Segmentation of 3D Carotid Ultrasound Images Using Weak Geometric Priors

by

Igor Solovey

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

Vascular diseases are among the leading causes of death in Canada and around the globe. A major underlying cause of most such medical conditions is atherosclerosis, a gradual accumulation of plaque on the walls of blood vessels. Particularly vulnerable to atherosclerosis is the carotid artery, which carries blood to the brain. Dangerous narrowing of the carotid artery can lead to embolism, a dislodgement of plaque fragments which travel to the brain and are the cause of most strokes. If this pathology can be detected early, such a deadly scenario can be potentially prevented through treatment or surgery. This not only improves the patient's prognosis, but also dramatically lowers the overall cost of their treatment.

Medical imaging is an indispensable tool for early detection of atherosclerosis, in particular since the exact location and shape of the plaque need to be known for accurate diagnosis. This can be achieved by locating the plaque inside the artery and measuring its volume or texture, a process which is greatly aided by image segmentation. In particular, the use of ultrasound imaging is desirable because it is a cost-effective and safe modality. However, ultrasonic images depict sound-reflecting properties of tissue, and thus suffer from a number of unique artifacts not present in other medical images, such as acoustic shadowing, speckle noise and discontinuous tissue boundaries. A robust ultrasound image segmentation technique must take these properties into account.

Prior to segmentation, an important pre-processing step is the extraction of a series of features from the image via application of various transforms and non-linear filters. A number of such features are explored and evaluated, many of them resulting in piecewise smooth images. It is also proposed to decompose the ultrasound image into several statistically distinct components. These components can be then used as features directly, or other features can be obtained from them instead of the original image. The decomposition scheme is derived using Maximum-a-Posteriori estimation framework and is efficiently computable.

Furthermore, this work presents and evaluates an algorithm for segmenting the carotid artery in 3D ultrasound images from other tissues. The algorithm incorporates information from different sources using an energy minimization framework. Using the ultrasound image itself, statistical differences between the region of interest and its background are exploited, and maximal overlap with strong image edges encouraged. In order to aid the convergence to anatomically accurate shapes, as well as to deal with the above-mentioned artifacts, prior knowledge is incorporated into the algorithm by using weak geometric priors. The performance of the algorithm is tested on a number of available 3D images, and encouraging results are obtained and discussed.

Acknowledgements

I would like to thank my supervisor Dr. Oleg Michailovich for tireless guidance and support. His enthusiasm, thorough knowledge and deep passion for medical imaging and many other scientific topics is a true inspiration to aspiring researchers. I also thank my fellow engineering graduate students Alan, Elad, Mahdi, Robert and Sudipto for many invaluable and fruitful discussions.

Dedication

This work is dedicated to my mother Olga and other family, and especially to my girlfriend Annalisa. Thank you all for your support and encouragement.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Computer-aided diagnosis is an increasingly successful and relevant field of biomedical engineering. A multitude of tedious, time-consuming or otherwise intractable medical tasks can be made feasible by combining medical expert knowledge with signal processing and pattern analysis tools. In particular, a wide variety of medical image processing techniques have been devised in order to solve the problems of registration of images from different sources, their transformation into a more informative form, or delineation of regions of interest by means of automatic segmentation. This often enables improved diagnosis of various medical conditions or makes it more cost-effective or less invasive.

1.1 Carotid atherosclerosis

Cardiovascular disease is among the leading causes of death and disability around the world[58], contributing 23% to the overall mortality in Canada[57]. In particular, *atherosclerosis* is a dangerous vascular condition which is difficult to diagnose at early, subclinical stages. A progressive buildup of atheromatous plaque on the walls of blood vessels due to deposits of fatty materials can eventually cause stenosis (abnormal narrowing) of the blood vessel, embolisms, and lead to death[63].

A location in the vascular system which is of particular concern for the diagnosis of atherosclerosis is the *carotid artery*, which carries blood to the head and face. The *carotid bifurcation* region, where the common carotid artery branches off into two arteries, is a particularly vulnerable spot for plaque formation. Because geometry of vessel walls in the bifurcation region is complex, the blood flow is characterized by rapid variations in the flow rates and significant shear stress, which result in higher than usual vulnerability of the vessel walls, and their susceptibility to the formation of atherosclerotic plaque[63]. At advanced stages of atherosclerosis in this region, a portion of plaque can break off and be

carried with the bloodstream into the brain, blocking the blood supply to brain tissue, eventually killing it due to ischemia. This mechanism is the cause of over 80% of the incidences of stroke[6]. Before such severe symptoms occur, the disease progresses subclinically, from initial fatty streaks on vessel walls to atheromatous plaque and calcifications[40].

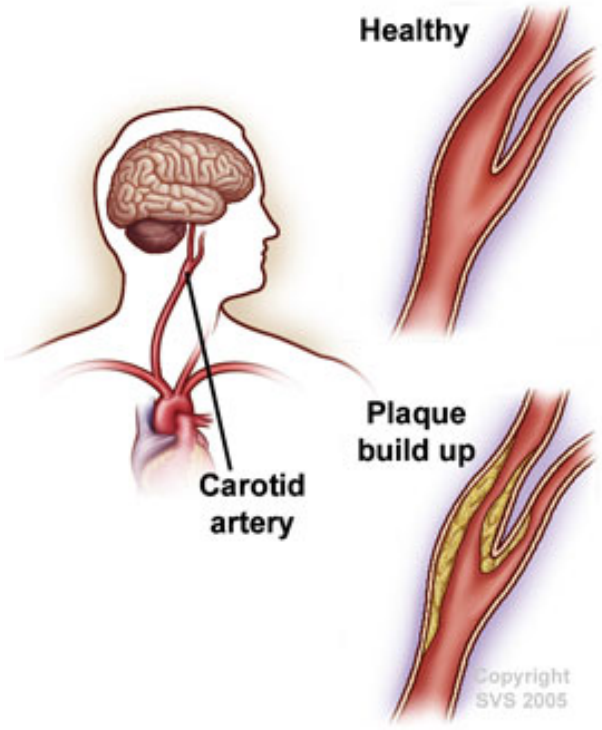


Figure 1.1: The location of the carotid bifurcation and an illustration of atherosclerotic plaque buildup[62].

1.2 Medical imaging of atherosclerosis

Diagnosis of carotid atherosclerosis, as well as medical trials assessing the efficacy of pharmaceutical treatment, rely on a variety of indicators to assess the health of the carotid artery. Indirect physiologic measurements such as blood pressure or cholesterol levels can provide indications of severe atherosclerosis[63], but they do not isolate it to the carotid artery or measure the plaque growth process itself, only factors correlated with it. On the other hand, medical imaging techniques provide the means to examine the plaque anatomy directly. Modalities that are most commonly used for plaque assessment in the carotid artery are described below. Their resolution and contrast is clearly a crucial

factor in determining their suitability for screening of the general population. However, other key parameters are the degree of invasiveness, reproducibility, availability and cost effectiveness[40].

In particular, treatment of atherosclerosis includes an extensive array of pharmaceuticals, such as anti-coagulants and anti-cholesterol drugs. In addition, since high blood pressure leads to increased strain on vessel walls, blood pressure lowering drugs are also administered[17]. A number of imaging techniques have been approved for use in validation of the efficacy of new pharmaceuticals in drug trials. Angiography has been a traditional tool used for this purpose, but other imaging modalities such as ultrasound have become increasingly viable[47, 26]. Biological markers derived from the images were determined to be indicative of atherosclerosis severity. One drawback of such approaches, as with diagnostic atherosclerotic imaging, is that the measurements are frequently carried out manually, and are thus subjective and time-consuming. For these reasons, it is desirable to develop automatic methods to perform the same tasks.

1.2.1 Angiography

Angiography is a collective term for imaging techniques that allow visualization of blood vessels by injection of a contrast agent into the bloodstream. This makes the blood significantly more discernible in the resulting image, highlighting any potential narrowing or occlusions in the imaged vessels. The traditional angiography modality is x-ray imaging, in which a contrast agent that absorbs x-rays is injected into the bloodstream and an x-ray image is acquired. Digital subtraction angiography (DSA) is a modification of the procedure that produces an improvement in contrast; two images are taken, one before and one after the agent is administered, and are subtracted to produce an image of the blood vessels without most of the surrounding tissue. X-ray-based angiography produces a two-dimensional projection image which may not be sufficient for visualization of blood flow and diagnosis of stenosis and obstruction in the vessels.

CT Angiography Computed Tomography (CT) Angiography produces an inherently three-dimensional depiction of the blood vessel structure by taking a series of x-ray images at regular spacing or angles, and computing a single 3D image. It is often used to identify atherosclerotic disease in the carotid bifurcation region[38]. Although it is widely available and cost-effective, CT angiography involves the use of iodine-based contrast material that can produce a severe allergic reaction; moreover, patients who have advanced diabetes or kidney problems should avoid this procedure because the contrast agent can further reduce their kidney function[38]. Furthermore, CT angiography exposes the patient to a significant dose of ionizing radiation. In cases when severe (or even life-threatening) atherosclerosis

is suspected, the benefits outweigh the risk. However, the use of CT angiography may not be justified for less severe cases.

MR Angiography Magnetic Resonance Imaging is a powerful imaging modality that has also been adapted for imaging the blood vessels, both with and without the use of contrast agent. Although MR angiography has been used successfully for detection of atherosclerosis, particularly in the arteries of head and neck such as the carotid artery, its biggest drawback is a comparatively high cost as well as low availability, in particular in Canada. Its chief advantage is the high quality of resulting images.[19].

Although angiography is superior in its ability to visualize the artery lumen, it only highlights occlusions and narrowing of the artery, without giving any further information about the composition and shape of the plaque itself. A more important drawback for the task of early detection of atherosclerosis is the use of ionizing radiation in traditional and CT angiography[38], and the prohibitive cost of MRI.

Although these imaging modalities have established uses in atherosclerosis imaging, a viable alternative exists in the form of ultrasound imaging. Its primary advantages over CT and MRI are cost-effectiveness and safety, as well as portability. Since ultrasound is at the centre of this research work, it is discussed in more detail in the succeeding chapter. The remainder of this thesis is organized as follows. Chapter 3 discusses the segmentation approaches, focusing on the ones suitable for ultrasound segmentation in 3D. Chapter 4 discusses features which can be extracted from the ultrasound images and used by the segmentation algorithm. Chapter 5 outlines the main contributions of this thesis: a novel ultrasound image decomposition technique and a weak geometric prior energy which is added to the segmentation algorithm to augment the observed information with prior knowledge. Chapter 6 discusses the implementation of the segmentation algorithm and the curve representation method, as well as other computational details. In Chapter 7, automatic segmentation boundary is compared to manual segmentation both qualitatively and quantitatively. Chapter 8 concludes this work with a summary and an outline of future work.

Chapter 2

Ultrasound Imaging

Due to its cost-effective and safe nature, the ultrasound scanner is the most readily available imaging tool in today's hospitals. In particular, ultrasound is one of the most widely used technologies in carotid atherosclerosis imaging, and the focus of this work. Despite its advantages, ultrasound images present a unique challenge for quantitative processing, necessitating the development of algorithms specifically tailored to it. For this reason, a more detailed look is required at the physics of ultrasound, the design of the ultrasound transducer and its unique statistical properties.

2.1 Ultrasound physics

An ultrasonic wave is an oscillatory variation in pressure, whose frequency is above the range of human hearing. As any wave, it possesses certain parameters such as wavelength and energy. Unlike an electromagnetic wave, it can only travel through a medium, in which it produces alternating regions of compressions and rarefactions as depicted in Figure 2.1. The speed with which the sound wave travels depends on such properties of the medium as density and elasticity.[20]

As it travels through a complex object like biological tissue, the ultrasound wave gradually loses its energy. Each medium *absorbs* a certain portion of the energy, transforming it into random heat energy. Whenever there is a change in the acoustic properties of the medium, as happens on the boundaries between two anatomical tissues, some energy of the ultrasound wave travelling through this region will be reflected back. These *reflections* form the basis of an ultrasound image.

When the wave travels across a large and smooth interface between two media with different associated speeds of sound, its behaviour is governed by the characteristic *acoustic*

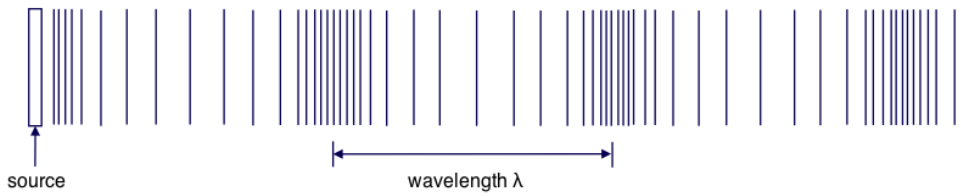


Figure 2.1: Propagation of a longitudinal wave, showing alternating sections of compressions and rarefactions.

impedance of the interface, which is proportional to the difference in the sound velocities on either side. A portion of the energy of the incident wave will be reflected, resulting in *specular reflection*, and the remainder will travel on further[52]. In addition to crossing large and smooth interfaces, ultrasonic waves also encounter relatively small objects, which result in *diffuse scattering* of the incident wave. The two modes are contrasted in Figure 2.2. Unlike specular reflection, the portion of the wave's energy scattered by such small objects has no predominant orientation; it travels diffusely in wide range of directions. This phenomenon introduces an incoherent interference component into the resulting image[20].

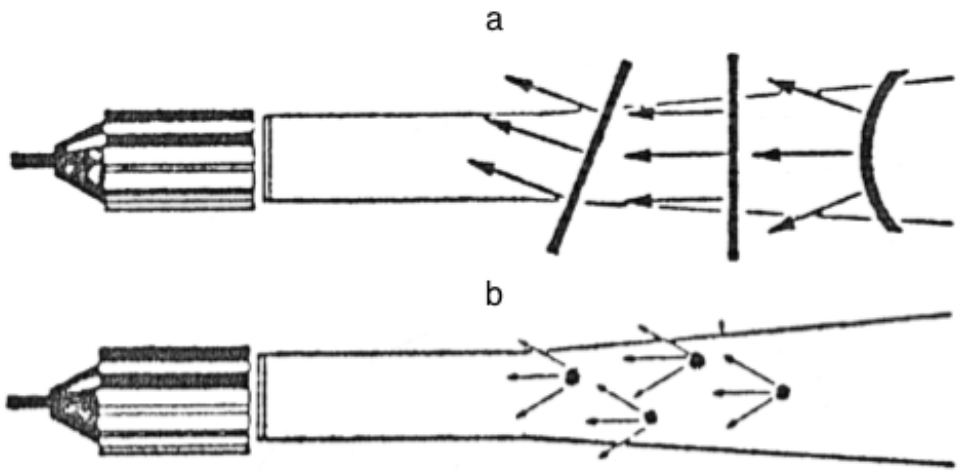


Figure 2.2: Ultrasound wave reflection types. a) Specular reflection from large and smooth acoustic interfaces; b) Scattering reflection from small objects;

In order to relate the amount of specular reflection to acoustic impedance, let Z_1 and Z_2 represent such acoustic impedances of two different media, related to their respective speed of sound, c_1 and c_2. Then the ratio of the reflected energy, known as the reflection

coefficient, is given by

$$r = \frac{Z_1 - Z_2}{Z_1 + Z_2} \approx \frac{c_1 - c_2}{c_1 + c_2}. \tag{2.1}$$

This expression suggests that the larger the difference between c_1 and c_2, the higher the percentage of energy reflected[20]. On the one hand, these strong reflections will result in strong boundaries in the ultrasound image. However, another implication is that any tissue interfaces further away from the wave's source will be exposed to a wave with much less energy, and thus will be less visible on the ultrasound image when reflected. Since the closer boundaries effectively shadow deeper ones, this phenomenon is generally known in ultrasound as *acoustic shadowing*. It degrades the quality of the image of those portions of the tissue which are "shadowed" by a strong tissue interface in front of them.

