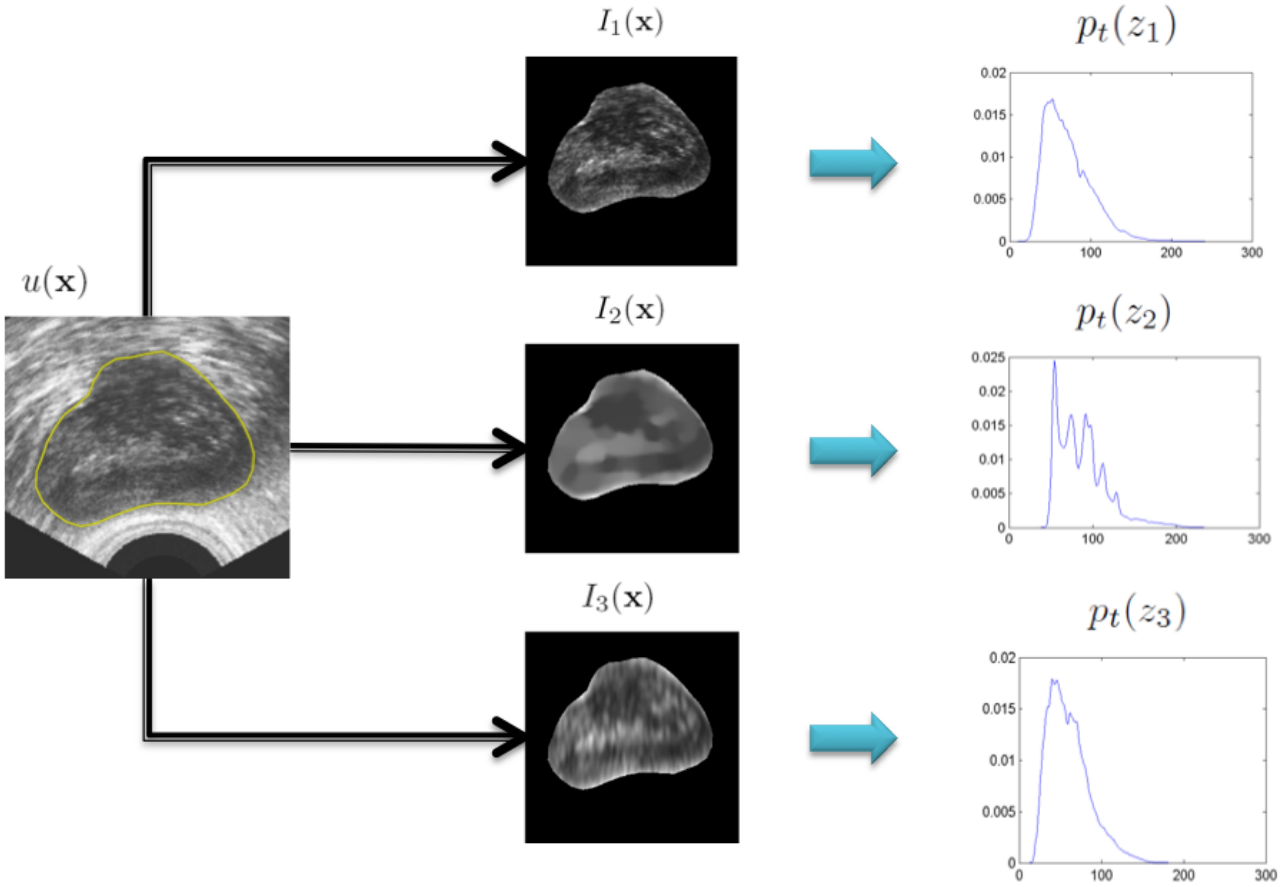

Figure 4.2: An example construction of photometric model distributions is shown here. The left-most column shows an original TRUS image u of a prostate with a manual delineation depicted in yellow. The middle column shows the extracted vector-valued image I corresponding to the known prostate region. Finally the right-most column shows the estimated pdf's of the corresponding features values computed via (4.5)

Now, let u be an observed image to be segmented, and I be its corresponding feature image. Also, let $\Omega_{in} \subseteq \Omega$ be a non-empty arbitrary subset of Ω. In this case, an approximation similar to (4.5) can be utilized to estimate the *empirical* pdf $P(\mathbf{z} \mid \Omega_{in})$ of the

features of u observed over Ω_{in}. Again, due to the assumption of statistical independence, $P(\mathbf{z} \mid \Omega_{in})$ can be factorized similarly to P_t with its d factors computed according to

$$p(z_k \mid \Omega_{in}) = \frac{\int_{\Omega_{in}} K(z_k - I_k(\mathbf{x}))\, d\mathbf{x}}{\int_{\Omega_{in}} d\mathbf{x}}, \quad \text{with } k = 1, 2, \ldots, d, \tag{4.6}$$

such that $P(\mathbf{z} \mid \Omega_{in}) = \prod_{k=1}^{d} p(z_k \mid \Omega_{in})$.

Ultimately, the objective of the proposed segmentation algorithm is to evolve the region Ω_{in} so as to closely approximate the subset Ω_t, which supports the ground truth region of the prostate. Moreover, whenever Ω_{in} coincides with Ω_t, the distribution $P(\mathbf{z} \mid \Omega_{in}$ and $P_t(\mathbf{z})$ should be close to one another in terms of a distance measure. In particular, the measure of similarity between the two probability distributions is chosen to be the Bhattacharyya coefficient B as defined in [26, 27, 32]

$$B(\Omega_{in}) = \int_{\mathbb{R}^d} \sqrt{P_t(\mathbf{z}) P(\mathbf{z} \mid \Omega_{in}) d\mathbf{z}}. \tag{4.7}$$

Alternatively, using the assumption of mutual independence of z_k, one can rewrite (4.7) as

$$B(\Omega_{in}) = \prod_{k=1}^{d} B_k(\Omega_{in}) = \prod_{k=1}^{d} \int \sqrt{p_t(z_k) p(z_k \mid \Omega_{in})}, \tag{4.8}$$

The Bhattacharyya coefficient $B(\Omega_{in})$ achieves its maximum of 1 when $p_t(z_k) = p(z_k \mid \Omega_{in}), \forall k$, which occurs when the sets Ω_{in} and Ω_t coincide with each other. Conversely, by maximizing B as a function of Ω_{in}, it is reasonable to expect a minimal discrepancy between Ω_{in} and Ω_t. This forms the basis of the segmentation procedure presented in [25], as well as the proposed algorithm.

4.3.2 Level-set Formulation

In recent years, the level-set framework [33] has gained wide popularity in the area of image segmentation because of its number of remarkable advantages. In particular, level-set methods allow one to perform numerical computations involving curves and surfaces without the need to explicitly parameterize these objects. Thus, in the proposed algorithm it was chosen to represent the subset Ω_{in} in the level-set framework as follows

$$\Omega_{in} = \{\mathbf{x} \in \Omega \mid \varphi(\mathbf{x}) \leq 0\}, \tag{4.9}$$

where $\varphi : \Omega \to \mathbb{R}$ is a *level-set function*, whose zero level-set $\{\mathbf{x} \in \Omega \mid \varphi(\mathbf{x}) = 0\}$ is used to implicitly define the corresponding active contour representing the estimated prostate

boundary. In general, the level-set function φ is defined as a *signed distance function* [33], which leads to convenient and efficient numerical computations. Specifically, φ is defined as follows

$$\varphi(\mathbf{x}) = \begin{cases} -d(\mathbf{x}, \partial\Omega_{in}), & \mathbf{x} \in \Omega_{in} \\ +d(\mathbf{x}, \partial\Omega_{in}), & \mathbf{x} \notin \Omega_{in} \end{cases} \tag{4.10}$$

where $\partial\Omega_{in}$ denotes the boundary of the region supported by Ω_{in}, and $d(\cdot)$ represents the minimal distance between point $\mathbf{x} \in \mathbb{R}^2$ in the image and the closest point on the boundary $\partial\Omega_{in}$.

Expressing Ω_{in} in terms of its corresponding level-set function φ leads to an alternative definition of the Bhattacharyya coefficient, which now becomes a function of φ i.e.

$$B(\varphi) = \prod_{k=1}^{d} B_k(\varphi) = \prod_{k=1}^{d} \int \sqrt{p_t(z_k)\, p(z_k \mid \varphi)}\, dz_k, \tag{4.11}$$

where the densities $p(z_k \mid \varphi)$ can be computed according to

$$p(z_k \mid \varphi) = \frac{\int_\Omega K(z_k - I_k(\mathbf{x}))\, \mathcal{H}(-\varphi(\mathbf{x}))\, d\mathbf{x}}{\int_\Omega \mathcal{H}(-\varphi(\mathbf{x}))\, d\mathbf{x}}, \tag{4.12}$$

with $\mathcal{H}(x) = (x)_+$ denoting the standard Heaviside function. Subsequently, the problem of finding an optimal Ω_{in} can be posed as an equivalent problem of finding an optimal level-set function φ^* as given by

$$\varphi^* = \arg\max_{\varphi \in \Phi} B(\varphi), \tag{4.13}$$

where Φ denotes a set of signed distance functions [33].

A local maximizer of $B(\varphi)$ can be computed via the *steepest ascent* procedure, which approximates φ^* as a stationary point of the sequence of solutions produced by the gradient flow

$$\frac{\partial\varphi(\mathbf{x}, \tau)}{\partial\tau} = \frac{\delta B(\varphi)}{\delta\varphi}, \tag{4.14}$$

with τ being an artificial time parameter introduced to represent the iterative update applied to the level-set function $\varphi(\mathbf{x}, \tau)$ ($\varphi(\mathbf{x}, 0) = \varphi_0(\mathbf{x})$ represents an initialization of the level-set function), and $\partial B(\varphi)/\partial\varphi$ standing for the first variation of $B(\varphi)$ computed with respect to the level-set function φ. The derivation of $\partial B(\varphi)/\partial\varphi$ is shown below.

To derive an expression for $\partial B(\varphi)/\partial\varphi$, it is first noted that the first variation of $p(z_k \mid \varphi)$ in (4.12) is equal to

$$\frac{\delta p(z_k \mid \varphi)}{\delta\varphi} = A^{-1}\delta(\varphi(\mathbf{x}))\Big(p(z_k \mid \varphi(\mathbf{x})) - K(z_k - I_k(\mathbf{x}))\Big), \tag{4.15}$$

where $\delta(\cdot)$ is a Dirac delta function, and

$$A = \int_{\Omega} \mathcal{H}(-\varphi(\mathbf{x}))\, d\mathbf{x}. \tag{4.16}$$

Consequently, each of the d coefficients $B_k(\varphi(\mathbf{x}))$ in (4.8) has its first variation defined as

$$\begin{aligned}\frac{\delta B_k(\varphi)}{\delta \varphi} &= \int_{\mathbb{R}} \frac{p_t(z_k)}{2\sqrt{p_t(z_k)p(z_k \mid \varphi(\mathbf{x}))}} \frac{\delta p(z_k \mid \varphi)}{\delta \varphi}\, dz_k \\ &= \delta(\varphi(\mathbf{x}))\, \frac{1}{2A} \left(B_k(\varphi(\mathbf{x})) - \int_{\mathbb{R}} \sqrt{\frac{p_t(z_k)}{p(z_k \mid \varphi(\mathbf{x}))}}\, K(z_k - I_k(\mathbf{x}))\, dz_k \right). \end{aligned} \tag{4.17}$$

The above expression can be rewritten in a more concise form as

$$\frac{\delta B_k(\varphi)}{\delta \varphi} = \delta(\varphi(\mathbf{x}))\, \frac{1}{2A} \left(B_k(\varphi(\mathbf{x})) - \Big[r(z_k \mid \varphi(\mathbf{x})) * K(z_k) \Big]_{z_k = I_k(\mathbf{x})} \right), \tag{4.18}$$

where $*$ stands for the operation of convolution and $r(z_k)$ is defined as

$$r(z_k \mid \varphi(\mathbf{x})) = \sqrt{\frac{p_t(z_k)}{p(z_k \mid \varphi(\mathbf{x}))}}. \tag{4.19}$$

Finally, the results in (4.18) can be combined together to yield the first variation of $B(\varphi)$ as given by

$$\frac{\delta B(\varphi)}{\delta \varphi} = \delta(\varphi(\mathbf{x}))\, V_B(\mathbf{x}), \tag{4.20}$$

where

$$V_B(\mathbf{x}) = \frac{1}{2A} \sum_{k=1}^{d} \alpha_k \left(B_k(\varphi(\mathbf{x})) - \Big[r(z_k \mid \varphi(\mathbf{x})) * K(z_k) \Big]_{z_k = I_k(\mathbf{x})} \right), \tag{4.21}$$

with

$$\alpha_k = \prod_{i=1, i \neq k}^{d} B_i(\varphi(\mathbf{x})). \tag{4.22}$$

Hence, the gradient flow in (4.14) can be rewritten as

$$\frac{\partial \varphi(\mathbf{x}, \tau)}{\partial \tau} = \frac{\delta B(\varphi)}{\delta \varphi} = \delta(\varphi(\mathbf{x}, \tau)) V_B(\mathbf{x}, \tau) \tag{4.23}$$

where $\delta(\cdot)$ stands for the Dirac delta function, and $\delta(\varphi(\cdot))V_B(\cdot) : \Omega \to \mathbb{R}$ is a velocity field governing the evolution of the level-sets in $\varphi(\mathbf{x})$.

It should be noted that the estimation of φ^* has been so far performed based on regional features of the observed image u. This approach is useful in the cases when the object of interest does not possess well-identified edges at all [34]. However, as seen in section 4.2, TRUS images filtered by the SRAD algorithm do possess a few well-defined edges along the true prostate boundary. Therefore, edges in the pre-processed TRUS images still present an important source of information. Specifically, edge-related information can be incorporated via using the framework of *geodesic active contours* [35, 36]. In this case, the optimal level-set function φ^* is estimated according to

$$\varphi^\star(\mathbf{x}) = \arg\max_{\varphi\in\Phi} \left\{ \alpha\, B(\varphi) - \int_\Omega g(\mathbf{x})\, \|\nabla \mathcal{H}(\varphi(\mathbf{x}))\|\, d\mathbf{x} \right\}, \tag{4.24}$$

where $\alpha > 0$ is a regularizing constant, and $g : \Omega \to \mathbb{R}^+$ is an edge detector function, which will be defined below. As a result, the gradient flow associated with the optimization problem in (4.24) can be shown to be equal to

$$\frac{\partial \varphi(\cdot,\tau)}{\partial t} = \delta(\varphi) \left(\alpha\, V_B + \operatorname{div}\left(g\, \frac{\nabla\varphi}{\|\nabla\varphi\|} \right) \right) \tag{4.25}$$

Note that the V_B term in (4.25) attempts to find a region Ω_{in} in the image domain Ω whose empirical density $P(\mathbf{z} \mid \Omega_{in})$ closely matches the model density $P_t(\mathbf{z})$ in terms of the Bhattacharyya distance measure. The divergence term, on the other hand, attempts to match the boundary of Ω_{in}, while maintaining minimal curvature of the related active contour.

Finally, it should be noted that in practice, the numerical support of the level-set evolution is extended by replacing $\delta(\cdot)$ with a smoothed version, which is defined [34] as

$$\delta_\epsilon(x) = \begin{cases} \frac{1}{2\epsilon}\left(1 + \cos(\frac{\pi x}{\epsilon})\right), & |x| \le \epsilon \\ 0, & \text{otherwise} \end{cases} \tag{4.26}$$

Note that the support of δ_ϵ is finite and controlled through the user-defined parameter ϵ. This is set to be equal to 2 for all of the experiments carried out in the results section.

4.3.3 Motivating Example

In practical computations, the boundary Γ_{in} of Ω_{in} represents a decision boundary that separates the object of interest from its surroundings. This boundary, conventionally referred to as an *active contour*, is defined by the zero level-set of $\varphi(\mathbf{x})$, i.e. $\Gamma_{in} \triangleq \{\mathbf{x} \in \Omega \mid \varphi(\mathbf{x}) = 0\}$. In the formulation of the segmentation procedure thus far, the active contour belongs to the family of geodesic active contours. In particular, the gradient flow

defined in (4.25) only considers regional image features and edge related information. This is particularly useful, when the object of interest has well-defined edges as well as distinguishable features from the image background. Unfortunately, in TRUS imaging, detected edges are not considered to be reliable features. Due to the presence of shadowing artifacts and speckle noise, edges corresponding to the prostate boundary are often discontinuous or may be an artifact of the imaging modality. Consequently, using the segmentation tools dependent on regional image and edge information alone is prone to converge to erroneous results in the case of ultrasound imaging.

The above mentioned problem is exemplified in Fig. 4.3, which shows an original prostate image (Subplot A), its corresponding edge map $g(\mathbf{x})$ (Subplot B), and the final achieved segmentation (Subplot C) as obtained by (4.25).

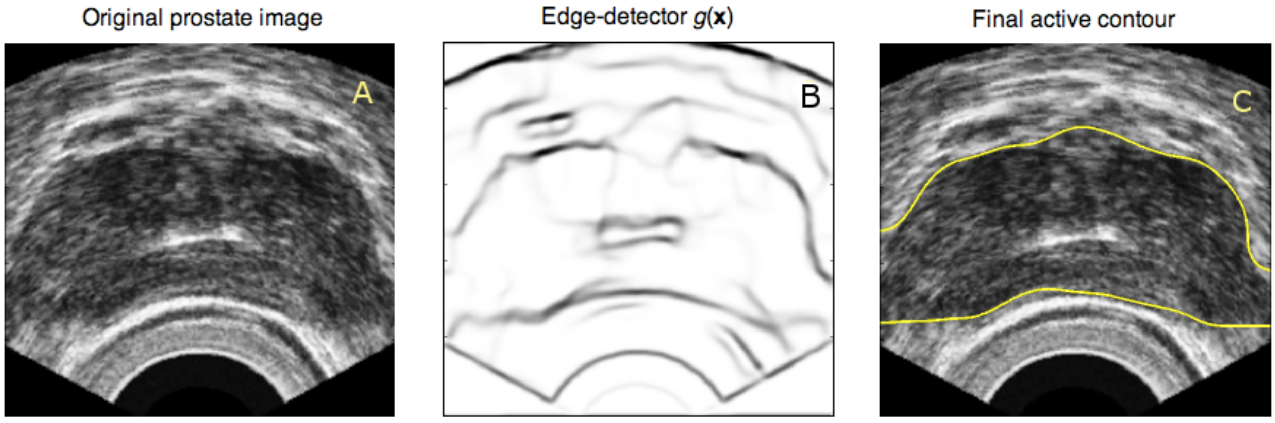

Figure 4.3: A segmentation example using only regional and edge information without incorporation of shape priors. (A) Original TRUS image. (B) Robust edge detector function $g(\mathbf{x})$. (C) Final segmentation without the use of shape priors.

Note that in this experiment, the edge detection function $g(\mathbf{x})$ is defined as

$$g(\mathbf{x}) = \frac{1}{1 + \lambda \left\| \nabla \tilde{u}(\mathbf{x}) \right\|^2}, \tag{4.27}$$

where $\tilde{u}$ stands for a de-speckled version of u as achieved by the SRAD filter [11], and λ is a positive constant set to be 3. Moreover, the Bhattacharyya function in (4.11) was defined using $I_1 = u$ and $I_2 = \tilde{u}$. From visual inspection, one can observe that the discontinuous prostate boundary along with shadowing artifacts have caused the active contour to "leak out" at the sides of the prostate region (Fig. 4.3 Subplot C). To alleviate

this problem, additional information on the expected geometry of the prostate boundary should be incorporated into the segmentation procedure, as explained in the following section.

4.3.4 Tracking of Shape Features using Weak Priors

In order to overcome the limitations presented in the above section, and to guarantee the convergence of the active contour to a useful segmentation, the gradient flow in (4.25) should be further constrained. A common solution to the aforementioned problem lies in the concept shape priors as exemplified by the works of [22, 23] as detailed in section 3.4. Another shape prior technique, which has shown promise, is described in the work of Tsai et al. [37], which was tested experimentally with MRI prostate data rather than TRUS image data. In their formulation, a shape model is derived by performing principle component analysis (PCA) on a training set of level-set functions. Subsequently, the level-set function representing the actual active contour is modelled as a linear combination of the principal components or *eigenshapes* [37], offset by the algebraic mean of the training functions. Specifically, the variable shape model is defined [37] as

$$\Phi[\mathbf{w}] = \bar{\Phi} + \sum_{i=1}^{k} w_i \Phi_i \tag{4.28}$$

where $\bar{\Phi}$ is level-set function corresponding to the mean training shape, $\Phi_i, i = 1 \dots k$ are the k different eigenshapes obtained by PCA, and $\mathbf{w} = \{w_1, w_2, \dots, w_k\}$ are the weights for the respective eigenshapes. This newly constructed level-set function Φ becomes an implicit shape model, as its zero level-set describes the shape boundary. Therefore, as the values of $\mathbf{w}$ vary, the implicit shape representation function Φ varies as well. However, the type and range of shape variability is directly related to the variability of the k eigenshapes, which also must depend on the variability of the training shapes. Thus, one of the main drawbacks of the PCA-based shape modelling methods lies in their dependence on the size of the training sets in use. In particular, small trainings sets can introduce significant estimation bias as they lack the ability to represent all possible appearances of the desired object. Note that, the above mentioned shape prior methods will, from now on, be considered as "strong" shape priors for the remainder of this thesis.