When the inter-medium boundary and the direction in which the wave propagates are not at a 90° angle, Snell's law applies: the angle of the reflected and refracted waves will be proportional to the incident angle, as shown in Figure 2.3. Moreover, the energies of the reflected and refracted waves will depend on the incident angle as well. This dependence on the incident angle has an important impact on the nature of images produced by ultrasonic devices: they are orientation-dependent[20]. *Weak and discontinuous boundaries* result when tissue boundaries are oriented in an axial or nearly axial direction.

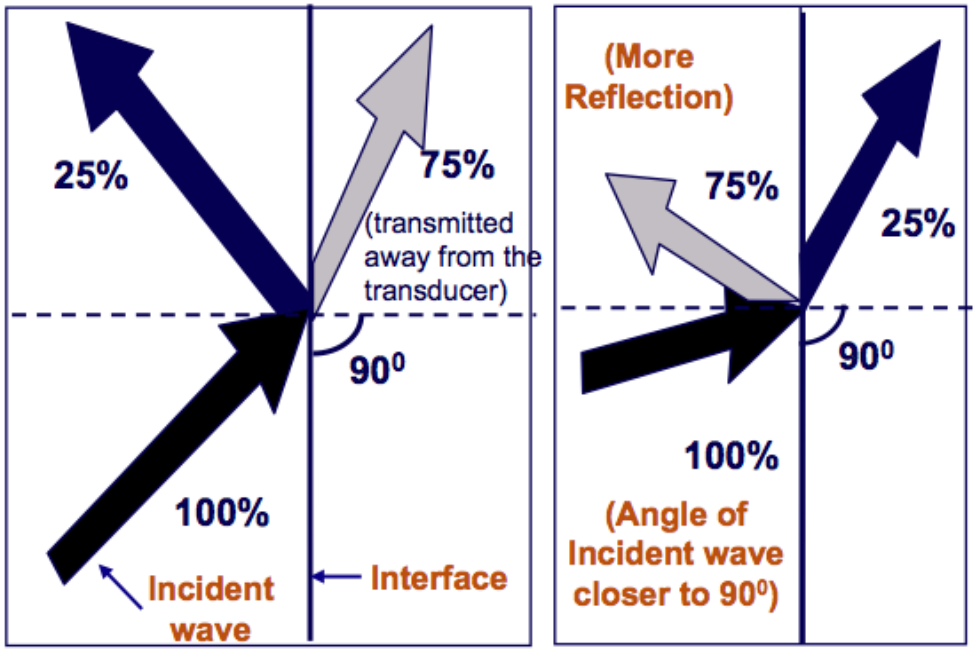


Figure 2.3: Dependence of reflection strength on incident angle of the ultrasound wave.

As a result, ultrasound-specific artifacts such as acoustic shadowing and orientation-

dependent dropouts degrade the quality of the image and make its interpretation more challenging, whether the latter is done quantitatively or qualitatively. Another important property of ultrasound envelope images is the *speckle* pattern, which is discussed in Section 2.3.

2.2 Ultrasound imaging systems

Ultrasound imaging tools generate ultrasonic waves whose behaviour is described above, as well as register the reflected waves and process them to result in an ultrasound image. In doing so they exploit the variations in acoustic response properties of the various tissues inside the human body to image them. Because the strongest reflections are produced by specular reflectors, the best imaged anatomical structures are the boundaries.

2.2.1 Generation and detection

An ultrasound pressure wave is both generated and retrieved by a transducer working on the principle of piezoelectricity. Piezoelectric crystals change their physical dimensions and cause a pressure variation when an electric voltage is applied to them. This is used to create an ultrasonic wave from an electrical pulse supplied through an electrode. Conversely, a pressure variation impacting the crystal induces an a voltage change in the electrode. A reflected ultrasonic wave is thus detected by the ultrasound imaging system.

A single such transducer can send and receive an ultrasonic wave and calculate the depth of the anatomical feature from which the wave is reflected from the time difference between the sent and received wave. Such one-dimensional depth profile is known as amplitude- or *A-mode* scanning. As previously described, the scattering and specular reflection results in reflected waves in many different directions. Only the waves reflected in the incident direction will be retrieved by the transducer.

If the A-mode amplitude profile is rendered as intensity dependent on the strength of the received echo, it can be considered a one-dimensional "image" of the tissue directly in front of the transducer. If many such scans are performed in series, however, by rotating or linearly translating the transducer probe to acquire the next image, the result will be a 2D intensity map reflecting the acoustic impedance of the imaged tissue. This is known as a *B-mode* image[52].

Although a B-scan image can be acquired by a single mechanically moveable transducer, modern ultrasound probes instead employ electronic *transducer arrays*. Multiple transducers are arranged in linear or circular order, and a single one-dimensional scan is produced by synthesizing the pulses from many transducers in ways that significantly improve the

resolution of the resulting image, a process known as *beamforming*. The pulses can be timed to focus the beam on a particular location inside the tissue, as shown for example in Figure 2.4. Similarly, the reflected waves are received with varying delays depending on the angle, and such differences are compensated for to get a clearer image.

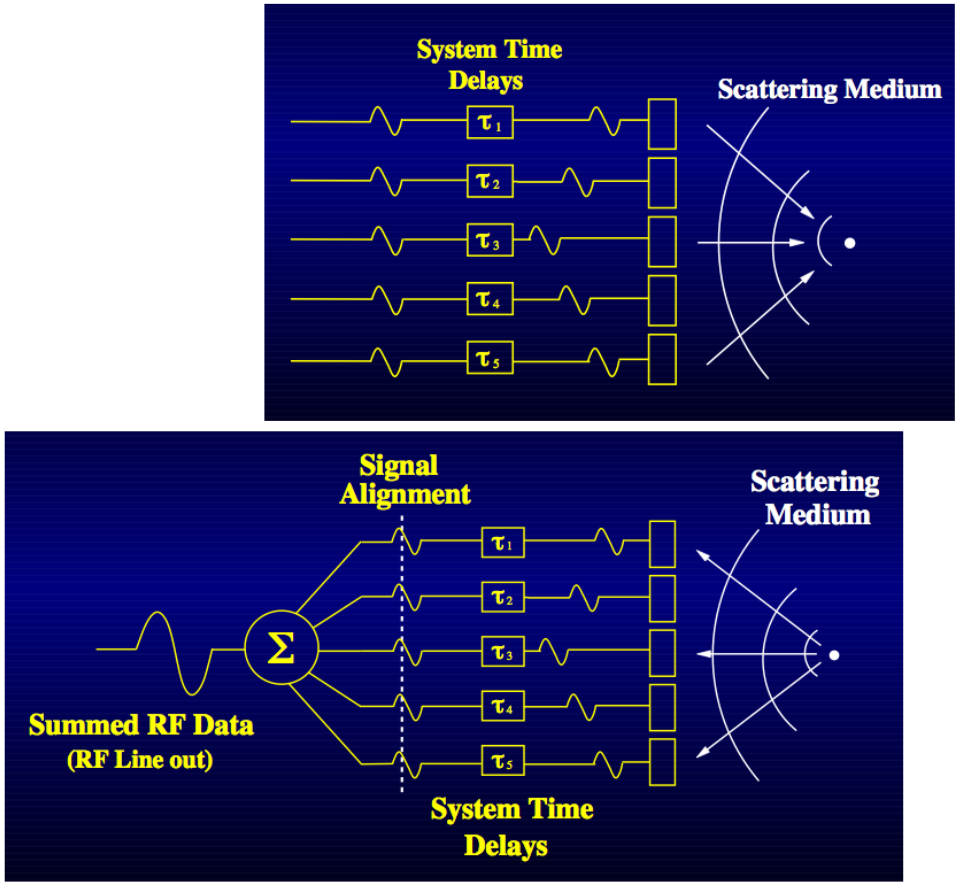


Figure 2.4: Ultrasound beamforming[43]

2.2.2 Post-processing

The signal received by a single transducer or a transducer array must be further processed to make it suitable for displaying on the screen or storing in a digital form for later viewing. These procedures are normally applied in the ultrasound system's hardware immediately upon acquisition of the signal.

The frequency of the transmitted wave is in the ultrasonic range, typically between 3 MHz and 18 MHz[52]. The first step in its processing is to demodulate this radio frequency

(RF) image, resulting in the complex-valued in-phase/quadrature (IQ) image. The absolute value of this image is obtained, resulting in the *envelope image* – the form which is suitable for display, with further post-processing. The diagram in Figure 2.5 shows this process as part of a typical B-mode ultrasound system.

Due to the combined effects of absorption, reflection and scattering, the ultrasound wave loses its energy as it propagates deeper into the imaged tissue. In fact, the signal strength decreases exponentially with distance from the transducer. An uncorrected image would show very strong structures corresponding to tissues close to the transducer and severely weakened ones for tissues at higher depths. Thus the received signal undergoes *time-gain compensation* (TGC), an amplification of the reflected wave that increases with time since transmission, and thus, with distance of the imaged tissue from the transducer. The degree of TGC can be adjusted depending on the tissue being imaged. In practice the amount of TGC is limited to reflections whose strength is above the noise level, and thus the depth of tissue which can be imaged is limited as well[52]. This step can either be performed before demodulation and envelope detection, or after, as Figure 2.5 shows.

Moreover, the reflected wave picked up by the transducer can have a wide range of amplitudes. Strong tissue interfaces close to the transducer will produce high-amplitude reflections, while deeper or weaker boundaries will result in much smaller amplitudes. In addition, specularly reflected echos have a much higher amplitude than echos from scattering reflectors. As a result, the dynamic range of the amplitude can be as high as 80 dB. Quantizing signals with such high dynamic ranges requires a lot of storage capacity to maintain low quantization error. Displaying such signals on monitors whose dynamic range is typically in the range of 40 dB will render many details in low contrast - the strongest reflections will dominate so much that everything else may not be discernible. The solution to this problem is to *compress* the dynamic range, which is done by the following type of function:

$$f(x) = A \log(Dx + B), \tag{2.2}$$

where the user-adjustable parameter D controls the degree of compression, and constants A and B depend on the signal's digital resolution. The signal transformed by the function f is both displayed and stored for later use[52]. An illustration of the effects of log-compression are presented in Figure 2.6.

2.3 Ultrasound statistics

In order to devise algorithms for processing of ultrasound images, including their denoising and segmentation, it is necessary to account for the statistics of their formation. This procedure must be justified with a choice of a suitable statistical model for the ultrasonic signal acquisition. Central to such development is the treatment of the *speckle pattern*

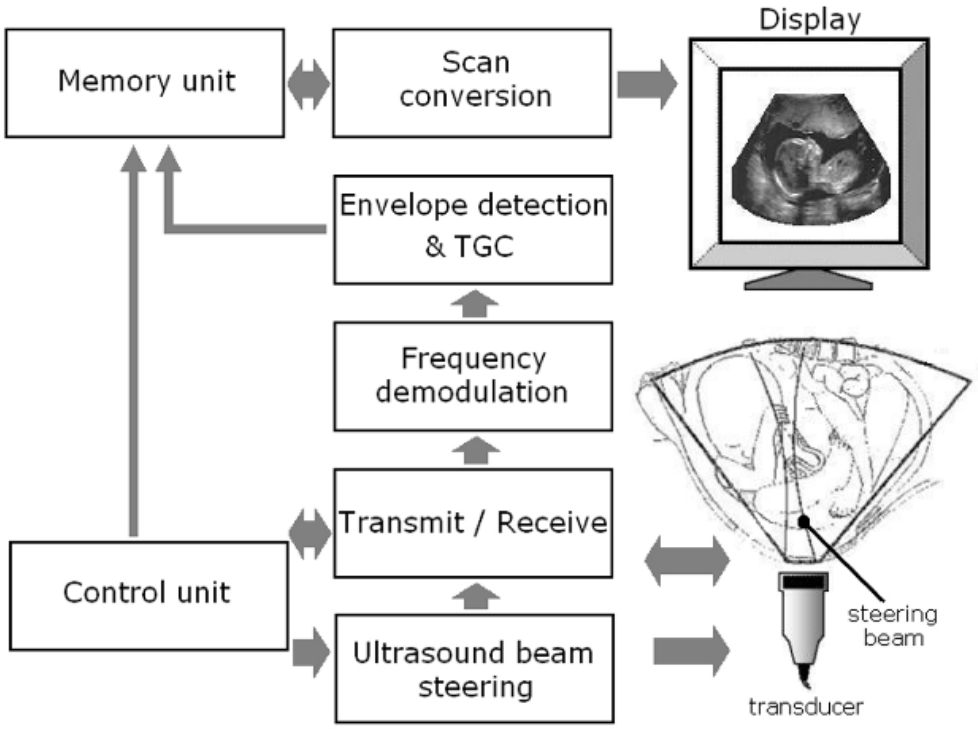


Figure 2.5: An overview of an ultrasound system. The received signal is demodulated, its absolute value is taken, time-gain conversion and scan conversion are performed. Dynamic range conversion is part of the latter procedure[23].

- a characteristic pattern which is present in all ultrasound images and is illustrated in Figure 2.6. Regardless of whether speckle is treated as a diagnostically useful image feature or as undesirable noise to be filtered away, its treatment is an essential part of a well-designed segmentation algorithm[39].

2.3.1 Ultrasound noise model

As described before, the reflected echos of the incident ultrasound wave may be broadly grouped into two types. Specular reflections are the echos from strong, relatively smooth and long interfaces between two tissues of substantially different echogeneity. Echos from scattered reflectors originate from more irregular boundaries and from objects which are small compared to the wavelength of the wave. This second type of reflection results in a speckle interference pattern which can be treated from the viewpoint of stochastic signal analysis.