A conceptually different solution to the problem of shape priors is presented in the proposed algorithm. Specifically, a "weak" shape prior is described, which is proposed to *track* a plausible shape configuration in the manner similar to the framework for tracking of photometric feature distributions. Such "weak" shape priors are used as a minimally restrictive regularization force, that is let to be dominated by the feature driven segmentation forces. However, it will be shown that the proposed "weak" shape prior is still restrictive enough to guide the active contour towards a useful segmentation.

As a means to create a weak shape model based on the distribution tracking framework, one must first extract a morphological feature that will be used to characterize the boundary of the object of interest. To this end, let φ_t be the matching level-set function corresponding to a known Ω_t and let $\Gamma_t = \{\mathbf{x} \in \Omega \mid \varphi_t(\mathbf{x}) = 0\}$ be the boundary of Ω_t. In particular, the weak shape priors constructed in the proposed algorithm use *curvature* of Γ_t as the morphological feature of choice. It is chosen for the fact that curvature has the property of being invariant under the group of Euclidean transformations; this allows the algorithm to forgo a pre-registration step via rotations and translations before curvature values associated with the model and segmentation shape contour can be compared.

Given a set of manually delineated images with known prostate ground truth regions (i.e. known Ω_t), the fast marching method [33] is used to compute the corresponding level-set functions φ_t as signed distance functions of the segmentation boundaries Γ_t. In this case, the values of the curvature Γ_t can be obtained from the values of the curvature map κ_t of the level-sets belonging to φ_t as obtained [33] by

$$\kappa_t(\mathbf{x}) = -\text{div}\left\{\frac{\nabla\varphi_t(\mathbf{x})}{\|\nabla\varphi_t(\mathbf{x})\|}\right\}. \tag{4.29}$$

Thus, given a training set of segmented images with corresponding Γ_t, φ_t and κ_t, the *model* pdf of the object curvature can be estimated as given by

$$C_t(\eta) = \mathcal{E}\left\{\frac{\int_\Omega \delta_\epsilon(\varphi_t(\mathbf{x}))\, K(\eta - \kappa_t(\mathbf{x}))\, d\mathbf{x}}{\int_\Omega \delta_\epsilon(\varphi_t(\mathbf{x}))\, d\mathbf{x}}\right\}, \tag{4.30}$$

where $\mathcal{E}$ denotes the averaging operator as before, with K again denoting a positive, normalized kernel function, while κ_t and δ_ϵ are defined by (4.29) and (4.26) respectively. Note that the calculation in (4.29) provides the curvature values of *all* the level sets of φ_t. Therefore, $\delta_\epsilon(\varphi(\mathbf{x}))$ term in (4.30) restricts the estimated pdf to only depend on the values of κ_t in close proximity of the active contour (zero level-set). Finally, the target curvature distribution $C_t(\eta)$ is obtained via averaging the curvature pdf's over all available training boundaries.

Now, given an arbitrary level-set function φ, corresponding to Ω_{in} of an unknown TRUS image, the *empirical* probability density $C(\eta \mid \varphi)$ of the curvature of its associated active contour can be estimated according to

$$C(\eta \mid \varphi) = \frac{\int_\Omega \delta_\epsilon(\varphi(\mathbf{x}))\, K\left(\eta - \kappa(\mathbf{x})\right)\, d\mathbf{x}}{\int_\Omega \delta_\epsilon(\varphi(\mathbf{x}))\, d\mathbf{x}}. \tag{4.31}$$

where $\kappa(\mathbf{x})$ is the curvature map values associated with the arbitrary level-set function φ and obtained in the same way, using the equation (4.29).

Subsequently, the similarity between the model and empirical densities can again be assessed using the Bhattacharyya coefficient, which in this case is given by

$$B_c(\varphi) = \int \sqrt{C_t(\eta)\, C(\eta \mid \varphi)}\, d\eta, \tag{4.32}$$

where the subscript c has been added to distinguish the above curvature Bhattacharyya coefficient with the photometric one in (4.11).

Analogously to the tracking of photometric variables, it is proposed to track the empirical density $C(\eta \mid \varphi)$ by forcing it to be as similar as possible to the model density $C_t(\eta)$ through the maximization of $B_c(\varphi)$ in (4.32). The latter methodology forms the basis of the proposed "weak" shape priors. The maximization of $B_c(\varphi)$ can be achieved using the same tool of steepest ascent in the direction of the first variation of $B_c(\varphi)$ computed with respect to φ. The derivation of the first variation $\partial B_c(\varphi)/\partial\varphi$ is shown below.

Note that the derivation of the above first variation is presented under the assumption that $\|\nabla\varphi\| = 1$, which is valid as long as φ is initialized and sustained to be a signed distance function, which is the case in the present study. Under this assumption, one can show that the first variation of $C(\eta \mid \varphi)$ with respect to φ is given by

$$\frac{\delta C(\eta \mid \varphi)}{\delta\varphi} = \frac{\Delta\left[\delta_\epsilon(\varphi(\mathbf{x}))K'(\eta - \kappa(\mathbf{x}))\right] + \delta'_\epsilon(\varphi(\mathbf{x}))\left[K(\eta - \kappa(\mathbf{x})) - C(\eta \mid \varphi(\mathbf{x}))\right]}{\int \delta_\epsilon(\varphi(\mathbf{x}))d\mathbf{x}}. \tag{4.33}$$

Consequently, the first variation of $B_c(\varphi(\mathbf{x}))$ in (4.32) is defined as

$$\frac{\delta B_c(\varphi)}{\delta\varphi} = \frac{1}{2}\int \sqrt{\frac{C_t(\eta)}{C(\eta \mid \varphi(\mathbf{x}))}}\, \frac{\delta C(\eta \mid \varphi)}{\delta\varphi}\, d\eta. \tag{4.34}$$

Let $L(\eta) = \left[C_t(\eta)/C(\eta \mid \varphi(\mathbf{x}))\right]^{1/2}$ and $A = \int \delta_\epsilon(\mathbf{x})d\mathbf{x}$. Then $\delta B_c(\varphi)/\delta(\varphi)$ becomes

$$\begin{aligned}
V_c(\mathbf{x}) \triangleq \frac{\delta B_c(\varphi)}{\delta(\varphi)} &= \frac{1}{2}\int L(\eta)\, \frac{\delta C(\eta \mid \varphi(\mathbf{x}))}{\delta\varphi(\mathbf{x})} d\eta = \\
&= \frac{1}{2\,A}\int L(\eta)\, \Delta\left[\delta_\epsilon(\varphi(\mathbf{x}))\, K'(\eta - \kappa(\mathbf{x}))\right] d\eta + \\
&\qquad + \frac{\delta'_\epsilon(\varphi(\mathbf{x}))}{2\,A}\int L(\eta)\, \left[K(\eta - \kappa(\mathbf{x})) - C(\eta \mid \varphi(\mathbf{x}))\right] d\eta = \\
&= \frac{1}{2\,A}\Bigg[\int L(\eta)\, \Delta\left[\delta_\epsilon(\varphi(\mathbf{x}))\, K'(\eta - \kappa(\mathbf{x}))\right] d\eta + \\
&\qquad + \delta'_\epsilon(\varphi(\mathbf{x}))\Big(\left[L(\eta) * K(\eta)\right]_{\eta=\kappa(\mathbf{x})} - B_c(\varphi(\mathbf{x}))\Big)\Bigg],
\end{aligned} \tag{4.35}$$

where $*$ stands for the operation of convolution as before.

Subsequently, the optimization problem of (4.24) can be extended to incorporate the additional prior shape information, which results in a new optimal φ^* computed according to

$$\varphi^* = \arg\max_{\varphi\in\Phi}\left\{\alpha\, B(\varphi) + \beta\, B_c(\varphi) - \int_\Omega g(\mathbf{x})\,\|\nabla\mathcal{H}(\varphi(\mathbf{x}))\|\,d\mathbf{x}\right\}, \tag{4.36}$$

where Φ again denotes the set of signed distance functions (which can be identified with the set of solutions of the eikonal equation $\|\nabla\varphi\| = 1$, subject to $\varphi(\mathbf{x})\,|_{\mathbf{x}\in\partial\Omega_{in}} = 0$), α and β are regularization parameters which are set to be 0.5 and 2.5 respectively in the experimental studies. Finally, the gradient flow that corresponds to the maximization problem (4.36) is given by

$$\frac{\partial\varphi(\mathbf{x})}{\partial\tau} = \delta_\epsilon(\varphi(\mathbf{x}))\left(\alpha\, V_B(\mathbf{x}) + \mathrm{div}\left(g(\mathbf{x})\frac{\nabla\varphi(\mathbf{x})}{\|\nabla\varphi(\mathbf{x})\|}\right)\right) + \beta\, V_c(\mathbf{x}), \tag{4.37}$$

where the velocity $V_c(\mathbf{x}) : \Omega \to \mathbb{R}$ attempts to maximize $B_c(\varphi)$. Finally, it should be noted that the gradient flow in (4.37) is implemented using an implicit discretization scheme based on the method of additive operator splitting (AOS) as detailed in [38, Ch. 3 - 4].

4.3.5 Technical Considerations

Regularization of Curvature

In numerical computations, the values of the curvatures κ of the level-sets of φ are commonly computed according [33] to

$$\kappa(\mathbf{x}) = -\mathrm{div}\left\{\frac{\nabla\varphi(\mathbf{x})}{\|\nabla\varphi(\mathbf{x})\|}\right\} = \frac{2\varphi_x\varphi_y\varphi_{xy} - \varphi_{xx}\varphi_y^2 - \varphi_{yy}\varphi_x^2}{(\varphi_x^2+\varphi_y^2)^{3/2}}, \tag{4.38}$$

where the subscripts x and y denote partial differentiation along the respective directions. Since the partial derivatives in (4.38) are usually approximated by means of finite-support numerical schemes, the practical values of κ can be biased by both quantization and measurement noises. In this case, the discrete curvature values may be a noisy and "jagged" function. Needless to say, such behaviour of the curvature could bias the estimation of its corresponding curvature pdf. This situation can be observed in Fig. 4.4, where Subplot A shows a signed distance function corresponding to a yellow active contour. In addition, Subplot A1 shows a zoomed section of Subplot A (where the effect of quantization can be clearly observed), while Subplot A2 superimposes the circle over an image of its related curvature κ. It is obvious from visual inspection that the approximated values of κ suffer from spurious variations, which contradict the theoretical value of κ.

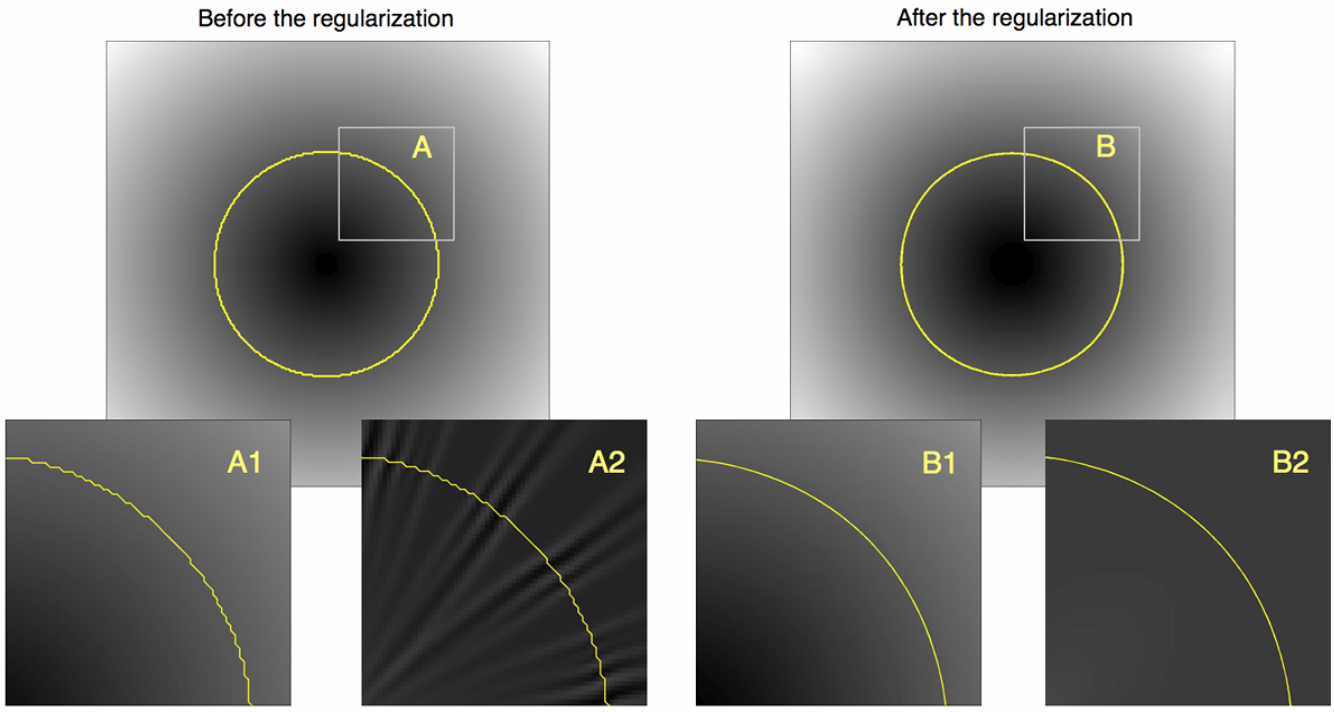

Figure 4.4: (Subplot A) A level-set function and its corresponding zero level-set before regularization. (Subplot A1) A zoomed section of the level-set function in Subplot A. (Subplot A2) The curvature of the level-sets in Subplot A1. (Subplot B) The regularized version of the level-set function in Subplot A. (Subplot B1) A zoomed portion of the regularized level-set function in Subplot B. (Subplot B2) The curvature of the level-sets in Subplot B1.

In order to alleviate the problem of irregularity of discrete κ, it was chosen to precede its computation by *regularizing* or smoothing of the corresponding φ, before φ is used to estimate the pdf of κ associated with the implicit active contour. Specifically, the pre-smoothing of φ can be achieved via an *anisotropic diffusion* defined by means of the following [39] equation

$$\frac{\partial \varphi(\mathbf{x},\tau)}{\partial \tau} = \operatorname{div}(D(\mathbf{x})\,\nabla\varphi(\mathbf{x},\tau)), \tag{4.39}$$

where $D(\mathbf{x}) \in \mathbb{R}^{2\times 2}$ denotes a diffusivity tensor at $\mathbf{x}$, and τ denotes an artificial diffusion (not to be confused with the one in (4.37)).

The smoothing effect of the diffusion (4.39) is controlled by the diffusion matrix $D(\mathbf{x})$, which should allow for smoothing to propagate along the direction *tangent* to the level-sets of φ. The diffusion in the tangential direction will preserve the shape of the active contour, while effectively suppressing the irregularities in φ due to numerical discretization schemes. In particular, the above goal can be achieved by setting $D(\mathbf{x}) = \gamma\vec{v_1}(\mathbf{x})\vec{v_1}(\mathbf{x})^T + \vec{v_2}(\mathbf{x})\vec{v_2}(\mathbf{x})^T$, with $\gamma \ll 1$, and with $\vec{v_1}(\mathbf{x})$ and $\vec{v_2}(\mathbf{x})$ being two unit vectors pointing in the directions parallel and perpendicular to $\nabla\varphi(\mathbf{x})$, respectively. Namely, $\vec{v_1}(\mathbf{x}) \parallel \nabla\varphi(\mathbf{x})$ and $\vec{v_2}(\mathbf{x}) \perp \nabla\varphi(\mathbf{x})$ at every $\mathbf{x} \in \Omega$.

In the experiments reported in the later section, the diffusion in (4.39) is performed using the AOS scheme detailed in [39], with $\gamma = 0.01$, and the number of iterations and diffusion time-step being equal to 4 and $\Delta\tau = 5$, respectively. The advantage of the proposed pre-smoothing procedure is shown in Subplots B1 and B2 of Fig. 4.4, which shows a regularized version of the level-set function of Subplot A and its corresponding smoothed curvature. One can see that the regularization of the discrete values of κ better complies with its theoretically predicted value.

Algorithm Summary

The overall structure of the proposed method for segmentation of TRUS images is summarized in the pseudocode of Algorithm 1. The algorithm receives as its input a TRUS image $u(\mathbf{x})$, the target pdf $P_t(\mathbf{z})$ of a set of photometric features, the target pdf $C_t(\eta)$ of the curvature of prostate boundary, as well as an initial level-set function $\varphi_0(\mathbf{x})$. At the output, the algorithm returns a final level-set function $\varphi_\tau(\mathbf{x})$, which can be used to indicate the prostate region as $\{\mathbf{x} \mid \varphi_\tau(\mathbf{x}) \le 0\}$, or by means of its boundary $\{\mathbf{x} \mid \varphi_\tau(\mathbf{x}) = 0\}$. Note that at Step 10 of Algorithm 1 below, the operation denoted by "AOS" refers to the procedure represented by equation (51) in [40] (the equation is provided below, please see the paper for additional details).

$$u^{k+1} = \frac{1}{m}\sum_{l=1}^{m} \left(I - m\tau A_l(u^k)\right)^{-1} u^k. \tag{4.40}$$

Algorithm 1 Proposed segmentation procedure

given: $u(\mathbf{x})$, $P_t(\mathbf{z})$, $C_t(\eta)$, $\varphi_0(\mathbf{x}) \equiv \varphi(\mathbf{x}, \tau = 0)$
preset: $\Delta\tau = 10$, $\alpha = 0.5$, $\beta = 2.5$, $\tau = 0$
compute: $\{I_k(\mathbf{x})\}_{k=1}^{d}$ and $g(\mathbf{x})$ using (4.27) (e.g. with $\lambda = 3$)
while $\delta > 10^{-4}$ **do**
    Diffuse $\varphi_\tau(\mathbf{x})$ using (4.39) to result in $\tilde{\varphi}_\tau(\mathbf{x})$
    Compute $\kappa(\mathbf{x})$ using $\tilde{\varphi}_\tau(\mathbf{x})$ and equation (4.38)
    Compute $\{p(z_k \mid \varphi_\tau(\mathbf{x}))\}_{k=1}^{d}$ using (4.12) and $C(\eta \mid \varphi_\tau(\mathbf{x}))$ using (4.31)
    Compute $V_B(\mathbf{x})$ using (4.21) and $V_c(\mathbf{x})$ using (4.35)
    $\varphi_{\tau+1}(\mathbf{x}) \Leftarrow \varphi_\tau(\mathbf{x}) + \Delta\tau\ (\alpha\, V_B(\mathbf{x}) + \beta\, V_c(\mathbf{x}))$
    $\varphi_{\tau+1}(\mathbf{x}) \Leftarrow \mathrm{AOS}\,(\varphi_{\tau+1}(\mathbf{x}), g(\mathbf{x}), \Delta t)$
    Redistance $\varphi_{\tau+1}(\mathbf{x})$ by fast marching [33]
    $\delta \Leftarrow \|\varphi_{\tau+1} - \varphi_\tau\|$
    $\tau \Leftarrow \tau + 1$
end while
return $\varphi_\tau(\mathbf{x})$

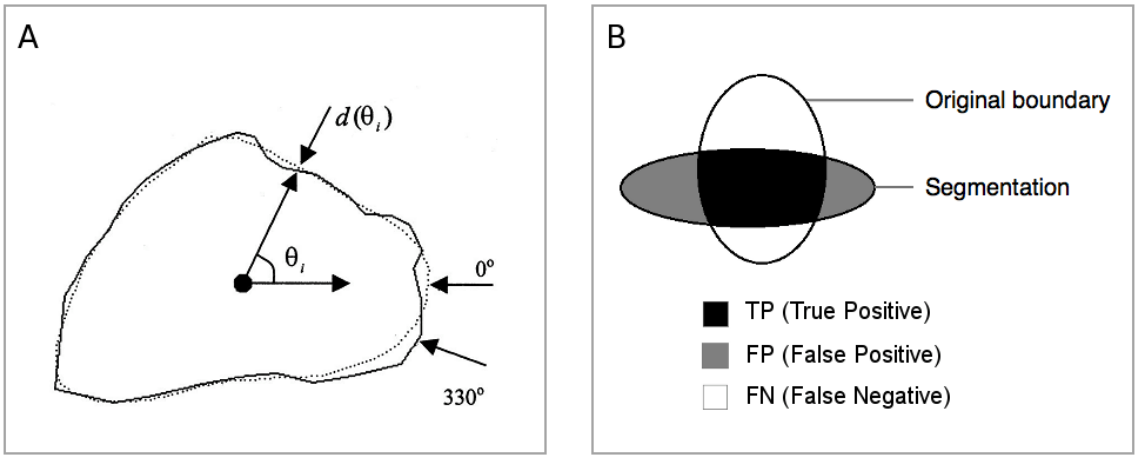

Figure 4.5: (Subplot A) Local difference $d(\theta_i)$ between the original and optimal boundaries; (Subplot B) Different regions within the original and optimal segmentations [8].