In general, a speckle pattern results from the mutual interference of a number of

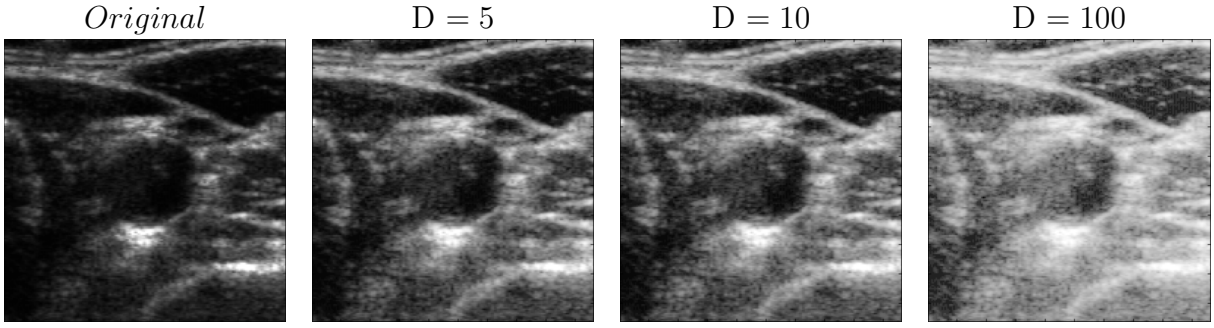


Figure 2.6: Illustration of the speckle pattern and the effects of log compression in a B-mode image, with different values of D from (2.2).

partially-coherent wavefronts, subject to small temporal and spatial fluctuations[1]. In ultrasound, this happens when a coherent ultrasonic wave is scattered due to heterogeneities in the objects it encounters, such as tissue parenchyma. The ultrasound RF image is based on the spatially varying echogenic properties of the imaged tissue, represented by its *tissue reflectivity function* (trf). It also depends on the parameters of the ultrasound scanner, especially its resolution. The resolution of an ultrasound beam, whether formed by a single transducer or a transducer array, is defined by its lateral width and temporal duration[1]. These parameters may be used to define a resolution cell, a volume of space whose width in the two transverse directions (orthogonal to the beam) is defined by the beam width, and the longitudinal direction (parallel to the beam) is defined by the beam duration. If two objects are located within the same resolution cell, an observer will not be able to resolve them as separate in the resulting ultrasound envelope image. Inside a homogenous section of the imaged tissue, there may be multiple such small-scale objects each resulting in diffuse scattering of the incident ultrasonic wave. The accumulation of scatterings from all these objects influence the behaviour of the resulting speckle pattern. Different tissue composition will result in different patterns and densities of such reflectors, and ultimately in different speckle textures in the final envelope image.

The stochastic behaviour of the speckle pattern in the envelope image is dependent on the distribution and number of scattering reflectors in the *tissue reflectivity function*, as well as their mutual dependence. In particular, a large number of independent Gaussian-distributed scatterers of zero mean and variance σ_f^2 in the trf will result in an envelope image whose speckle noise is Rayleigh distributed[64]:

$$p_Z(z) = \frac{z}{\sigma_f^2} \exp\left\{-\frac{1}{2}\left(\frac{z}{\sigma_f}\right)^2\right\},\ z \geq 0. \tag{2.3}$$

After the application of log compression, the random variable obeys the Fisher-Tippett

distribution[36]:

$$p_Y(y) = 2\exp\{2\,(y - \ln\sigma_f) - \exp\{2\,(y - \ln\sigma_f)\}\}, \tag{2.4}$$

with σ_f^2 denoting the variance of the tissue reflectivity function. It is characterized by an asymmetric pdf with a long left tail, as illustrated in the left subplot of Figure 2.7. A realization of this random variable, shown in the right subplot of Figure 2.7 contains large spikes, corresponding with the left tail in the pdf.

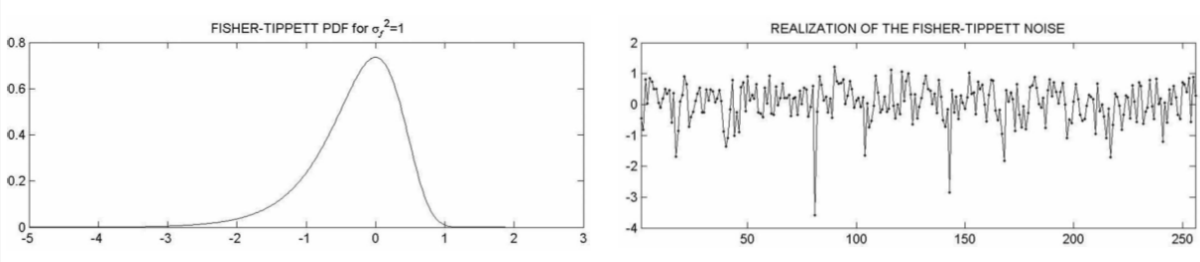


Figure 2.7: (Left subplot) The Fisher-Tippett distribution, with $\sigma_f^2 = 1$; (Right subplot) A realization of Fisher-Tippett-distributed random variable[36]

.

The contamination of the ultrasound envelope image with this speckle pattern is often defined using the following noise model:

$$I(\mathbf{x}) = u(\mathbf{x})v(\mathbf{x}) + e(\mathbf{x}), \tag{2.5}$$

where $\mathbf{x}$ is a spatial coordinate, I is the observed envelope image, u is the original image (a function of the tissue echogeneity), v and e are the multiplicative and additive components of speckle pattern, respectively[36]. It should be noted that although this multiplicative model of speckle has been widely used, it is not the only possible formation model, nor is it free of limitations, such as its limited accuracy regarding fine details unresolvable by the imaging system[60]. In ultrasound, the additive component e[36] is usually negligible, so it can be dropped. In order to be able to remove or at least separate the signal-intensity-dependent speckle v, it is necessary to resort to *homomorphic* processing, whereby the logarithm operator transforms the model into that of additive log-speckle:

$$I_l(\mathbf{x}) = \log(I(\mathbf{x})) \approx \log(u(\mathbf{x})v(\mathbf{x})) = u_l(\mathbf{x}) + v_l(\mathbf{x}). \tag{2.6}$$

Using this model, the envelope image can be post-processed from a statistical estimation point of view. This technique will be used later in this work to suggest a new feature extraction technique based on image decomposition.

2.4 Application to carotid atherosclerosis imaging

Ultrasound imaging is used for two main purposes in vascular medicine: to guide invasive medical procedures, and to visualize the degree of stenosis for diagnosis of atherosclerosis. In addition to the B-mode ultrasound described above, several other types of ultrasound are used for these purposes.

Intravascular ultrasound (IVUS) images the carotid artery from the inside: through an incision, a catheter is inserted into the artery with an ultrasound probe at its tip. Although primarily used for imaging of the coronary arteries[48], IVUS is used for plaque assessment in the carotid artery as well. It is a highly accurate alternative to angiography and has been used to evaluate the efficacy of drug therapies in treating atherosclerosis[48]. The primary drawback of this surgical technique is its invasiveness, as well as the resulting high cost.

Doppler ultrasound is the imaging technique most commonly used to detect the location and extent of atherosclerotic plaque in those blood vessels which are easily accessible to an external transducer, one of which is the carotid artery. Such assessment is done by measuring the blood flow velocity in the artery by exploiting the Doppler effect, which relates the velocity of a moving sound pulse reflector to the (easily measurable) frequency shift of the pulse. Such measurements can be used to estimate actual velocity, or to color-code it and overlay it with the ultrasound image itself. Additionally, the frequency spectrum of the returning signal is analysed to provide information about the blood flow dynamics.

2.4.1 Three-dimensional ultrasound

Ultrasound imaging, in particular the standard B-mode ultrasound, is traditionally a two-dimensional imaging modality. Since the anatomical tissues are inherently three-dimensional objects, the use of B-mode ultrasound to diagnose medical conditions such as atherosclerosis involves an interactive process of acquiring a sequence of 2D images and building a mental picture of the 3D anatomy. The quality of the results thus depends heavily on the experience of the clinician guiding the ultrasound probe[16], and is subject to considerable inter- and intra-observer variability. In particular, diagnosing and monitoring atherosclerotic disease by 2D ultrasound imaging is problematic due to such variability, as well as the difficulty of re-positioning the ultrasound probe over the same region of the plaque at a later time. Additionally, the anatomy of the patient sometimes results in inaccessibility of a good acquisition angle[13].

Several methods exist for creating three-dimensional B-mode ultrasound imagery, resulting in data sets that can both be qualitatively analyzed by human observers and quantitatively used in computer-aided diagnosis systems. Although two-dimensional phased array transducers are an increasingly viable technology and would be the most straightforward to use, their use for imaging the carotid artery is limited by the prohibitive cost, limited spatial resolution, as well as practical limitations in the available technology, such as the difficulty of transmitting many signals over a practical cable[16]. All other technologies involve a mechanically augmented 1D ultrasound transducer.

One popular method involves manual acquisition of 2D images, which results in their inexact positioning, as is shown in Figure 2.8. In order to correct this artifact, the system keeps track of the position and orientation of the 2D images using a position monitor, which is usually a magnetic sensor. With such information recorded, the 2D images are combined into a 3D volume by reconstruction software, which determines the values of the voxels that have not been acquired by interpolation. The reconstruction is fast enough for the 3D volume to be accessible immediately. Such *free-hand* scanning systems have the advantage of being small and as flexible, and thus not interfering with the acquisition. However, several drawbacks limit the use of freehand 3D acquisition systems. Firstly, the magnetic tracking sensor used in most freehand systems is affected by electromagnetic interference from monitors, AC power cables and the transducer, which may result in significant distortions in the position information stored with each 2D image. Other position tracking approaches, such as acoustic devices, or position estimation by speckle de-correlation, have also been used. Additionally, the operator must ensure that the acquired images adequately cover the 3D volume with no significant gaps[16, 13].

An alternative acquisition method is to control the movement of the transducer probe by a mechanical device with a motor. In the case of imaging the carotid artery, 2D images are acquired continuously while the mechanical device linearly translates the transducer in the transverse direction. Rotational and tilting movements are used for other applications, such as obstetrics. The speed of the mechanical motion, as well as the transducer's elevational resolution, determine the spatial resolution in the direction of motion. A *mechanical linear scanning* approach can be used to acquire a 3D volume of the carotid bifurcation region in around 20 seconds. The chief advantage of this technique is that, aside from initial positioning, the acquisition process does not depend on the operator, thus removing the possibility of any motion tracking artifacts[13, 12]. Additionally, the need for 3D volume reconstruction by interpolation is eliminated, and the volume is a direct stacking of 2D B-scan images with predefined positions, as shown in Figure 2.8. This makes quantitative processing of the data more robust, since it better preserves the statistical nature of the data, and speckle patterns, than 3D volumes reconstructed from freehand scans, in which a significant number of voxels are averaged or interpolated. Linear scanning 3D ultrasound has been employed, for example, to examine the extent and morphology of plaque before

and after administration of the drug treatment[26].

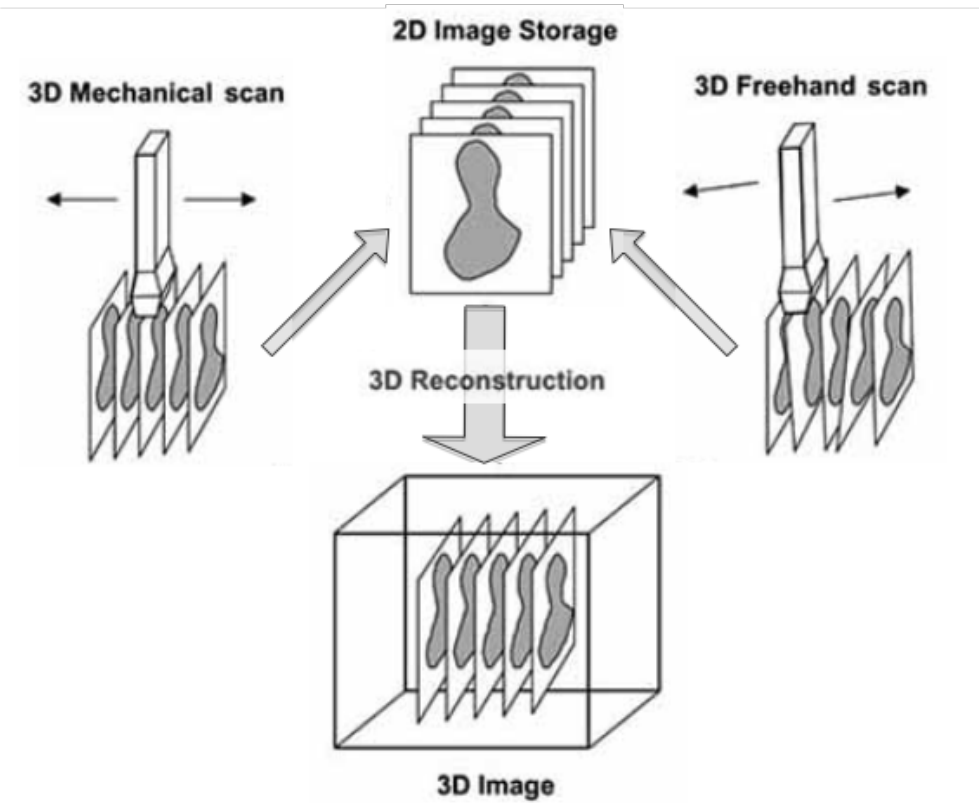


Figure 2.8: Three-dimensional Ultrasound acquisition methods. A series of 2D images are acquired either by a mechanical mechanism with exact spacings and orientations, or by a freehand system which records their position and orientation. The images are stored and processed to construct a 3D image.[13]

Regardless of the acquisition methodology, a number of important artifacts affect the quality of the data, with their prominence increasing when the volume takes longer to acquire. Firstly, as is the case with all imaging modalities, the patient needs to remain motionless so that motion artifacts do not corrupt the image. Depending on the length of acquisition, respiratory and pulsatile motion may need to be accounted for. Motion due to blood flow is an especially significant artifact in cardiovascular imaging, and therefore is often compensated for by recording the images at the same stage of the cardiac cycle; this process is known as cardiac gating. Additionally, the acquisition procedure itself is a source of error due to the deformation of the anatomy caused by pressing the ultrasound probe to the skin; freehand scanning is more prone to this artifact[16]. Nevertheless, 3D ultrasound is a much richer source of information when imaging three-dimensional

anatomical structure, one that is less dependent on the skill of the probe operator and more useful for off-line quantitative analysis.

Chapter 3

Image segmentation

Assessing the condition of the carotid artery bifurcation directly from the 3D ultrasound volume is a difficult task. Fundamentally, it involves locating the boundaries of the artery and examining the extent and volume of plaque. Due to the nature of ultrasonic images, such a task is difficult to perform manually and directly using the 3D dataset: it is time consuming, as well as prone to both inter- and intra-observer variability[26]. For this reason, automatic or semi-automatic segmentation approaches must be examined and customized to account for unique properties of ultrasound imaging.

The task of pinpointing the outer boundary of the carotid artery may be solved by means of an *image segmentation* algorithm. In general, let an image $I(\mathbf{x})$ be defined on a subdomain $\Omega \subset \mathcal{R}^n$ (where $n = 3$ in the case of 3D volume), so that $\mathbf{x} \in \Omega$. The image $I(\mathbf{x})$ may be scalar-valued, in which case a single scalar measurement is associated with each $\mathbf{x}$, or it can be a vector-valued image containing d such measurements or *features* at each coordinate $\mathbf{x}$, i.e. $I(\mathbf{x}) \in \mathcal{R}^d$. The object being depicted in $I(\mathbf{x})$ can be subdivided into K meaningful regions Ω_i, such that $\Omega = \bigcup_{i=0}^{K} \Omega_i$, $\Omega_i \cap \Omega_j = \{0\}, j \neq i$. Usually, the precise boundaries $\partial\Omega_i$ are not known in advance, and their estimate is needed for subsequent image processing tasks. The goal of image segmentation then is to classify each image pixel $\mathbf{x} \in \Omega$ as belonging to one of subregions Ω_i such that as few pixels as possible are mis-classified.