Chapter 5

Experimental Results

5.1 Segmentation Evaluation Metrics

In this section, to quantitatively assess the performance of the proposed segmentation method, a number of evaluation metrics are used. Specifically, given the true boundary of a desired object and its estimate, both contours can be sampled at the points defined by their intersection with L rays extending from the common center of the contours (see Subplot A of Fig. 4.5 for an example). Subsequently, along the directions of the rays (defined by angles $\theta_i = 2\pi i/L$, with $i = 0, 1, \ldots, L$), the signed distances $\{d(\theta_i)\}_{i=1}^{L}$ between the original and estimated points can be calculated and used to compute mean difference (MD), mean absolute difference (MAD), and maximum difference (MAXD) defined [8] as

$$\mathrm{MD} = \frac{1}{L}\sum_{i=1}^{L} d(\theta_i), \tag{5.1a}$$

$$\mathrm{MAD} = \frac{1}{L}\sum_{i=1}^{L} |d(\theta_i)|, \tag{5.1b}$$

$$\mathrm{MAXD} = \max_{1\le i\le L}\{|d(\theta_i)|\}. \tag{5.1c}$$

Whereas MD provides useful information on the size and direction of the segmentation bias, MAD and MAXD quantify the overall segmentation error in terms of the average and maximum value of $\{|d(\theta_i)|\}_{i=1}^{L}$, respectively.

Additional performance measures can be derived from area-based metrics, which (as opposed to MD, MAD, and MAXD) are insensitive to the local geometry of segmentation boundaries. In order to compute the area-based metrics, one must first define the *true*

positive, *false positive*, and *false negative* subsets of Ω as shown in Subplot B of Fig. 4.5, with the corresponding areas denoted by TP, FP, and FN, respectively. Then, one can define three area-based metrics as given by [8]

$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{5.2a}$$

$$\text{ACC} = 1 - \frac{\text{FP} + \text{FN}}{\text{TP} + \text{FN}}, \tag{5.2b}$$

$$\text{AO} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}, \tag{5.2c}$$

where TP + FN being equal to the area of the ground truth region, and SEN, ACC, and AO represents sensitivity, accuracy, and area-overlap of the actual segmentation respectively. In general, the higher the values of SEN, ACC, and AO, the better the performance of an image segmentation under consideration.

Finally, the segmentation accuracy can also be assessed in terms of the standard *mean squared error* (MSE) criterion. Specifically, let F be the $M \times N$ matrix of the true (binary) segmentation mask, and $\tilde{F}$ be its estimate. Then, the MSE is defined as

$$\text{MSE} = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left| \tilde{F}(x,y) - F(x,y) \right|^2. \tag{5.3}$$

5.2 *In-silico* Experiments

In this section, the performance of the proposed algorithm is validated through a series of *in silico* experiments. In particular, 2-D prostate shapes were synthetically generated to mimic the typical geometry of a prostate gland with the addition of noise and shadowing artifacts to simulate the challenges of a typical TRUS image. Simulated prostate images of different levels of contrast are illustrated in Fig. 5.1.

In the present study, 200 simulated prostate images were generated for each contrast category with half of them being used as training images and the rest used for algorithm validation. In particular, the training images, together with their SRAD filtered versions were used to pre-compute the target pdf's $P_t(\mathbf{z})$ and $C_t(\eta)$. The test images, on the other hand were used for segmentation, followed by a quantitative analysis of its performance.

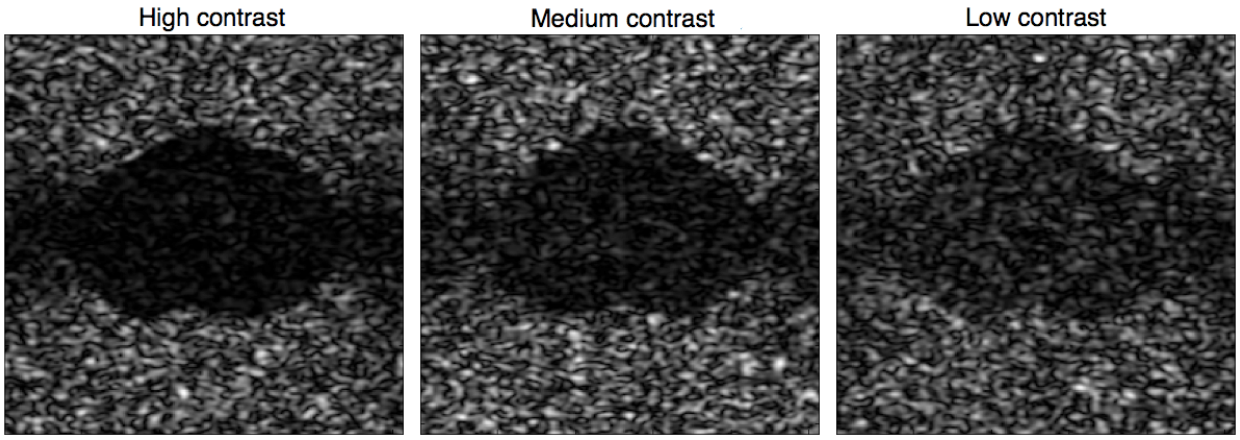

Figure 5.1: Simulated prostate images with varying levels of contrast and corrupted by noise and shadowing artifacts.

A subset of typical segmentation results are shown in Fig. 5.2. In particular, the upper row of subplots of the figure depicts the ground truth regions of the simulated prostate shapes, while the subplots in the second row show the corresponding simulated prostate images. In the first trial run, these images were segmented by Algorithm 1 without the use of "weak" shape priors (i.e. $\beta = 0$). The results of this segmentation are demonstrated by the third row of subplots in Fig. 5.2. It can be observed that the discontinuity of prostate boundaries along with the low contrast of shadowed regions have caused the segmentation to converge to an erroneous result. To validate the feasibility of the "weak" shape prior algorithm, the same images were segmented by Algorithm 1 with the inclusion of shape priors (i.e. $\beta = 2.5$). The resulting segmentation is shown in the bottom row of subplots of Fig. 5.2. It can be seen that the segmentation is regularized enough to converge to reasonable configurations.

Note that since the synthetic experiments are only used to validate the feasibility of using "weak" shape priors and the fact that image features are not representative of actual TRUS images, the only feature used in this set of experiments is the original grey-level intensities of the simulated images.

The mean segmentation results from the 100 testing images are summarized in Table 5.1, which shows that the inclusion of shape priors results in significant increase in the segmentation accuracy for most of the evaluation metrics presented (as expected). In case of the SEN metric, it can be biased towards segmentation without shape priors because it is only concerned about the percentage of the TP area (from the segmentation estimate) with respect to the area of the ground truth region. Hence, the over-estimation caused by the segmentation without shape priors sometimes produces a better SEN metric.

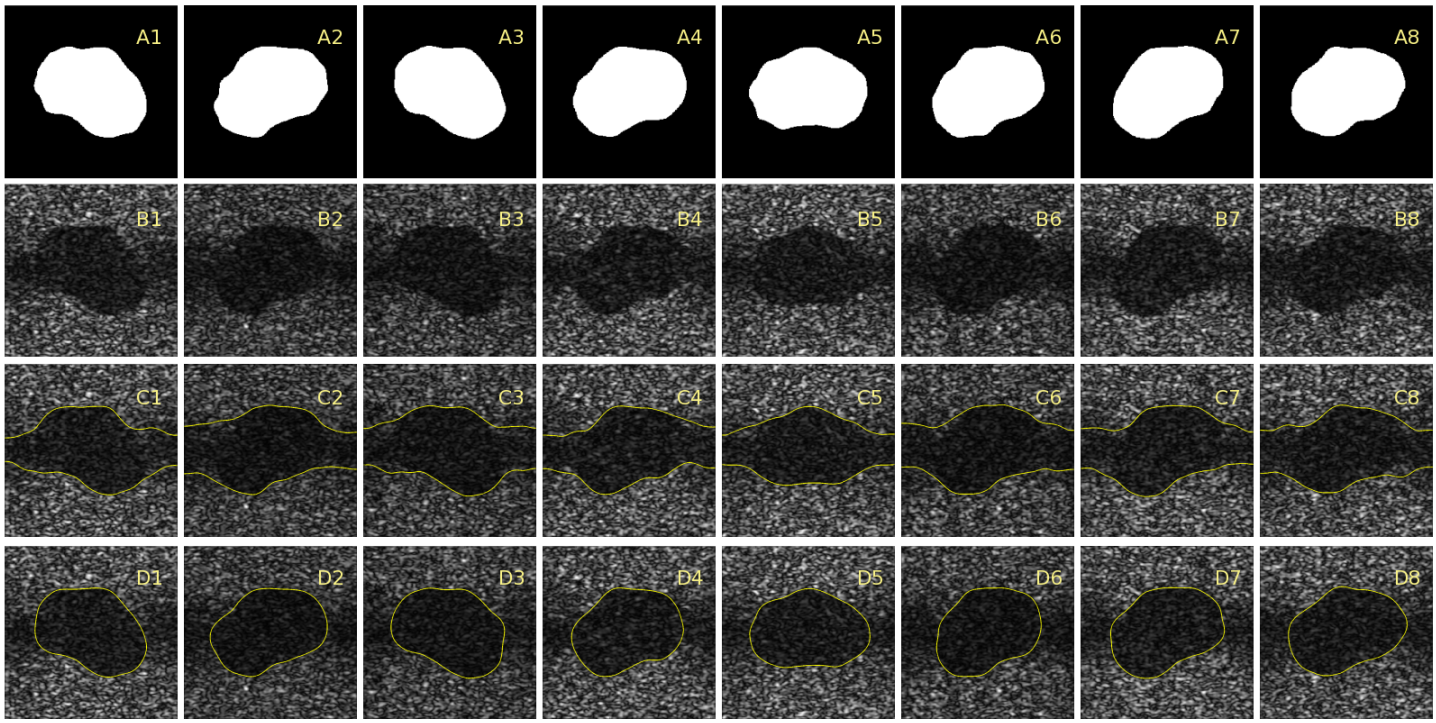

Figure 5.2: (First row of subplots) True segmentation profiles used for simulation of prostate images. (Second row of subplots) The corresponding prostate images (medium contrast) corrupted by shadowing artifacts. (Third row of subplots) Segmentation results obtained by Algorithm 1 with $\beta = 0$ (i.e. without shape priors). (Fourth row of subplots) Segmentation results obtained by Algorithm 1 incorporating weak shape priors.

Table 5.1: Segmentation Results for Synthetic Prostate Data

	High Contrast		Medium Contrast		Low Contrast	
Method	Weak Priors	No Priors	Weak Priors	No Priors	Weak Priors	No Priors
MD	**0.45**	−5.70	**0.642**	−8.242	**0.994**	-11.3
MAD	**2.31**	8.50	**1.80**	9.90	**3.42**	11.9
MAXD	**10**	55	**8**	56	**24**	59
SEN	**0.9626**	0.9625	0.964	**0.978**	0.929	**0.992**
ACC	**0.932**	0.633	**0.947**	0.588	**0.905**	0.536
AO	**0.934**	0.724	**0.948**	0.704	**0.907**	0.681
MSE	**0.0170**	0.0918	**0.0132**	0.123	**0.0238**	0.116

5.3 *In-vivo* Experiments

In addition to the *in silico* experiments, *in vivo* experiments were also conducted to further test the performance of the proposed algorithm. In the *in vivo* experiments, the ultrasound images were obtained during clinical TRUS sessions using the Aloka 2000 ultrasound machine with a broadband 7 MHz sector array and a field of view of approximately 6 cm. Subsequently, the obtained set of clinical TRUS images were manually delineated by an expert radiologist.

Aside from providing the ground truth for a comparative analysis, the manually delineated images were also used to learn the target pdf's $P_t(\mathbf{z})$ and $C_t(\eta)$. While $C_t(\eta)$ is learned based on the curvature values of the known prostate boundary, photometric features used to construct the model pdf $P_t(\mathbf{z})$ are described in the section below.

5.3.1 Photometric Feature Selection

Since the proposed algorithm is mainly driven by the photometric features in the TRUS images, the proper choice of such features is of great importance to the segmentation success. To this end, a number of prevalent image features including grey-level intensities, SRAD filter coefficients, Laws' texture images, and Gabor filtered images were considered in the construction of the model $P_t(\mathbf{z})$. Specifically, it would be desirable to select features, which show a high degree of dissimilarity between the features corresponding to the prostate and background regions respectively. Moreover, in the current segmentation framework, the feature pdf's corresponding to the prostate region should be maximally *dissimilar* to the feature pdf's of the background region as indicated by a *low* value Bhattacharyya coefficient as defined by

$$B_k = \int \sqrt{p_k(z_k \mid \Omega_p) p_k(z_k \mid \Omega_b)} dz_k, \tag{5.4}$$

where the subscript k denotes the kth feature, while Ω_p and Ω_b denotes the regions supporting the prostate and background respectively. Examples of feature selection evaluations are illustrated in Fig. 5.3. In each of the Subplots (a)-(d), the feature image was separated into the ground truth prostate and background regions as depicted in the top row. Then, the pdf distributions were calculated for the respective regions of interest and displayed underneath the corresponding prostate and background image regions.

Upon visual inspection of the example features illustrated in Fig. 5.3, it can be observed that the pdf's corresponding to grey-level intensities, SRAD filtered images, and Gabor filtered images showed a great level of dissimilarity between the prostate and background regions. Conversely, the Laws texture energy image in Fig. 5.3 Subplot (d) showed highly

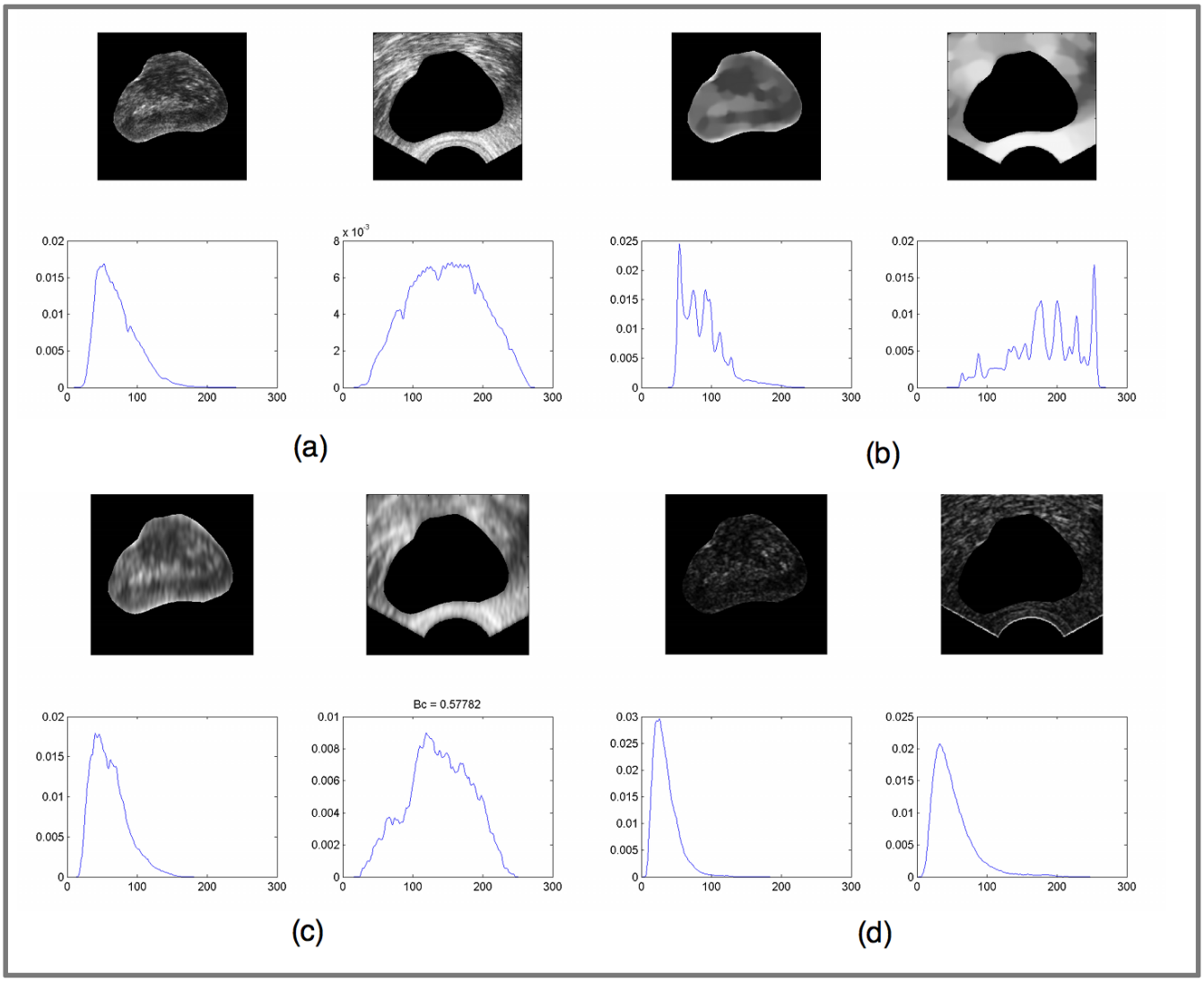

Figure 5.3: Examples of feature images and the corresponding feature value pdf's with respect to both the prostate and background regions are presented: (a) original grey-level intensities, (b) SRAD filtered image, (c) Gabor filtered image, (d) Laws' textured energy image.

similar distributions for the regions of interest. Although the pdf's of the Gabor filtered images does show a certain level of dissimilarity, their overall distribution shapes appear to be quite similar to the pdf's corresponding to original grey-level intensities. Ultimately, after a series of experimental segmentation tests, it was decided to use the original grey-level intensities and SRAD filtered versions as the photometric features for constructing $P_t(\mathbf{z})$, as they provided the best segmentation performance from the list of considered features.

5.3.2 Segmentation Process

In the *in vivo* experiments, a typical segmentation process is demonstrated in Fig. 5.4. Starting at the left most column and transitioning to the right, each column represents the results of initial, intermediate, and final segmentation stages. In particular, the upper row of subplots depicts the convergence of the active contour from its initial to final configuration, while Subplots B1-B3 and Subplots C1-C3 compare the target (dashed) and empirical (solid) pdf's. Specifically, the pdf's corresponding to the intensity values of the TRUS images (as observed within the boundaries of the active contour) are plotted in B1-B3. At the same time, Subplots C1-C3 shows the evolution of the curvature pdf's. As the segmentation converges, the appearances of the target and empirical distributions progressively becomes more and more similar.

Note that the segmented TRUS image in 5.4 is the same image presented in 4.3. In the previous case, the prostate was unsuccessfully segmented using only edge and regional based features of the prostate (i.e. the gradient flow equation in 4.37 was used with $\beta = 0$). The occurrence of this result is due to the presence of shadowing artifacts, which presents two significant challenges. First of which is the lack of well-defined edges near the left and right sides of the prostate. Second challenge is that the regional image features pertaining to the prostate sides and its immediate surroundings are hardly distinguishable. On the other hand, the proposed "weak" shape priors sufficiently regularizes the segmentation to a more reasonable result.