Although the number of regions K is unrestricted, it is often simpler to discuss segmentation approaches by setting $K = 2$. In particular, the goal of the segmentation of carotid ultrasound is to identify all pixels $\mathbf{x} \in \Omega$ as belonging to either the subdomain Ω_A of the artery (including vessel wall, plaque and lumen), or the subdomain Ω_B containing the remainder of the image, or the background. Although the basis of this segmentation process is the observed image I, in practice the artifacts associated with ultrasound imagery mean that the assignment of some pixel locations to either of the two subdomains will be ambiguous and will strongly depend on factors other than their the image values at

these locations. The subdomain Ω_A can be further subdivided into the plaque and lumen subregions, Ω_P and Ω_L. The goal of this thesis is to focus on the former task of segmenting the artery Ω_A from the background. Once this is successful, the artery can be further segmented into plaque and lumen. The advantage of this two-stage approach lies in the fact that the number of voxels belonging to the artery is usually small compared to the background. Once the artery has been found, delineating the plaque inside it is a much simpler and less computationally demanding task.

3.1 Main challenges

An application of image segmentation to 3D carotid ultrasound presents a number of unique challenges. Though a large number of general segmentation algorithms exist, a successful segmentation of the carotid artery must address these challenges and be adapted specifically for 3D ultrasound data.

Firstly, as detailed in Chapter 2, due to the way an ultrasound signal is formed, the resulting image is contaminated with an array of artifacts, such as incomplete or missing tissue boundaries, as well as interference known as speckle noise. One approach at overcoming these difficulties is to supplement the information derived from the ultrasound image with prior knowledge about the object to be segmented; this augments the observation with expert knowledge, expressed in a suitable mathematical formulation, and bears close resemblance to how medical imaging data is interpreted by human experts in clinical practice.

Secondly, operating on a three-dimensional data set is an extremely resource-intensive task. Many two-dimensional segmentation algorithms have conceptually straightforward extensions to 3D. However, the amount of data and the number of computations is orders of magnitude higher, sometimes rendering such extensions intractable. Since the goal of this work is to investigate not only robust and accurate, but computationally feasible approaches to carotid ultrasound segmentation, the issues of computation and memory efficiency must be carefully considered.

3.2 Segmentation approaches

Image segmentation is a fundamental task in medical image analysis, with a great many applications in medical diagnosis in particular. As a result, there have been a multitude of approaches to the problem of segmenting an image into regions of interest. Although it is impossible to categorize them all in a meaningful fashion, a review of segmentation strategies is presented here, with a particular focus on medical imaging, using as examples

both general algorithms and those specifically created for ultrasound imaging. It generally holds true that a successful segmentation approach is significantly customized to the domain of its application. Since ultrasound images are inherently noisy and due to their echogenic nature possess properties that are distinct from all other medical imaging modalities, the need for significant adaptation of any segmentation approach is especially great – no "canned solutions" will produce accurate and robust results.

3.2.1 Decision boundary

A segmentation problem can be characterized as a special case of the classification/clustering problem. Data samples, consisting of a single measurement or a vector of such measurements, are assigned labels corresponding to meaningful sub-classes of the data. This can be done by either utilizing previously labelled data points as training examples (classification or supervised learning) or without any known examples (clustering or unsupervised learning). This process can be visualized as drawing a decision boundary separating the data into distinct subsets. Two general types of decision boundaries can be distinguished, based on the domain in which they are defined.

In general, a d-dimensional data sample vector can be visualized as a point in a d-dimensional vector space, $\mathcal{V}^d$. Each of these d dimensions represents the value of a certain feature. In classical classification and clustering algorithms, the decision boundary separating the data points is defined as a manifold in this *feature space*, often a linear one, i.e. a hyperplane. This is the approach taken in classification of a variety of data types, including images. It usually rests upon building a statistical model of the data in the feature space, in the hopes that the data belonging to distinct classes or clusters will be separable by an *algebraically defined decision boundary* in the feature space.

However, imagery data contains additional structure: it is defined in a geometric coordinate space. The voxel $\mathbf{x}$ is treated as a geometric point positioned in a subdomain Ω of $\mathcal{R}^3$. In particular, in the discrete setting, the data belongs to a structured coordinate grid. The data thus possesses additional geometric properties, such as neighbour relationships and spatial distances. In most image processing tasks, an object of interest represents a spatially coherent grouping of pixels. This fact justifies an alternate domain of definition for the classification decision boundary - the coordinate space. Most image segmentation approaches use such *geometrically* defined decision boundaries because they deal with the spatial nature of coordinate-based data directly.

An important example of an algebraic decision boundary approach is segmentation by *thresholding*. This approach is statistically justified if, in the feature space, the data points belonging to distinct sub-regions of the image may be separated by a linear boundary with minimal error. Note that the quality of such delineation heavily depends on the choice of the

features themselves. In the discrete-valued domain, the feature space can be alternatively viewed as a histogram. The simplest algorithm involves a binary segmentation task and a single feature - image intensity. A global threshold is applied to the intensity histogram. In the case when segmentation into more than two regions is sought, the decision boundary is specified in terms of average values μ_i and tolerances ϵ_i, such that the subregions Ω_i are defined as

$$\Omega_i = \{\mathbf{x} \mid I(\mathbf{x}) \in [\mu_i - \epsilon_i, \mu_i + \epsilon_i]\}. \tag{3.1}$$

If the image $I(\mathbf{x})$ is vector-valued, then μ_i and ϵ_i are vectors as well. In general, there are two steps associated with applying the thresholding segmentation. Firstly, *threshold selection* either relies on an analyzing of the image histogram to detect distinguishable object types, or on prior knowledge of the statistical nature of the data. As an example of such prior knowledge, intensity of CT images has a well-characterized scale known as the Hounsfield scale. Different types of imaged anatomical objects, such as bones, muscle tissues and blood, have a known range of values in Hounsfield units, and thus the average values and tolerances are experimentally justified[46]. The second step is to post-process the result obtained by an algebraic histogram decision boundary to compensate for the shortcomings of this approach. This is frequently done by means of connected component analysis which requires manual user initialization by selecting a pixel inside the region of interest (ROI), and then automatically "growing" this selection into a sub-domain by iteratively attaching all neighbouring pixels that belong to the same segmentation region. Various morphological operations can also be performed to discard small misclassified regions and "smooth" the segmentation boundaries[46]. Such an approach has been frequently taken in CT image processing, for example in segmentation of lungs[22] and liver[53]. This approach has also been used in pre-segmenting the carotid bifurcation region in CT angiography data sets. In this case, thresholding was based on previous studies determining appropriate Hounsfield unit values for vessel tissue, and the reduced data set was subsequently re-segmented with a region-growing algorithm to remove outliers[56].

A large number of general thresholding segmentation algorithms have been devised, although their application toward ultrasonic data is problematic due to the statistical nature of the echogenic image formation, as previously discussed. When such an approach is taken toward segmenting B-scan carotid ultrasound, the chief problem is the noisiness of the decision boundary, as exemplified by a selection of such results in Figure 3.1, a comparison performed in[14]. A variety of statistical assumptions on the nature of the pixel data may be made, for example that the image intensity follows a Gaussian mixture model[25] (Figure 3.1(c)). Alternatively, a different feature can be used for the thresholding, such as minimum cross entropy, designed to minimize overlap between classes from an information-theoretic point of view[29] (Figure 3.1(d)). In all such results, the delineation is noisy and contains a lot of spurious sub-regions all over the image domain.

Other algebraic decision boundary algorithms stem from more involved statistical anal-

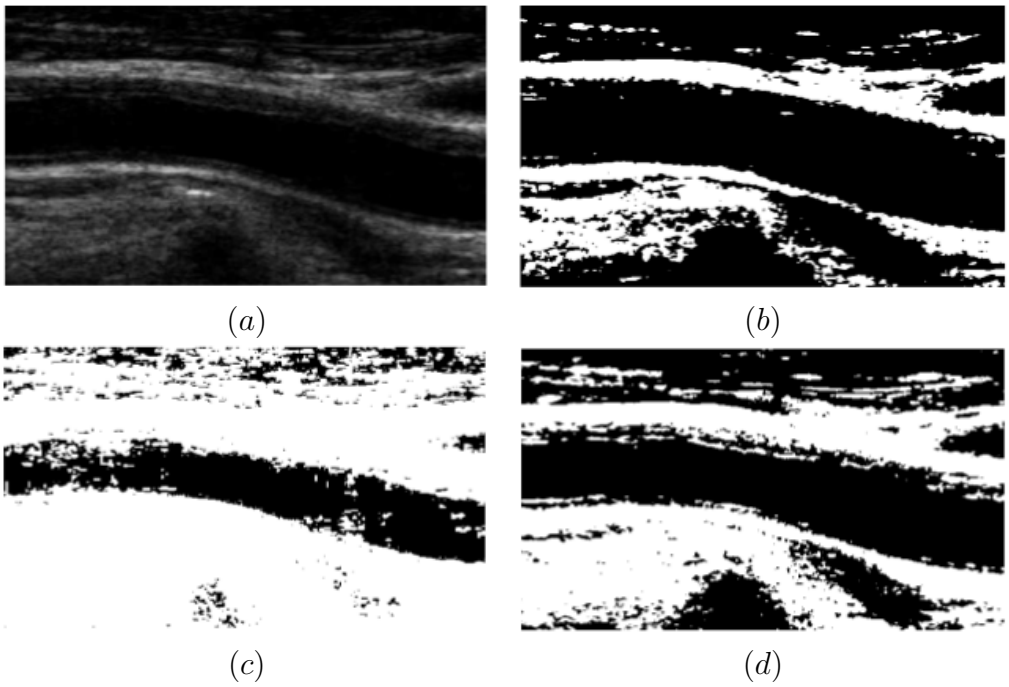


Figure 3.1: Application of several general thresholding segmentation algorithms to B-scan carotid ultrasound images[14]: a) Original 2D ultrasound image; b) Segmentation using Otsu's bimodal thresholding[42]; c) Segmentation using Kittler's Gaussian mixture method[25]; d) Segmentation using Minimum Cross Entropy (MCE) thresholding[29]

ysis of the image pixels, frequently resulting in a non-linear decision boundary in the feature space. For example, artificial neural networks have been used to segment spatio-temporal sequences of echocardiographic images[2]. The structure of the two-layer ANN is depicted in Figure 3.2 subplot (d) and it shows that the inputs are local image features. This is a rare example of an ultrasound segmentation method that does not take into account spatial consistency of the delineation, and it illustrates the chief resulting problem: extreme noisiness of the delineation, as exemplified by subplot (b) in Figure 3.2. The boundary of interest (endocardium) is present, but is quite noisy and obscured by the presence of multiple other objects, resulting either from other anatomical structures or simply having the same values of intensity and other derived photometric features. Post processing is required to obtain a good delineation as shown in subplot (c) of the figure.

The fact that decision boundaries defined in the feature space are not designed to expressly account for the spatial relationships between pixels results in a number of unwanted artifacts in the resulting delineations. Although there are various post-processing techniques which can overcome this difficulty, a much better approach is to acknowledge

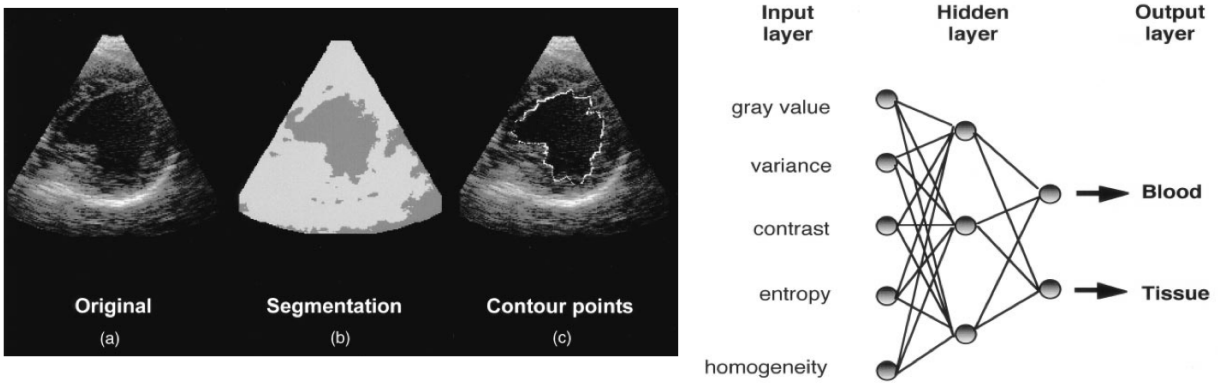


Figure 3.2: Endocardial contour detection using a 2-layer Artificial Neural Network: Net architecture and an example of segmentation result. (a) Original image. (b) Results of image segmentation performed by the ANN (within the region of interest); light grey color corresponds to tissue (myocardium) and dark grey to the blood pool region. (c) Contour point detection after the radial search algorithm was applied. (d) the architecture of the neural net.[2]

the intrinsically geometric nature of the problem from the beginning. For image modalities with better quality of data, such as CT and MRI, algebraic decision boundaries combined with post-processing are frequently sufficient, and may be preferred due to their relative simplicity and low computational complexity. However, with ultrasound data, which represents interfaces between regions better than it distinguishes various tissue types, such histogram-based approaches have a low chance of success. Most ultrasound segmentation algorithms are based on geometric decision boundaries - although they utilize the observed values of the features, the decision boundary is ultimately defined on the geometric coordinate grid.

Segmentation algorithms utilizing geometric decision boundaries impose some form of spatial constraints on the segmentation boundary. The most useful such constraint, controllable either implicitly or explicitly, is the regularity or smoothness of the solution. This is a desirable property in many segmentation tasks in which the ROI is known to be smoothly varying. Such a constraint can effectively reject the noise present in the data.

3.2.2 Energy minimization framework

Generally, a segmentation method amalgamates the various sources of information, as well as constraints on them, in a single mathematical descriptor. Many algorithms are conveniently cast in an energy minimization framework. For an image $I(\mathbf{x}), \mathbf{x} \in \Omega$, there

are many possible decision boundaries (contours) C delineating Ω into sub-domains Ω_A and Ω_B. These decision boundaries are geometrical objects and can be thought of as curves in 2D or surfaces in 3D. A particular segmentation approach defines an energy or cost function $E(C)$ which quantifies the suitability of C under a variety of criteria. The smaller the "energy" of C, the more desirable it is as the final segmentation result. Since C is a continuous geometric object, $E(C)$ is often called an energy functional. The goal of an image segmentation procedure then is to find the best possible curve by minimizing $E(C)$:

$$C^* = \arg\min_C E(C). \tag{3.2}$$

There are various approaches at solving this optimization problem, as well as the definitions of the curve C itself. A variety of constraints can be incorporated into an energy functional $E(C)$ in a form of energies, in which case $E(C)$ becomes a weighted sum of a number of components:

$$E(C) = \sum_k w_k E_k(C), \tag{3.3}$$

where w_k represent the tuneable weights. These approaches frequently require one to define the notion of the "derivative" of E with respect to the continuous curve C. Following the standard nomenclature of variational calculus, which deals with functionals whose inputs are continuous functions, this derivative is known as the *first variation* and is denoted as $\frac{\delta E}{\delta C}$.

These "sub-energies" can be broadly grouped into two types. Internal energies impose a variety of geometric constraints on the shape of the decision boundary itself. The most common of these is a "regularization" energy measuring the length of the curve. Minimizing this length will tend to make the decision boundary smooth. These energies bear no relation to the data contained in the image $I(\mathbf{x})$ itself. External energies, on the other hand, incorporate this observed information into the segmentation procedure.

The vast majority of modern segmentation approaches dealing with ultrasound imaging can be considered as iterative energy minimization procedures of the type defined in (3.2). Thus, it is instructive to examine in detail the possible energies of this form and their applicability to carotid ultrasound imaging.

3.2.3 Photometric information type

All segmentation algorithms analyse the values of the observed image $I(\mathbf{x})$, whether as a scalar intensity value by itself or as a feature vector. The different ways to incorporate this information may be subdivided into two types: an analysis of the statistical properties of all image pixels within a common region, or an analysis of transitions between different regions, or edges.

The edge-based image segmentation approach is widely used in medical imaging[39]. At first glance, it is especially justifiable for ultrasonic data, since an ultrasound image most strongly depicts tissue boundaries, and can thus be considered as an image of edges. However, there is a drawback: due to shadowing and orientation dependency artifacts previously discussed, those edges are weak or discontinuous and may not be sufficient for accurate segmentation.

A popular approach to incorporating edge information in an energy-minimizing segmentation setup is by introducing a *geodesic length* energy. If (some) edges in an image overlap with the boundary of the object to be segmented, then it is desirable that the segmentation boundary overlaps with edges as much as possible. The geodesic length measures such an overlap:

$$E_G(C) = \int_C g(I(s))\, ds, \tag{3.4}$$

where g is an edge-detector function proportional to the strength of the edges in I. This formulation incorporates two (sometimes competing) objectives: the overall length of the curve C should not be very large, and it should overlap with image edges as much as possible. Application of the geodesic length energy to image segmentation problems is frequently done in the geodesic active contours formulation developed by Caselles *et al.* in [7], or in the level set formulation for which efficient solutions are possible[18]. In these works, the algorithm is demonstrated using simple definitions of the geodesic function g, for example, defined in [10] as

$$g(I) = \frac{1}{1 + |\nabla I(\mathbf{x})|^P}, \quad P \geq 1, \tag{3.5}$$

where $|\nabla I(\mathbf{x})|$ is the gradient magnitude of a scalar-valued intensity image $I(\mathbf{x})$. However, more involved edge functions are possible, including replacing $I(\mathbf{x})$ in (3.5) by a pre-processed image with sharper, less noisy edges.

Commonly cited shortcomings of approaches based purely on edge information, such as the geodesic active contours, is their tendency to get trapped in local minima which do not correspond to desirable delineations, as well as their being heavily initialization-dependent. Both of these problems stem from noise contamination[11].

In addition to purely edge-based energies, information from all pixels belonging to a segmented region can be used[10]. Statistics of the noise contaminating the image are frequently used to distinguish between different regions inside the image. Each "true" sub-region Ω_j of the image domain Ω is associated with a probability density function (pdf) $p_j(\mathbf{z})$. For a given segmentation boundary C, one can also empirically estimate two pdfs P_A and P_B from pixels belonging to current estimates of Ω_A and Ω_B respectively. It is desirable to find a set of features to include in $I(\mathbf{x})$ such that densities P_A and P_B of the two sub-regions Ω_A and Ω_B are maximally *distant* from each other when C properly

delineates the image. Such an optimal segmentation contour C will minimize a measure of similarity between P_A and P_B.

The energies measuring similarity between empirical pdf estimates of different regions can be distinguished by whether they assume a *parametric model* for these pdf's. For example, a frequently made assumption that the image $I(\mathbf{x})$ is contaminated by additive white Gaussian noise (AWGN), and that samples of this noise are independent, leads to a well known minimal variance energy of Chan and Vese[10]. The energy is defined as

$$E_{CV}(C) = \int_{\Omega_A} (I(\mathbf{x}) - \mu_A)^2 d\mathbf{x} + \int_{\Omega_B} (I(\mathbf{x}) - \mu_B)^2 d\mathbf{x}, \quad \mu_k = \frac{\int_{\Omega_k} I(\mathbf{x}) d\mathbf{x}}{\int_{\Omega_k} d\mathbf{x}} \tag{3.6}$$

where μ_k is the mean vector, $k \in \{A, B\}$. The advantages of the region-based approach of the Chan-Vese energy (augmented in the original paper by a curve-shortening energy like (3.4) with $g = 1$) are the improved noise performance, successful segmentation with weak or discontinuous edges, and no dependence on initialization[10].

The above result can be considered a special case of *Maximum-Likelihood (ML)* estimation of the segmentation boundary[33]. A likelihood function expresses the conditional probability of observing the image I given a set of model parameters:

$$L[I \mid C, \theta_A, \theta_B] = P(\chi_A \,|\, C, \theta_A) \cdot P(\chi_B \mid C, \theta_B), \quad P(\chi_j | C, \theta_j) = \prod_{\mathbf{x} \in \Omega_j} p(\mathbf{x}|\theta_j), \tag{3.7}$$

where $j \in \{A, B\}$, $\chi_j = \{I(\mathbf{x}) \mid \mathbf{x} \in \Omega_j\}$, and θ_j is a vector of parameters specifying the assumed pdf model. It is assumed that the pixel samples in the image I are independent. The segmentation algorithm then amounts to minimizing the negative log-likelihood, and thus maximizing the log-likelihood:

$$C^* = \arg\min_C \left(-\log L[I \mid C, \theta_A, \theta_B] \right). \tag{3.8}$$

If the image pixels are assumed to be contaminated with AWGN and the variance is constant throughout the image, this formulation reduces to that of Chan and Vese in (3.6)[11]. However, the setup in (3.7-3.8) can generally be applied to images under the assumption of a variety of noise models, as is detailed in [33].

An alternative way to segment an image based on inter-regional differences in statistical distributions is to use a non-parametric empirical estimate of the density functions. Using a positive scalar-valued *kernel function* $K(\cdot)$ of unit integral, an empirical density can be computed as

$$P_j(\mathbf{z}|\Omega_j) = \frac{\int_{\Omega_j} K(\mathbf{z} - I(\mathbf{x})) d\mathbf{x}}{\int_{\Omega_j} d\mathbf{x}}, j \in \{A, B\} \tag{3.9}$$

where $\mathbf{z}$ stands for the vector of d features, considered random quantities. The biggest advantage of this approach is its generality; it does not impose any parametric statistical model on the image regions. The equation in (3.9) defines a smoothed and normalized histogram, dependent only on the definition of the kernel function $K(\cdot)$. Even though the computational cost is increased due to having to operate on the whole empirical density instead of the distribution parameters, such a non-parametric definition generalizes the approach to all distribution types.

A segmentation algorithm using this definition of the densities can be defined using the ML framework in (3.8). An alternative, as used in [35], is to find an optimal segmentation curve C^* which results in the maximum possible discrepancy between P_A and P_B in terms of the *Bhattacharyya metric* defined by

$$E_{bh}(C) = \int_{\mathbf{z}\in\mathbb{R}^d} \sqrt{P_A(\mathbf{z})P_B(\mathbf{z})}\; d\mathbf{z}, \tag{3.10}$$

where integration over z proceeds over all the values that the image I can attain. This energy measures the similarity between P_A and P_B; it is proportional to the degree of correlation between the two empirical densities.

Each of the two ways of incorporating photometric information into a segmentation framework has its distinct advantages and disadvantages. For the purposes of ultrasound segmentation, the presence of edges is a useful indication of the boundary between the carotid artery and the surrounding tissue. Such edges are, however, weak and discontinuous, and relying on an edge-based energy alone would not be sufficient. On the other hand, using regional statistics alone would ignore this valuable source of discriminatory information when it is present. In the present work, a hybrid approach is used, with a linear combination of an edge-based and a region-based energy.

3.2.4 Prior knowledge

An alternative statistical estimation methodology for region-based segmentation is obtained through the *Maximum A-Posteriori (MAP)* statistical framework. Unlike the ML estimation in (3.7), the a-posteriori probability of the parameters given the observed image is used to estimate an optimal contour as

$$C^* = \arg\max_C \; P(C \mid I) = P(I \mid C) \cdot P(C), \tag{3.11}$$

where $P(C \mid I)$ and $P(C)$ are respectively the posterior and prior probabilities, and $P(I \mid C)$ is the likelihood function. A notable change from the ML approach is the presence of a *prior* probability term $P(C)$. This enables convenient incorporation of prior knowledge about the segmentation outcome. Such prior information can be incorporated in a variety

of forms. Some priors are obtained *statistically* from a training set of manually segmented images. This methodology encodes prior knowledge about the object of interest in the form of probability distributions. The information can be photometric (e.g., distributions of the image values inside and outside the ROI), or it can be geometric (e.g., statistical shape priors). Alternatively, rather than obtaining the prior knowledge by means of machine learning, it can be encoded explicitly by means of general *geometric priors*.

Statistically learned priors are employed when a large number of manually segmented training examples are available. From this training set, a representative template is retrieved of either the distribution of pixels inside the object, or its shape. During the segmentation, a prior energy is introduced which penalizes deviations from the template. For example, the pixel distribution inside an object can be non-parametrically estimated by using a histogram-smoothing method of (3.9) for each image in the training set. The empirical densities can then be combined by, for example, averaging, to create a template empirical p.d.f. of pixels inside an object of interest, $P_T(\mathbf{z})$. The same Bhattacharyya distance measure as given by (3.10) is used, except it now compares the template density with that of the current object. This energy will enforce maximal overlap between the two.

Alternatively, the object shape itself can be subjected to a statistically learned constraint. There is a multitude of segmentation approaches in which such statistical shape priors are employed. In particular, it has been applied toward segmenting medical imagery data (2D cardiac and 3D prostate MRI) by means of principal components analysis (PCA) on a level-set representation of the segmentation contour[59]. Although the introduction of a-priori shape information into the segmentation algorithm improves the segmentation results, it nevertheless suffers from several drawbacks. Firstly, the shape model must be aligned with the segmentation boundary at each iteration of the segmentation routine, introducing additional complexity. Furthermore, it usually imposes a strong constraint on the shape of the segmentation boundary or the distribution of pixels within it, which may not be flexible enough to be applied to some medical imaging data in which the inter-patient variability is high. One of such types of data is the carotid ultrasound. Age, health and other biological factors increase the inter-patient variability, while the nature of ultrasound increases variability between scans. Lastly, the efficacy of this approach depends heavily on the quality of the training set, usually manually segmented by many experts and thus suffering from both inter- and intra-observer variability.

One way to deviate from "strong" shape priors is to impose a constraint on the shape of the object in a weaker fashion. A combination of geometric shape priors and density-tracking has been used in [72] by extracting a distribution of curvature along the boundary of the prostate gland in 2D trans-rectal ultrasound (TRUS) images[72]. This approach removes the need for alignment since curvature is invariant under the group of Euclidean transformations. Additionally, using the distribution of curvature is minimally restrictive, which allows for more variability than more restrictive approaches such as those discussed

in [59].

3.3 Contour representation

Within an energy-minimization framework, there exist a number of ways in which the contour C is represented in a segmentation algorithm. They differ considerably in their computational complexity, limitations and the ease with which the geometric transformations of the contours are handled.

One of the most popular segmentation approaches is the use of active contour models also known as *snakes*. In two dimensions, the delineation curve C is represented parametrically as

$$C : [0, 1] \rightarrow \Omega, C(p) \equiv \{x(p), y(p)\}, 0 \leq p \leq 1. \tag{3.12}$$

This parametric curve is deformable, with its dynamic behaviour governed by a combination of internal and external forces as in (3.3). The curve is typically represented using splines, as in its first popular implementation by Kass et al. [24]. The dynamic behaviour is imposed on a set of control points, with the curve between the points defined by means of spline polynomial functions.

This approach is very attractive from the viewpoint of numerical efficiency: only a finite set of control points is propagated around the image, with the rest of the curve being implicitly defined. Moreover, the use of spline polynomial functions imposes smoothness and continuity constraints, and thus restricts the variation of possible curve configurations to that of curves with a pre-defined amount of elasticity[3]. However, there are a number of caveats. Firstly, undesirable behaviour of the curve such as self-intersection or excessive irregularity of the spacing of control points (too close together, too far apart) require "non-elegant" regularization steps such as re-gridding. More importantly, snakes require cumbersome additional steps to handle geometric transformations such as topological changes (splitting and joining of closed curves)[11]. In the context of carotid artery segmentation, such a topological change occurs at the bifurcation area - the common carotid artery splits into two. The extra steps required to handle these problems become even more cumbersome when the segmentation is performed on 3D images.

An alternative approach which implicitly handles the above issues is the *level set method*. Its implicit definition of the segmentation curve as a zero level set of a 3D function allows it to avoid the need for re-parameterization and special handling of topological changes. This is the curve evolution method used in the rest of this thesis, and is explained in detail in Section 6.1. Before describing its implementation and use, the next two chapters examine other important aspects of the proposed segmentation algorithm.

Chapter 4

Image features

The extraction of relevant features forms an important part of many image processing tasks. In particular, a variety of possible transformations of the observed intensity image can be used as features by the segmentation algorithm. Firstly, a number of them can be combined into a feature image $I(\mathbf{x})$ and used in region-based segmentation energies. In addition, some features which more faithfully represent the coarse structure of the imaged tissue can be used in the geodesic edge-detecting functions defined in (3.5).

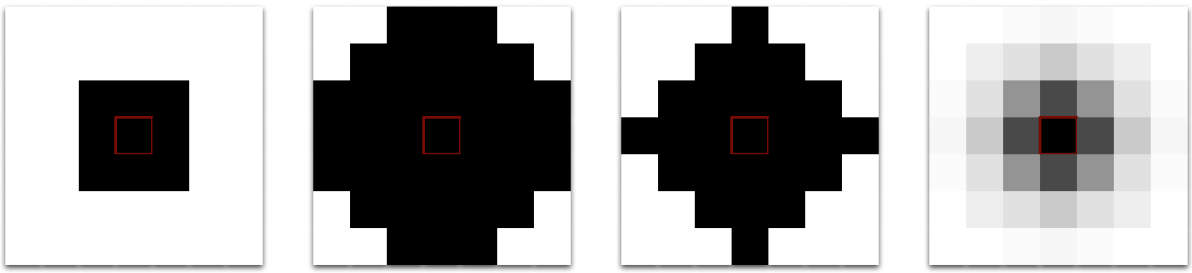
Although a wide array of possible image features exists, this work focuses on two major types. Firstly, statistical measures may be computed over a finite spatial neighbourhood around each pixel. Secondly, non-linear filtering techniques, including diffusion filtering, may be applied to obtain regularized images while preserving their edges.

4.1 Local statistics

The input to the regional segmentation energy is a set of features in which the region of interest is as statistically distinct from the background as possible. One method in which such features may be obtained is by examining the local behaviour of the image statistics in a spatial neighbourhood around each pixel location:

$$\bar{I} : \mathbf{x} \mapsto f(\mathbf{x} \mid \mathcal{N}(\mathbf{x})), \tag{4.1}$$

where $\mathcal{N}(\mathbf{x})$ is a set of pixel locations comprising a spatial neighbourhood of $\mathbf{x}$ and $f(\mathbf{x})$ is a local statistic computed based on this neighbourhood set. Examples of spatial neighbourhoods in two dimensions are given in Figure 4.1. The same principle can be straightforwardly extended to three dimensions. Aside from these mask-type neighbourhood definitions, one can define a neighbourhood where each location $j \in \mathcal{N}$ is assigned a weight $w_j \in [0, 1]$, as shown in Figure 4.1 subplot (d).