Additional segmentation examples are demonstrated in Fig. 5.5, whose Subplots A1-A3 show the results of three different manual segmentations performed by an expert. Note that the shown images represent the principal challenges of TRUS image segmentation, as they suffer from both severe shadowing artifacts and low contrast. Moreover, the boundaries of the shown prostates are discontinuous due to the aforementioned imaging artifacts. Under these conditions, the segmentation produced by Algorithm 1 without the use of shape priors (i.e. $\beta = 0$) "leaks out" near the shadow regions of the prostate boundary, as demonstrated in Subplots B1-B3. Considerably better segmentation results are produced by the method of [41] and depicted in C1-C3, which uses a strong shape prior obtained via applying PCA

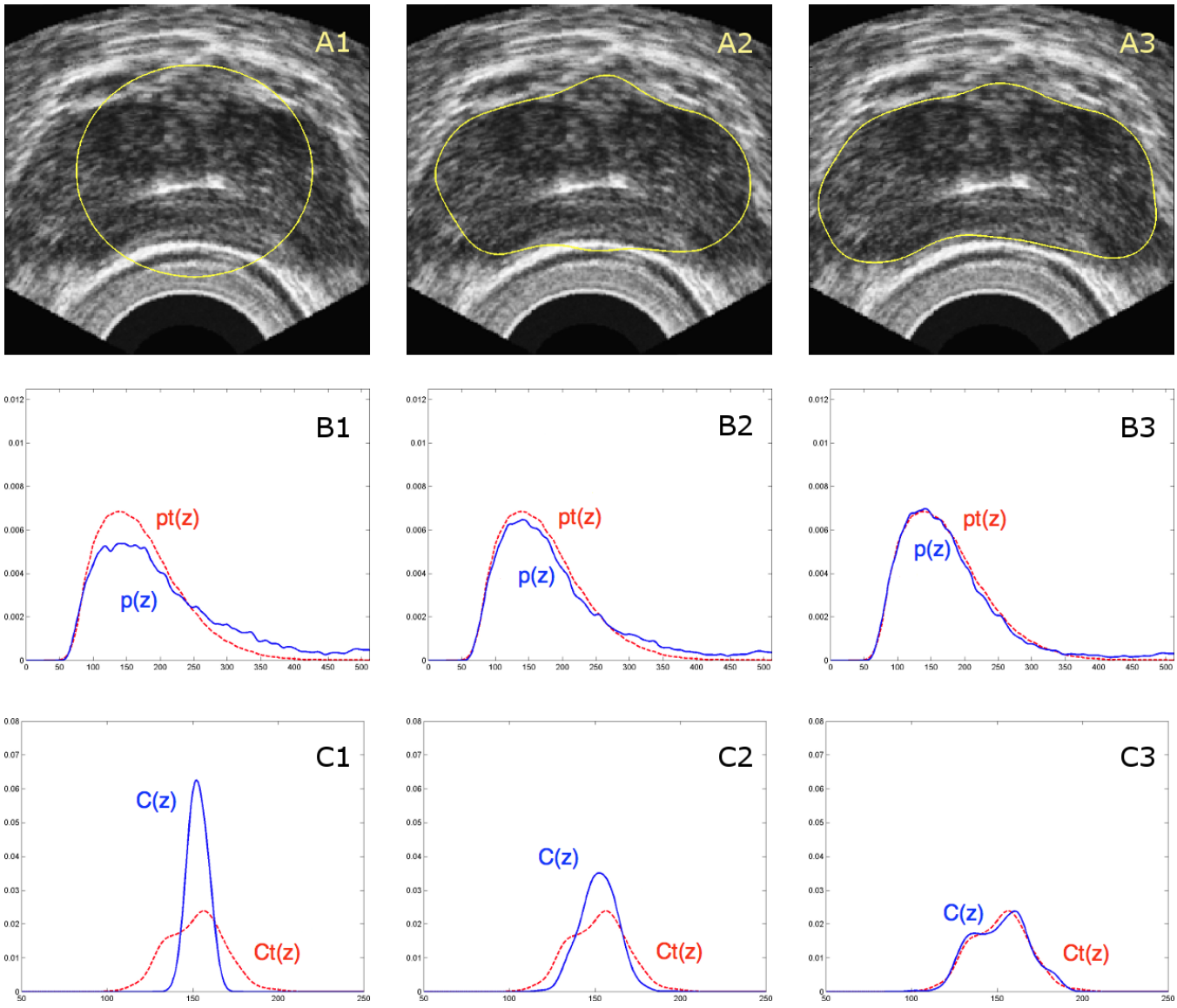

Figure 5.4: (Subplots A1-A3) Initial, intermediate, and final segmentations of a prostate image. (Subplots B1-B3) The target (dashed) and empirical (solid) pdf's of the image grey-levels corresponding to the segmented regions shown in Subplots A1-A3, respectively. (Subplots C1-C3) The target (dashed) and empirical (solid) pdf's of the curvature of the active contours shown in Subplots A1-A3, respectively.

on a set of training shapes. Finally, the segmentation results obtained from the proposed segmentation method are illustrated in Subplots D1-D3 of Fig. 5.5.

The method of [41] and the proposed algorithm are compared because both strategies use a distribution tracking approach towards the problem of segmentation. In particular, the same photometric distribution tracking framework is adopted by both algorithms, while the only difference lies in the choice of shape priors. In the work presented by Freedman et al [41], the chosen shape prior is obtained via applying PCA to a set of training shapes, and the subsequent active contour is represented by the linear combination of a set of eigenshapes and the mean shape. In contrast, the shape model used in the proposed algorithm is the "weak" shape prior as described in the previous section 4.3.4.

A quantitative comparison of the segmentation performance for the method described by Freedman et al. [41] and the proposed algorithm are summarized in Table 5.2. Specifically, given a set of six manually segmented TRUS images, five are used for training and the one remaining image is used for testing. Thus, six such leave-one-out tests are performed using both algorithms and the mean segmentation evaluation metrics are presented in Table 5.2. It can be observed that the segmentation results of the proposed weak shape priors are comparable to the strong priors [41] in most categories, and slightly favourable in certain categories. Thus, the concept of weak shape priors has shown some promise in terms of both visual inspection as well as quantitative performance analysis.

Table 5.2: Segmentation Results for *In-vivo* Prostate Data

Method	Weak Priors $\pm\sigma$	Strong Priors [41] $\pm\sigma$
MD	4.60 ± 1.96	$\mathbf{1.55} \pm 2.70$
MAD	$\mathbf{5.39} \pm 1.51$	6.53 ± 2.10
MAXD	$\mathbf{7.12} \pm 0.708$	9.87 ± 0.849
SEN	$\mathbf{0.939} \pm 0.032$	0.902 ± 0.039
ACC	$\mathbf{0.873} \pm 0.029$	0.850 ± 0.052
AO	$\mathbf{0.881} \pm 0.027$	0.858 ± 0.047
MSE	$\mathbf{0.031} \pm 0.008$	0.0527 ± 0.019

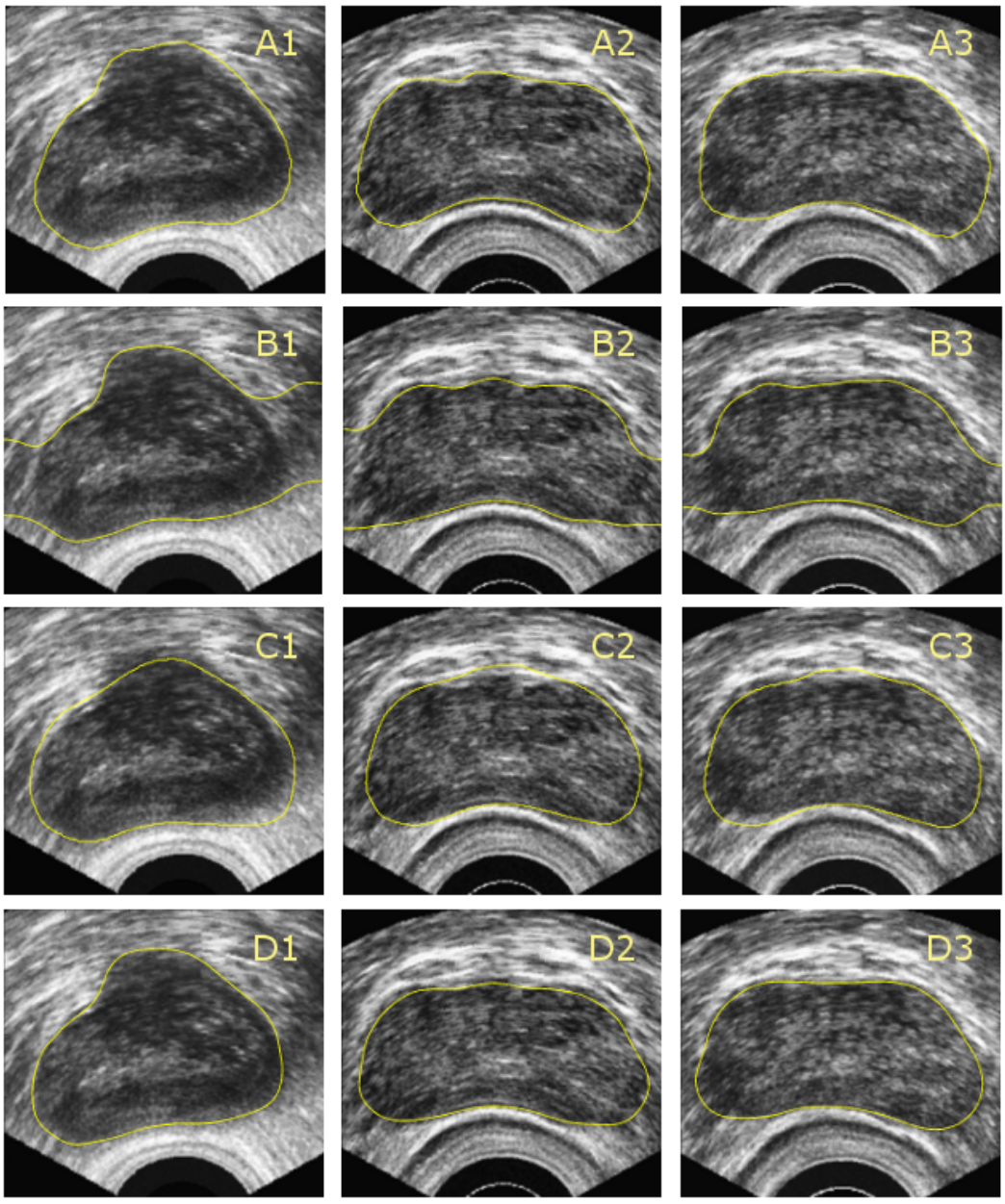

Figure 5.5: Examples of *in vivo* image segmentation: (Subplots A1-A3) Expert manual delineations. (Subplots B1-B3) Results produced by Algorithm 1 without the use of shape priors ($\beta = 0$). (Subplots C1-C3) Results produced using the strong shape priors [41]. (Subplots D1-D3) Results produced by Algorithm 1 ($\beta = 2.5$).

Chapter 6

Discussions and Conclusions

An alternative approach to segmentation of TRUS images of prostates was presented in this study. The method can be categorized as a learning-based technique, since it takes advantage of the availability of manually delineated TRUS data. Moreover, the proposed algorithm can be seen as a natural extension to the distribution tracking approach of [25], in which the segmentation process is performed via minimizing the distance between empirical and model pdf's of image features. Learning the model densities is based on training samples, which consists of a set of representative images and their corresponding ground truth regions. Based on the latter ground truth regions, a set of photometric distributions can be learned and constructed to be the model pdf to be used in the subsequent segmentation procedure. Moreover, the outer boundary of the ground truth regions can also be used to estimate the pdf of geometric parameters associated with the overall prostate structure. This allows the approach of [25] to be extended to tracking the "morphological distributions", in which case the optimal active contour during the segmentation process is required to minimize the distance between the model pdf of the chosen geometric parameter (curvature in this case) and its empirical counterpart. In the current study, the Bhattacharyya coefficient has been employed to assess the distances between the model and empirical distributions for both photometric and morphological features.