(a) (b) (c) (d)

Figure 4.1: Examples of spatial neighbourhoods in two dimensions: a) a 3-by-3 square neighbourhood (8 nearest neighbours); b) a circular neighbourhood; c) a "diamond" neighbourhood; d) Gaussian-weighted neighbourhood;

One possible set of local statistics which could be computed on such neighbourhoods is the set of first and second moments (i.e., the local mean and variance):

$$f_\mu(\mathbf{x}) = \frac{1}{|\{\mathcal{N}(\mathbf{x})|} \sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} I(\mathbf{y}) \tag{4.2}$$

$$f_v(\mathbf{x}) = \frac{1}{|\mathcal{N}(\mathbf{x})|} \sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} |I(\mathbf{y}) - f_\mu(\mathbf{x})|^2, \tag{4.3}$$

where $|\mathcal{N}|$ is the size of $\mathcal{N}$. As illustrated in Figure 7.1, this variance emphasizes image edges and thus can potentially be used for edge detection. In the presence of outliers, such as in the ultrasound images, a robust version of these two statistics can be expressed by

$$f_{median}(\mathbf{x}) = \text{median}_{\mathbf{y}\in\mathcal{N}(\mathbf{x})}(I(\mathbf{y})), \tag{4.4}$$

$$f_{MAD}(\mathbf{x}) = \text{median}_{\mathbf{y}\in\mathcal{N}(\mathbf{x})}(|I(\mathbf{y}_i) - f_{median}(\mathbf{x})|), \tag{4.5}$$

where "MAD" stands for Median Absolute Deviation. Another statistical measure that can be obtained is the local entropy:

$$f_e(\mathbf{x}) = -\sum_{\mathbf{z}} p_{\mathbf{x}}(\mathbf{z}) \log p_{\mathbf{x}}(\mathbf{z}), \tag{4.6}$$

where $p_{\mathbf{x}}$ is the local distribution of values of I over the neighbourhood of $\mathbf{x}$,

$$p_{\mathbf{x}}(\mathbf{z}) = \frac{1}{\#\{\mathcal{N}(\mathbf{x})\}} \sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} \delta(I(\mathbf{y}) - \mathbf{z}). \tag{4.7}$$

Local entropy measures how structured the local p.d.f. is, and is thus related to variance. It has been proposed for use in ultrasound segmentation in [75], where it was shown to improve contrast in the ultrasound image, and thus provide better discriminability for the chosen threshold-based segmentation approach. The benefit of using this measure is its invariance to such transformations as log compression, as well as changes in the gray level of a particular region in the image due to acoustic shadowing[75]. The local entropy has also been applied as the weighting term for adaptive median filtering of ultrasound images[73]. Such filtering is meant to preserve image details, as indicated by high local entropy, while smoothing homogenous regions more strongly.

4.2 Non-linear filtering

A variety of non-linear filtering techniques have been devised to pre-process images. Some of these techniques have been specifically adapted to ultrasonic imaging, while others are general-purpose but nevertheless useful.

The most popular way to remove noise from a d-dimensional image I is to filter it with a linear low-pass filter. A frequently used example of the latter is the Gaussian filter, which can be defined in the following way:

$$\bar{I}(\mathbf{x}) = \sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} G_\sigma(\mathbf{x}-\mathbf{y}) \cdot I(\mathbf{y}), \tag{4.8}$$

where $G_\sigma(\cdot)$ is a Gaussian kernel, in which the degree of smoothing (the cutoff frequency of the low-pass filter) is proportional to the standard deviation σ:

$$G_\sigma(x) = \frac{1}{(2\pi\sigma^2)^{d/2}} \exp(-\frac{|x|^2}{2\sigma^2}). \tag{4.9}$$

Although this type of filter succeeds in removing the high-frequency noise, this is done at the expense of smoothing image edges by removing high-frequency components of transitions in the image. As a result, more involved denoising techniques have been created, which do not overly smooth the edges at the expense of complexity, and, frequently, non-linearity. One such technique is a conceptually straightforward extension of the linear filter in (4.8) known as the *bilateral filter*, which is defined as follows:

$$\bar{I}(\mathbf{x}) = \frac{1}{W_\mathbf{x}} \sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} G_{\sigma_s}(\mathbf{x}-\mathbf{y}) \cdot G_{\sigma_r}(I(\mathbf{x})-I(\mathbf{y})) \cdot I(\mathbf{y}), \tag{4.10}$$

$$W_\mathbf{x} = \sum_{\mathbf{y}\in\mathcal{N}(\mathbf{x})} G_{\sigma_s}(\mathbf{x}-\mathbf{y}) \cdot G_{\sigma_r}(I(\mathbf{x})-I(\mathbf{y})), \tag{4.11}$$

where $W_{\mathbf{x}}$ is the normalization factor which preserves the energy. In addition to low-pass filtering in the spatial domain as the linear filter does, the bilateral filter also takes the pixel values themselves into account by also applying the filtering operation in the range (intensity) domain (illustration). Multiplication by $G_{\sigma_r}(I(\mathbf{x}) - I(\mathbf{y}))$ in (4.10) makes the contribution of $I(\mathbf{y})$ to the final output at $\mathbf{x}$ dependent on how close the intensity values at $\mathbf{x}$ and $\mathbf{y}$ are to each other. Since its introduction[5], the bilateral filter has attracted considerable attention due to its robust performance resulting in piecewise-smooth denoised images with preserved edges. Although conceptually simple, the straighforward implementation of (4.10) is very slow; however, a significantly faster implementation has been developed, exploiting signal processing concepts to arrive at a good approximation of the original filter[44]. Furthermore, although constructed here with Gaussian functions G_{σ_s} and G_{σ_r}, the choice of the weighting function is general and includes, for example, the simple box function[68].

4.2.1 Diffusion filtering

Another major means of obtaining images with reduced noise and preserved edges is by means of non-linear diffusion filtering. These methods are based on solutions of partial differential equations derived from the heat equation, a physical description of the diffusion of heat in materials.

This approach looks at image filtering as a process occurring over time, and thus a non-linear diffusion filtering process results in a sequence $I(\mathbf{x}, t)$ of solutions to the diffusion equation,

$$\frac{\partial I}{\partial t} = \operatorname{div}(D_I \nabla I), \tag{4.12}$$

where t is the artificial parameter of "time" since initialization, D_I is an image-dependent diffusivity, and the diffusion process is initialized by the original image I_0:

$$I(\mathbf{x}, 0) = I_0. \tag{4.13}$$

In order for this diffusion process to be energy-preserving, the following boundary conditions are imposed:

$$\nabla I \cdot \vec{\mathbf{n}} = 0 \text{ at } \partial\Omega = 0, \tag{4.14}$$

s where $\vec{\mathbf{n}}$ is the direction normal to the image boundary $\partial\Omega$.

The diffusivity term D_I in (4.12) can be defined in a number of ways, resulting in conceptually different diffusion results. If D_I is a scalar constant, the diffusion process is linear, reducing simply to the Gaussian filter with standard deviation $\sigma = \sqrt{2t}$[65]. If, on the other hand, D_I is a spatially-varying scalar function, the result can be described as isotropic and non-homogenous diffusion - the diffusion at each pixel proceeds equally in

all directions, but its strength varies from pixel to pixel. Although in this case $D_I(\mathbf{x})$ is a general function, it is usually defined to reduce the amount of smoothing on the edges. A typical indicator of the presence of image edges is the magnitude of the gradient, denoted $|\nabla I|$. It is used to compute the diffusion strength, as originally defined by Perona and Malik[45]:

$$D_I(\mathbf{x}) = g(|\nabla I|) \in [0, 1], \tag{4.15}$$

where the scalar diffusivity function $g(\cdot)$ is designed to approach 1 in homogenous regions where no edges of interest are present, while approaching 0 when overlapping with edges. The computation of the gradient in practical images is performed using finite differencing, which amplifies any noise present in the image. In order to extract useful edge information from the gradient, as well as its magnitude, it is often computed on a pre-smoothed version of the image:

$$\nabla_\sigma I = \nabla(G_\sigma * I) = G_\sigma * \nabla I, \tag{4.16}$$

where G_σ is defined as in (4.9).

One example of such a diffusivity function in (4.15) is the following exponential function [67]:

$$g(s) = \begin{cases} 1 & s \leq 0 \\ 1 - \exp\left(\frac{-3.315}{(s/\lambda)^4}\right) & s > 0 \end{cases} \tag{4.17}$$

The parameter λ allows selection of a scale for edge preservation: edge-like structures for which $|\nabla I| > \lambda$ will be preserved ($g(s) \to 0$), while those for which $|\nabla I| > \lambda$ will be smoothed ($g(s) \to 1$). Other definitions of $g(s)$ for general images have similar behaviour, with slightly different transitions[10, 45].

Due to the unique characteristics of acoustic images, an application of non-homogenous diffusion filtering to ultrasound data performs better with custom approaches. Specifically, the diffusivity function $g(s)$ needs to be adapted to selectively preserving or removing the speckle pattern. The speckle-reducing anisotropic diffusion (SRAD) approach[74] achieves this by defining an adaptive speckle-detecting function. This adaptive function depends on an instantaneous coefficient of variation, which performs edge detection in speckle-contaminated imagery[54]:

$$q(\mathbf{x}, t) = \sqrt{\frac{(1/d)(|\nabla_\sigma I|/I)^2 - (1/2d)^2(\nabla^2 I/I)^2}{(1 + (1/2d)(\nabla^2 I/I))^2}}. \tag{4.18}$$

The normalization of the gradient magnitude $|\nabla_\sigma I|$ and the Laplacian operator $\nabla^2 I$ by the image intensity in (4.18) adapts this edge detector to the variations in average intensity – speckle will be detected in dark and bright regions. Additionally, a speckle scale function, indicating the "average" speckle dimensions, can either be found as a function of mean

and variance over a homogenous region of I[74], or be defined as discussed in [54]:

$$q_0(t) = \frac{\lambda \int_\Omega q^2(\mathbf{x}, t) d\mathbf{x}}{\int_\Omega d\mathbf{x}}, \tag{4.19}$$

where the parameter λ controls the scale selection in a similar manner to (4.17). The diffusivity function $g(s)$ is then defined as:

$$g(s) = \frac{1}{1 + [q^2(\mathbf{x}, t) - q_0^2(t)]/[q_0^2(t)(1 + q_0^2(t))]}. \tag{4.20}$$

This diffusion method extracts piecewise continuous regions of the ultrasound image by selectively removing the speckle pattern depending on its scale, while attempting to preserve anatomically relevant (non-speckle) edges.

A particularly useful non-linear denoising algorithm which can be cast into the diffusion filtering framework is *total variation* (TV) denoising[49]. This approach aims to recover a piecewise-smooth image from its noise-contaminated version by minimizing its total variation semi-norm[1], which for a 2D image is given by

$$\|I\|_{TV} := \int_\Omega |\nabla I| \, d\mathbf{x}. \tag{4.21}$$

This optimization problem leads to a solution via diffusion filtering as in (4.12), with the scalar diffusivity function given by[65]

$$D_I = \frac{1}{\|\nabla I\|}. \tag{4.22}$$

The strength of TV denoising diffusion is inversely proportional to the magnitude of the image gradient, and thus is small along the edges and stronger in more homogenous regions. As a result, the reconstructed images appear piecewise smooth.

The above definition of diffusivity D_I as a scalar function results in diffusion which is *isotropic* at all pixel locations – it proceeds with the same strength in all directions. An *anisotropic* diffusion process does not just preserve image edges, but can also be used for their enhancement. In order to accomplish this, diffusion needs to be directed along structures representing interfaces between regions, such as edges in 2D or surfaces in 3D, and inhibited in the orthogonal direction, along which the transition occurs. The diffusivity D_I is now defined as a tensor, and the term $D_I \nabla I$ is a matrix-vector product. Diffusivity D_I is defined using the *structure tensor*, a second-moment matrix which encodes the local coherence and orientation of the image transitions[66]. This matrix is based on the outer

[1]Generally speaking, this definition is applicable to differentiable images.

product of the pre-smoothed image gradient, which is further smoothed component-wise with another Gaussian kernel with standard deviation ρ:

$$S_I = G_\rho * (\nabla_\sigma I \otimes \nabla_\sigma I) \tag{4.23}$$

Due to smoothing the gradient with G_σ, and smoothing the components of the outer product with G_ρ, the structure tensor $S_I(\mathbf{x})$ integrates information from a local neighbourhood of $\mathbf{x}$. The eigenvectors and eigenvalues of this matrix now contain spatial orientation information at $\mathbf{x}$ and can be used to construct a tensor with a pre-defined degree of diffusion in the direction of largest change (perpendicular to the edge) and the direction(s) orthogonal to it. Additionally, the degree of homogeneity of the neighbourhood of $\mathbf{x}$ (i.e. absence of edges) can be estimated by looking at the differences between the size of the eigenvalues – approximately equal eigenvalues indicate homogeneity[66].

4.3 Conclusions

In the preceding chapters, the unique characteristics of ultrasound images were discussed, as well as various techniques for segmenting them. In this chapter, a number of possible definitions of image features were outlined, which can be used both for segmentation and edge detection. Next, the research part of this work is presented, starting with the main contributions and implementation details and ending with results and conclusions.

Chapter 5

Main contributions

Two chief contributions are presented in this thesis. Firstly, a novel approach aiming to improve the quality of features extracted from 3D carotid ultrasound images is proposed. The image is first decomposed into statistically distinct components based on the MAP framework incorporating signal and noise models which are appropriate for the ultrasound setting. Secondly, a novel form of incorporating the prior knowledge about the artery shape into the segmentation algorithm is used. The convexity shape prior is a weak prior energy which encourages the segmentation boundary to be almost everywhere convex, while not being overly restrictive. This strategy stems from the inherent variability of the carotid artery shapes in the 3D ultrasound data, which a "strong" prior energy (for example statistical shape priors learned from a training set of manual segmentations) may not be able to adequately encode, leading to misclassification errors. Both of these contributions are combined into the segmentation algorithm discussed in detail in the next chapter.

5.1 Ultrasound image decomposition

In a preprocessing stage of a typical segmentation algorithm, extraction of features is performed in order to improve the discriminability of the regions of interest with respect to their background. As previously discussed in Chapter 4, local statistics such as geometric moments or entropy are routinely used in the feature selection step. However, images are generally complex combinations of components possessing different statistical properties. In this work, it is proposed to first decompose the ultrasound image into statistically distinct components, and then to use these components separately as inputs to the feature extraction step, or as features in their own right, as outlined in Figure 5.1.

As discussed before, the speckle pattern contaminating the ultrasound image can drastically reduce the discernibility of the tissues of interest in the extracted image features.