It should be noted that given a pdf of a geometric parameter such as curvature, there exists a large number of possible shapes whose boundaries would have their empirical pdf's be identical or almost identical to a model curvature distribution. Thus, the shape model has been considered "weak" in the sense that it is minimally restrictive in nature, and is only considered as a regularization force rather than the dominant driving force in the segmentation procedure. Nevertheless, it was shown in a number of numerical experiments that the "weak" priors are capable of helping the convergence of the active contour to a useful segmentation. This is promising, because in contrary to the "strong" shape priors presented in [22,23,41], the "weak" shape prior does not force the segmentation to converge

to a set of previously observed shapes. Its definition naturally leads to a generalization to a family of shapes, which share common attributes, but are not strictly defined.

Another advantage associated with the definition of the "weak" shape priors using curvatures lies in its invariance under the group of Euclidean transformations. In most shape prior methods, the images comprising a given training set are normally registered to a common position, orientation, and scale before the specific prostate shape model can be learned. In the case of an unknown TRUS image, the prostate is in general unaligned with the pose of the defined shape model. Thus, many segmentation algorithms are bound to use an image registration procedure as a means to bring an evolving active contour into correspondence with its training model before the discrepancy between them can be assessed. Since the registration task is yet another optimization problem, it introduces an extra layer of difficulty to the problem of segmentation. Fortunately, the use of curvature values in the modelling of "weak" shape priors eliminates the above problem, as the geometric descriptor is invariant under the group of Euclidean transformations by definition.

Although the proposed method has shown some promise when applied towards the problem of segmentation of TRUS images, there are many possible improvements or extensions to the current technique. For example, the use of curvature, while being invariant under Euclidean transformations, would not be applicable when more complex (e.g. affine) deformations are expected. In this case, other geometric invariant descriptors should be considered. Furthermore, the current choice of photometric features consisting of grey-level intensities and SRAD filtered coefficients are sufficient, but perhaps more descriptive texture features exist. Moreover, the natural and important extension of the proposed method to 3-D TRUS data as well as its verification via more extensive clinical data is well warranted. Conceptually, the same distribution tracking framework of both photometric and morphological features can be used in the case of 3-D data. However, it will require the implementation of 3-D active contours as well as the utilization of 3-D shape descriptors such as Gaussian curvature.

In conclusion, a novel segmentation technique using "weak" shape priors has been presented. Conceptually, it has been validated in a number of numerical experiments, and its performance has shown promise when compared to a prevalent PCA-based shape modelling technique. However, there are still many possible enhancements that could be applied to the existing algorithm, and the previously described extensions present a main direction for future work.

Bibliography

[1] U. of Pittsburgh Medical Centre. (2009, Mar.) Prostate cancer: Transrectal ultrasound. [Online]. Available: http://www.upmccancercenters.com/cancer/prostate/biopsyultrasound.html 1, 13

[2] D. M. Parkin, F. Bray, J. Ferlay, and P. Pisani, "Global cancer statistics 2002," *CA Cancer J. Clin.*, vol. 55, no. 2, pp. 74–108, Mar. 2005. 1

[3] C. Guy and D. ffytche, *An Introduction to the Principles of Medical Imaging.* Covent Garden, London: Imperial College Press, 2005. 4, 5, 6, 7, 8, 9, 10, 11, 12

[4] S. D. Pathak, V. Chalana, D. R. Haynor, and Y. Kim, "Edge-guided boundary delineation in prostate ultrasound images," *IEEE Transactions on Medical Imaging*, vol. 19, no. 12, pp. 1211–1219, Dec. 2000. 14, 15, 16, 17, 26, 44, 57

[5] R. N. Czerwinski, D. L. Jones, and W. D. O. Jr., "Detection of lines and boundaries in speckle images-Application to medical ultrasound," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 126–136, 1999. 15

[6] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990. 16

[7] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 679–6989, 1986. 17

[8] H. M. Ladak, F. Mao, Y. Wang, D. B. Downey, D. A. Steinman, and A. Fenster, "Prostate boundary segmentation from 2D ultrasound images," *Med. Phys.*, vol. 27, pp. 1777–1788, 2000. 17, 18, 19, 20, 21, 26, 73, 74, 75

[9] S. Lobregt and M. A. Viergever, "A discrete dynamic contour model," *IEEE Transactions on Medical Imaging*, vol. 14, no. 1, pp. 12–24, Mar. 1995. 17

[10] Y. Yu, J. A. Molloy, and S. T. Acton, "Segmentation of the prostate from suprapubic ultrasound images," *Medical Physics*, vol. 31, no. 12, pp. 3474–3484, Dec. 2004. 22, 23, 24, 25, 26

[11] Y. Yu and S. Acton, "Speckle reducing anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 11, no. 11, pp. 1260–1270, Nov. 2002. 25, 26, 57, 58, 66

[12] J. S. Prater and W. D. Richard, "Segmenting ultrasound images of the prostate using neural networks," *Ultrasonic Imaging*, vol. 14, no. 2, pp. 186–211, 1992. 27, 28, 29, 30, 32, 41

[13] E. Barnard and R. Cole, *A Neural-Net Training Program Based on Conjugate Gradient Optimization*, 1989. 28

[14] W. D. Richard and C. G. Keen, "Automated texture-based segmentation of ultrasound images of the prostate," *Comput. Med. Imag. Graph.*, vol. 20, no. 3, pp. 131–140, May 1996. 31, 33, 35, 36, 37, 41, 42

[15] K. I. Laws, "Textured image segmentation," Ph.D. dissertation, Univ. of Southern California, Los Angeles, CA, 1980. 31, 33, 42, 60

[16] J. Y. Hsiao and A. Sawchuk, "Unsupervised textured image segmentation using feature smoothing probabilistic relaxation techniques," *Comput. Vision Graph. Image Process.*, vol. 48, no. 1, pp. 1–21, 1989. 31, 33

[17] A. Zaim, T. Yi, and R. Keck, "Feature-based classification of prostate ultrasound images using multiwavelet and kernel support vector machines," in *International Joint Conference on Neural Networks*, Orlando, Florida, Aug. 2007. 37, 38, 39, 40, 41

[18] H. W. Lee, J. W. Kim, C. S. Kim, and S. J. Kim, "Limited rotation-invariant pattern recognition using optical wavelet circular harmonic function mace filter," in *Proceedings of SPIE*, vol. 2491, 1995, pp. 63–71. 37

[19] O. Rashkovskiy, L. Sadovnik, and N. Caviris, "Scale, rotation and shift invariant wavelet transform," in *Proceedings of SPIE*, vol. 2237, 1994, pp. 390–401. 37

[20] J. S. Geronimo, D. P. Hardin, and P. R. Massopust, "Fractal functions and wavelet expansions based on several scaling functions," *J. Approx. Theory*, vol. 78, pp. 373–401, Sept. 1994. 38

[21] K. I. Kim, K. Jung, S. H. Park, and H. J. Kim, "Support vector machines for texture classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 11, pp. 1542–1550, 2002. 40

[22] L. Gong, S. D. Pathak, D. R. Haynor, P. Cho, and Y. Kim, "Parametric shape modeling using deformable superellipses for prostate segmentation," *IEEE Transactions on Medical Imaging*, vol. 23, no. 3, pp. 340–349, Mar. 2004. 42, 43, 44, 45, 46, 55, 67, 84

[23] D. Shen, Y. Zhan, and C. Davatzikos, "Segmentation of prostate boundaries from ultrasound images using statistical shape model," *IEEE Transactions on Medical Imaging*, vol. 22, no. 4, pp. 539–551, Apr. 2003. 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 67, 84

[24] B. Manjunath and W. Ma, "Texture features for browsing and retrieval of image data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 837–842, August 1996. 49

[25] D. Freedman and T. Zhang, "Active contours for tracking distributions," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 518–526, Apr. 2004. 56, 57, 62, 84

[26] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Transactions on Communication Technology*, vol. COM-15, no. 1, pp. 52–60, Feb. 1967. 57, 62

[27] F. Goudail, P. Refregier, and G. Delyon, "Bhattacharyya distance as a contrast parameter for statistical processing of noisy optical images," *J. Opt. Soc. Am. A*, vol. 21, no. 7, pp. 1231–1240, July 2004. 57, 62

[28] O. Michailovich, Y. Rathi, and A. Tannenbaum, "Image segmentation using active-contours driven by the Bhattacharyya gradient flow," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2787–2801, Dec. 2007. 57

[29] A. Lopes, R. Touzi, and E. Nezry, "Adaptive speckle filters and scene heterogeneity," *IEEE. Trans. Geosci. Remote Sensing*, vol. 28, pp. 992–1000, 1990. 57

[30] A. Lopes, E. Nezry, R. Touzi, and H. Laur, "Structure detection and statistical adaptive speckle filtering in sar images," *Int. J. Remote Sensing*, vol. 14, no. 9, pp. 1735–1758, 1993. 57

[31] B. W. Silverman, *Density estimation for statistics and data analysis.* CRC Press, 1986. 60

[32] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc.*, vol. 35, pp. 99–109, 1943. 62

[33] J. A. Sethian, *Level set methods and fast marching methods*, 2nd ed. Cambridge University Press, 1999. 62, 63, 68, 70, 73

[34] T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001. 65

[35] V. Caselles, R. Kimmel, and G. Sapiro, "On geodesic active contours," *Int. J. Comput. Vis.*, vol. 22, no. 1, pp. 61–79, 1997. 65

[36] A. Yezzi, S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum, "A geometric snake model for segmentation of medical imagery," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 199–209, Apr. 1997. 65

[37] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 137–154, 2003. 67

[38] S. Osher and N. Paragios, *Geometric level set methods in imaging, vision, and graphics*. Springer, 2003. 70

[39] J. Weickert, "Coherence-enhancing diffusion filtering," *Int. J. Comput. Vision*, vol. 31, no. 2-3, pp. 111–127, 1999. 72

[40] J. Weickert, B. M. T. H. Romeny, and M. A. Viergever, "Efficient and reliable schemes for nonlinear diffusion filtering," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 398–410, Mar. 1998. 72

[41] D. Freedman, R. J. Radke, T. Zhang, Y. Jeong, D. M. Lovelock, and G. T. Y. Chen, "Model-based segmentation of medical imagery by matching distributions," *IEEE Transactions on Medical Imaging*, vol. 24, no. 3, pp. 281–292, March 2005. 80, 82, 83, 84