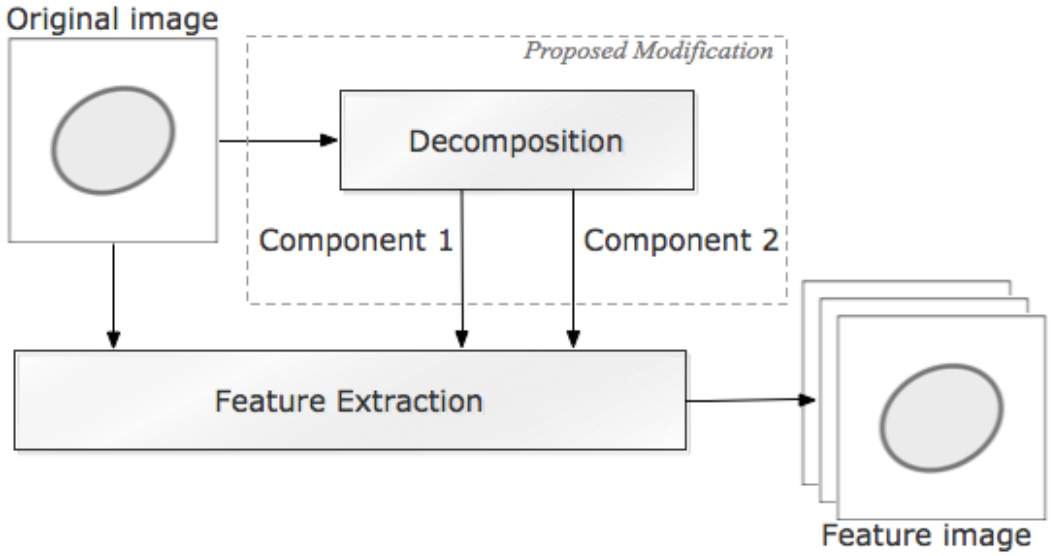


Figure 5.1: Feature extraction after decomposition: complex observed image is separated into statistically distinct components, which are used separately during feature extraction. Other features may be extracted from the original image directly.

Anatomically relevant details may be rendered unintelligible by the subsequent segmentation algorithm. However, the nature of the speckle pattern varies depending on the echogenicity of the underlying tissue. Thus, it is not desirable to simply discard speckle as a noise to be filtered out. In fact, despite being a realization of a random process as described in Chapter 1, it contains information which can be useful in differentiating between different anatomical tissues[36]. The other major component of the ultrasound image is a piecewise smooth rendering of the underlying tissue. Separately, the speckle and the piecewise smooth components are both useful, but together they interfere with and obscure each other, reducing the segmentation quality.

Therefore, an ultrasound-specific image decomposition algorithm is proposed here. It is designed to extract three components from the image. Each of these components belongs to distinct functional and statistical classes and may be pertinent to selectivity of different anatomical details. The piecewise smooth component represents the regional echogenicity of the anatomical structures. The speckle component is found by utilizing the statistical model presented in Section 2.3. Finally, the residual component accounts for any phenomena not adequately accounted for by the previous two components.

The proposed method uses the multiplicative model of speckle formation in (2.5), ap-

plying the logarithm operator to arrive at an additive log-speckle model as in (2.6). Thus, the model used as a basis for the decomposition algorithm,

$$g_l = u_l + v_l + e_l, \tag{5.1}$$

consists of the speckle-free piecewise smooth component u_l, the speckle pattern component v_l and the residual e_l.

5.1.1 Decomposition in the MAP estimation framework

The proposed decomposition method is developed using the Bayesian framework of MAP estimation. This statistical approach allows one to use a-priori knowledge about the quantity of interest, which typically improves the stability and robustness of the overall estimation procedure. The estimates of u_l and v_l are defined as maximizers of the posterior probability $P(u_l, v_l \mid g_l)$, which according to the Bayes theorem can be expressed as

$$P(u_l, v_l \mid g_l) \propto P(g_l \mid u_l, v_l)\, P(u_l)\, P(v_l), \tag{5.2}$$

where $P(g_l \mid u_l, v_l)$ is a likelihood term which describes a discrepancy between g_l and $u_l + v_l$ (referred to as e_l in 5.1) and u_l and g_l have been assumed to be conditionally independent. A commonly used alternative to performing maximization of $P(u_l, v_l \mid g_l)$ is minimization of the negative log-likelihood given by

$$L(u_l, v_l) := -\log P(u_l, v_l \mid g_l) \propto -\log P(g_l \mid u_l, v_l) - \log P(u_l) - \log P(v_l). \tag{5.3}$$

Before specifying an algorithm for minimization of $L(u_l, v_l)$, each component in (5.3) needs to be explicitly defined. Specifically, the residual component e_l is assumed to be zero-mean Gaussian noise (an assumption that can be justified by the Central Limit Theorem), in which case its associated log-likelihood term is proportional to the following ℓ_2-norm

$$-\log P(g_l \mid u_l, v_l) \propto \|g_l - (u_l + v_l)\|_2^2. \tag{5.4}$$

The component u_l, an estimate of the logarithm of the piecewise smooth "true" image, can be generally explained by a Markovian model of the type[30]

$$\begin{aligned} P(u_l) &= \frac{1}{Z} \prod_{\mathbf{x}\in\Omega} P(u_l(\mathbf{x})), \text{ with} \\ P(u_l(\mathbf{x})) &= \exp\left(- \sum_{\mathbf{y}\in C(\mathbf{x})} V_c(u_l(\mathbf{y})) \right), \end{aligned} \tag{5.5}$$

where $C(\mathbf{x})$ represents a neighborhood of $\mathbf{x}$ (i.e., its clique), $V_c(\cdot)$ is a Gibbs distribution with interaction potential $\{V_c, c \in \mathcal{C}\}$, and Z is a partition function whose role is to normalize the probability in (5.5). In the ultrasound setup, the image u_l represents local echogenicity of soft tissues, which is expected to be a piecewise-smooth function. In this case, it seems reasonable to choose V_c as an edge-preserving potential, which leads to the following likelihood term:

$$-\log P(u_l) \propto \sum_{\mathbf{x}\in\Omega} |\nabla u_l(\mathbf{x})| = \sum_{\mathbf{x}\in\Omega} \sqrt{\left(\Delta_x u_l(\mathbf{x})\right)^2 + \left(\Delta_y u_l(\mathbf{x})\right)^2} = \|u_l\|_{TV}, \tag{5.6}$$

where $\mathbf{x} = (x, y)$ is a particular pixel location in the image u_l, $\Delta_x u_l(\mathbf{x})$ and $\Delta_y u_l(\mathbf{x})$ are finite difference approximations of partial derivatives, and $\|\cdot\|_{TV}$ stands for the total variation (semi-) norm. This particular norm is defined on a space of functions of bounded variation [49],

$$BV(\Omega) \;=\; \{f : \Omega \to \mathbb{R} \mid \|f\|_{TV} < \infty\}, \tag{5.7}$$

whose choice is further justified by the fact that projecting g_l onto $BV(\Omega)$ allows for preservation of discontinuities in the images while excluding noise and oscillatory components[9, 30].

Finally, the component v_l is defined as an estimate of the speckle pattern. As previously discussed in Section 2.3.1, if the samples of the underlying tissue reflectivity function follow a generalized Gamma distribution, the log-compressed speckle noise can be shown to follow the Fisher-Tippett distribution:

$$v_l \sim 2\exp\left(2\left[v_l - \log\alpha\right] - \exp\left(2\left[v_l - \log\alpha\right]\right)\right). \tag{5.8}$$

The parameter α of this distribution has the effect of shifting its mean by $\log(\alpha)$. This constant shift can instead be accounted for by the bounded-variation component u_l, which can be facilitated by setting $\log(\alpha)$ to zero (or setting α to unity). Consequently, assuming the pixels of v_l to be *i.i.d.*, the speckle noise component of the observed image is represented in this optimization problem by

$$-\log P(v_l) \propto \varphi(v_l) := \sum_{\mathbf{x}\in\Omega} \left(\exp 2v_l(\mathbf{x}) - 2v_l(\mathbf{x})\right). \tag{5.9}$$

The assumption of the i.i.d. on the pixels of v_l may be regarded as an oversimplification due to the band-limitedness of the ultrasound acquisition system, which results in correlation between neighboring pixels. One way to overcome this shortcoming is to apply a de-correlation procedure which results in an image with uncorrelated noise[36]. This procedure, omitted for simplicity and computational speed from this analysis, is a potential area for improvement.

Using the results (5.4), (5.6) and (5.9), the minimization of (5.3) is equivalent to

$$\arg\min_{u_l, v_l} E(u_l, v_l), \quad E(u_l, v_l) = \|g_l - (u_l + v_l)\|_2^2 + \lambda_1 \|u_l\|_{TV} + \lambda_2 \varphi(v_l), \tag{5.10}$$

where $\lambda_1 > 0,\ \lambda_2 > 0$ are related to parameters of the distributions that their likelihood terms are derived from. These parameters control the degree to which the optimization problem in (5.10) balances out the competing tasks of minimizing the residual component and enforcing the adherence of the u_l and v_l components to their respective models.

Note that the functional $E(u_l, v_l)$ defined in (5.10) is not convex, which considerably complicates the task of solving this problem in a stable and unique manner. However, if we regard u_l as constant and modify $E(u_l, v_l)$ accordingly, the result,

$$E_v(v_l \mid u_l) = \|g_l - u_l - v_l\|_2^2 + \lambda_2 \varphi(v_l), \tag{5.11}$$

is in fact convex in v_l. Likewise, holding v_l as constant, the functional

$$E_u(u_l \mid v_l) = \|g_l - u_l - v_l\|_2^2 + \lambda_1 \|u_l\|_{TV} \tag{5.12}$$

also becomes convex in u_l[49]. Minimizing these two functionals is considerably simpler and a stable estimate is guaranteed. Accordingly, the proposed ultrasound decomposition scheme makes use of the alternating minimization approach, which is based on sequentially minimizing (5.11) and (5.12). Even though this procedure does not guarantee convergence to the global minimum of (5.10), it does guarantee the reduction in its value at each step.

The implementation of this decomposition algorithm is discussed in Chapter 6, where the implementation details of the algorithms proposed in this thesis are shown.

5.2 Convexity as a weak geometric prior

Reliability of carotid segmentation can be substantially improved by supplementing the observed data with prior information, which can be defined in the form of either probabilistic or geometric models. Unfortunately, parametrically defined geometric models tend to over-constrain the segmentation. Moreover, their application becomes problematic whenever the segmentation boundary needs to undergo topological changes (which is always the case in the region of carotid bifurcation) [21, 37]. Moreover, most of the segmentation methods based on statistical shape analysis[59] require the use of large training sets. When applied to carotid ultrasound imagery, however, such methods face difficulties in dealing with high patient-to-patient variability of the data, which results from factors such as the clinical condition of the arteries, age or gender.

To constrain the convergence of segmentation to anatomically plausible shapes, the proposed prior energy term controls the *curvature* of the segmentation boundary. Intuitively, curvature is a measure of the deviation of a geometric object from being flat. In the case of a 2D curve, the curvature measures deviation of this curve from a straight line. At any point along the curve, this deviation may be visualized by placing a circle that most closely matches the curvedness in the vicinity of the point, as shown in Figure 5.2. curvature is inversely proportional to the radius of this osculating circle. The sign of curvature is also useful in differentiating the behaviour of the curve. When the curve is closed, the sign of curvature indicates whether the curve is *convex* or *concave*. These two concepts are illustrated in Figure 5.3

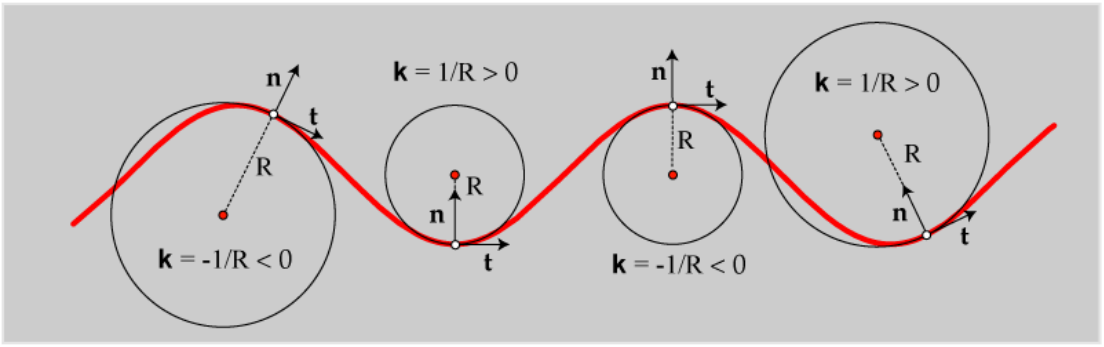


Figure 5.2: Illustration of the relationship between the signed curvature of a curve and its osculating circle[61].

Since no explicit assumptions on curve geometry are made, this method is referred to as a *weak geometric prior*. It should be noted that the latter is independent of various anatomical factors while being descriptive enough to improve the stability and reliability of the overall segmentation procedure. A further advantage of using curvature is that its measurement is invariant under Euclidean transformation (i.e. translation and rotation).

The development of the weak prior energy is rooted in the anatomy of arteries in general, and the carotid artery in particular. Due to the pressure applied to the arterial walls by blood, the cross-sectional shape of any artery is convex. In the transverse direction, the position of the carotid artery in particular is such that no concave regions are present, with the notable exception of the bifurcation itself. This is an idealizing assumption which, however, should not be violated if the patient's neck is properly positioned for the 3D ultrasound scan. On the other hand, any anatomically plausible segmentation of the bifurcation region itself will contain regions with negative curvature. Figure 5.4 depicts a simulated carotid artery bifurcation, as well as the sign of curvature in selected regions. An undesirable artifact, denoted as region A, departs from the anatomically plausible restriction that all cross-sections should be convex, as illustrated in the plots of Column A depicting the

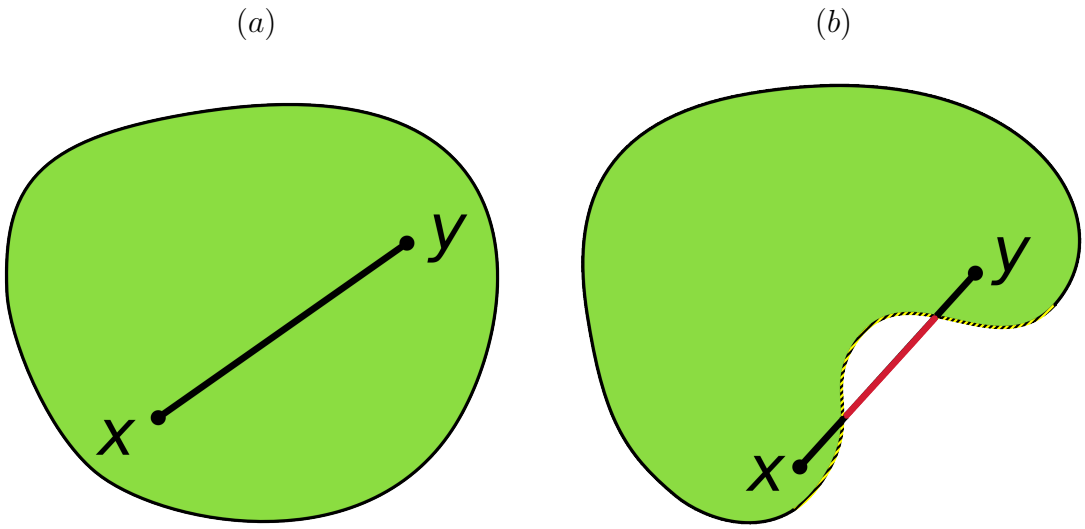


Figure 5.3: Convexity and concavity for objects in Euclidean space. (a) An object is convex if every line segment between points inside the object is contained entirely within the object; (b) When this is not the case, the object is concave.[70].

sign of curvature for selected cross-sections in region A. The bifurcation region, denoted as region B, also contains cross-sections with negative curvature, as shown in the plots of Column B. Thus, the shape of the carotid artery will contain a small region of negative curvature around the bifurcation, while remaining convex otherwise.

The weak prior energy proposed in this work encourages an anatomically plausible segmentation boundary of the carotid artery through the use of the sign of curvature. At the same time, it does not excessively penalize the shapes which contain small regions with negative curvature, since the carotid bifurcation invariably must be partially non-convex. Let $L(C)$ represent the length of a delineation curve C:

$$L(C) = \int_C ds, \tag{5.13}$$

where s is the arclength of curve C. In 3D, C is more accurately denoted as a surface, while $L(C)$ measures the surface area. Let $L_C(C)$ denote the length of only the portion of the curve C which is convex. This measure can be readily made by examining the sign of the curvature κ along C:

$$L_C(C) = \int_C \mathcal{H}(\kappa(s))\, ds, \tag{5.14}$$

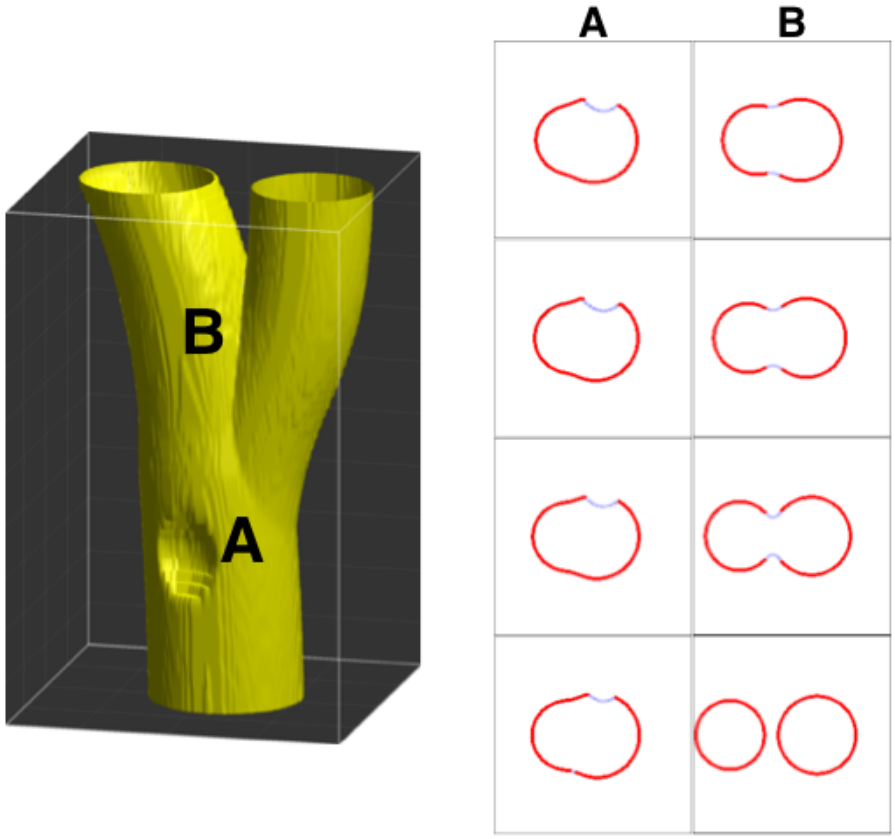


Figure 5.4: The behaviour of curvature in a simulated carotid artery bifurcation. Region A represents an undesirable segmentation boundary artifact. Region B is the bifurcation region. Columns A and B depict the sign of curvature at several cross-sections of these regions.

where $\mathcal{H}$ is a standardly defined *Heaviside function*,

$$\mathcal{H}(z) \quad = \quad \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases}, \tag{5.15}$$

whose role is to discard sections of the curve with negative curvature by setting them to zero, while counting the sections with positive curvature. The weak convexity prior energy is defined as

$$E_C(C) = -\log\left(\frac{L_C(C)}{L(C)}\right). \tag{5.16}$$

The ratio inside the logarithm is a number between 0 and 1 denoting the fraction of the curve which is convex. When the curve C is entirely convex, $E_C(C)$ equals zero. When,

on the other hand, the concave portion of the curve starts dominating its convex portion, the energy $E_C(C)$ tends to infinity. Thus, minimizing $E(C)$ will encourage the curve to be as convex as possible.

When used in combination with other energies discussed in Section 3.2, this energy will encourage the overall convexity without overly constraining the curve evolution. Furthermore, the rationale for using the logarithm can be understood by examining the possible values of $E(C)$ graphically, as shown in Figure 5.5. Segmentation boundaries with a large percentage of non-convex regions ($L_C\ /L \ll 1$) are penalized more severely than those with a smaller percentage ($L_C\ /L \approx 1$). Non-convex artifacts, such as region A in Figure 5.4, can come in a variety of sizes. Because of the use of logarithm, the energy $E_C(C)$ as defined in (5.16) will be more sensitive to those artifacts that have a large surface area than those with low surface area. Small artifacts will not strongly affect the ratio of L_C to L and will not be penalized. Such artifacts, however, will result in a more irregular surface, which will be locally penalized by the curve-shortening energy such as (3.4). As a result, the energy E_C penalizes large artifacts, detected by their deviation from convexity, while being robust enough to allow for the concavity around the bifurcation region to remain.

5.3 Proposed segmentation algorithm

The algorithm for segmenting the three-dimensional carotid ultrasound forms the main contribution of this work. As a pre-processing step, however, the image is first subjected to decomposition. Other features are extracted from the components and from the original image using an algorithm explained in Chapter 7. In addition, some of the features are used for the detection of edges, with examples and comparisons provided in the same chapter.

The segmentation algorithm is designed to robustly incorporate the prior knowledge about the anatomy of the artery with the information coming from the observed image. The proposed algorithm combines these sources of information in an energy minimization framework. The segmentation boundary is evolved in an active contours approach, which restricts the segmentation boundary to geometrically plausible shapes. The statistical dissimilarity between the artery and the background is exploited using a non-parametric distribution tracking approach performed on a variety of features derived from the observed intensity image. The edge information in the image is incorporated to aid the convergence of the algorithm to tissue interfaces by means of a geodesic active contours energy. Finally, the prior information about the shape of the carotid artery is incorporated in the form of a "weak" prior. Such an approach imposes an a-priori constraint on the shape of the segmented object - encouraging it to be convex - without overly restricting this shape in an inflexible fashion.

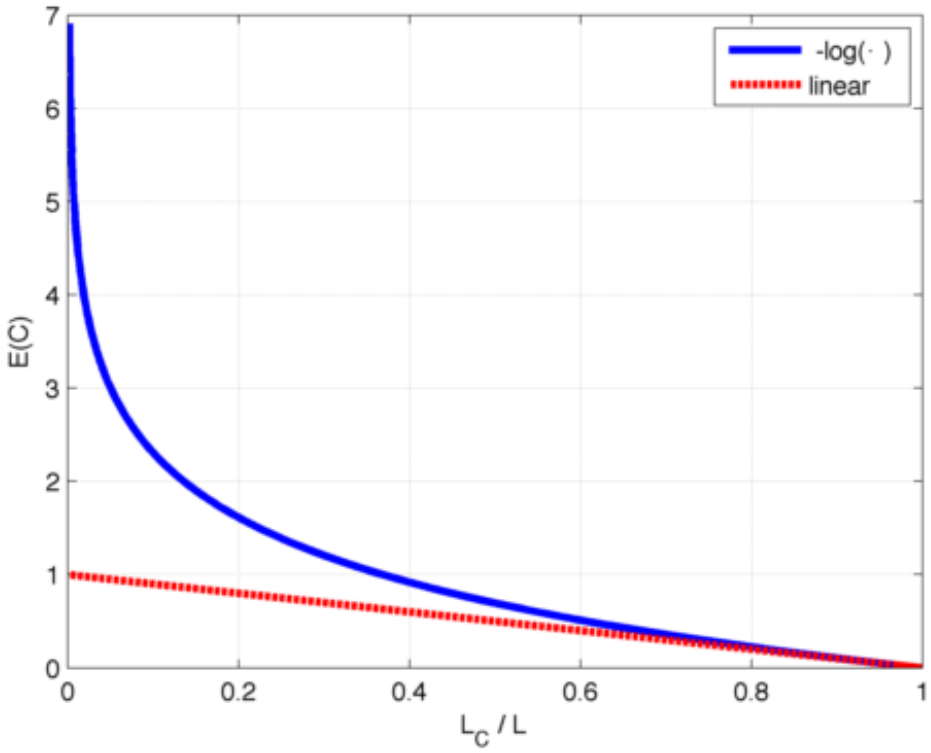


Figure 5.5: The behaviour of the energy $E_C(C)$ in (5.16) (solid line), compared to a linear function without the use of $\log(\cdot)$.

Chapter 6

Implementation of the proposed algorithm

In this chapter, the implementation of the proposed segmentation algorithm is detailed. Aside from the theoretical background of the proposed approach, computational details are presented which are instrumental in making the algorithm computationally tractable.

6.1 Level set method

The proposed segmentation algorithm attempts to solve an energy minimization problem, as presented in Section 3.2.2. In particular, the segmentation boundary is evolved in the *active contours* framework using the *level set method*[50]. Its unique advantages, in particular the elimination of the need to parameterize the segmentation curve, as well as its graceful handling of topological changes, have resulted in its widespread popularity in image segmentation tasks[20, 3], including medical imaging. In particular, the level set method has been widely applied to ultrasound imagery (see [39] for an extensive overview).

The curve representing the segmentation boundary is defined as follows. Let $\phi(\mathbf{x}) : \Omega \mapsto \mathbb{R}$ be a *level-set function* by means of which one can separate the domain Ω into inside and outside sub-domains, over which the level-set function is respectively positive and negative:

$$\Omega_A = \{\mathbf{x} \mid \phi(\mathbf{x}) > 0, \mathbf{x} \in \Omega\} \tag{6.1}$$

$$\Omega_B = \{\mathbf{x} \mid \phi(\mathbf{x}) \leq 0, \mathbf{x} \in \Omega\}. \tag{6.2}$$

In this case, the *zero level set* $\Gamma=\{\mathbf{x} \mid \phi(\mathbf{x}) = 0, \ \mathbf{x} \in \Omega\}$ is known as an active contour, which serves the purpose of a decision (segmentation) boundary. As Figure 6.1 illustrates,

Γ is an implicit representation of the boundary: the segmentation algorithm operates on the level set function ϕ, and its zero level set Γ implicitly evolves with it. One particularly notable advantage of the implicit definition of the segmentation contour is also illustrated in Figure 6.1: the curve Γ underwent a topological change by splitting in two. As previously discussed in Section 3.3, alternative contour evolution schemes, such as snakes, require additional logic to handle such a transition, whereas in the level set method, the transition is implicit and seamless.

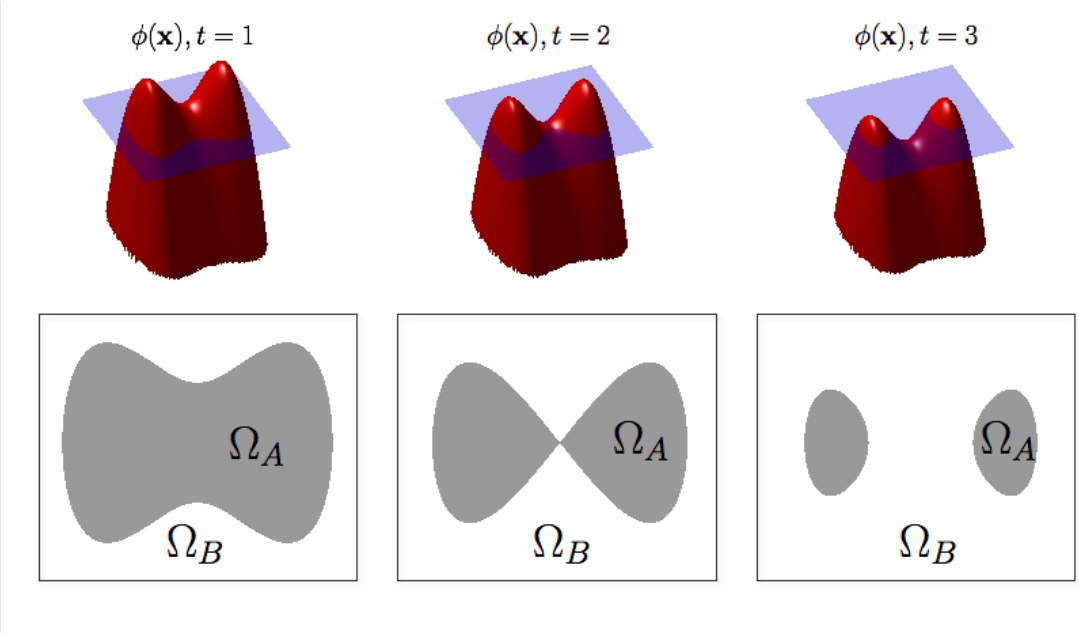


Figure 6.1: Implicit curve representation in the level set method[71]. Upper row: A level set function $\phi(\mathbf{x})$ (in red) evolving in time and its zero cross-section (in blue). Lower row: the same cross-section, illustrating the definitions in (6.1) and (6.2).

The dynamic behaviour of the level set function ϕ is governed by the following gradient flow:

$$\frac{\partial \phi}{\partial t} = -\frac{\delta E}{\delta \phi} |\nabla \phi| \; = \; V \, |\nabla \phi|, \tag{6.3}$$

where $\frac{\delta E}{\delta \phi}$ is the first variation of $E(\phi)$, a combined energy functional defined in Section 3.2, and V is the "velocity" which controls the evolution of ϕ with respect to the artificial time parameter t.

Although the level set function can in general be arbitrarily defined, as long as it is smoothly varying[50], in practice it is usually defined to be a *signed distance function*. Specifically, the function ϕ is defined in terms of a pre-existing contour C dividing the

image domain Ω into Ω_A and Ω_B as

$$\phi(\mathbf{x}) = \begin{cases} d(\mathbf{x}, C), & \mathbf{x} \in \Omega_A \\ -d(\mathbf{x}, C), & \mathbf{x} \in \Omega_B, \end{cases} \tag{6.4}$$

where $d(\mathbf{x}, C)$ represents the minimum distance between $\mathbf{x}$ and the contour C. The chief advantage of the signed distance function is that they are solutions to the eikonal equation,

$$\begin{aligned} |\nabla\phi| &= 1, \\ \phi(\mathbf{y}) &= 0,\ \mathbf{y} \in \partial\Omega \end{aligned} \tag{6.5}$$

which simplifies the implementation of the curve evolution by gradient flow in (6.3) since $|\nabla\phi|$ is universally unity.

6.1.1 Definition of segmentation energies

Using the level set framework, the energy terms which govern the dynamic behaviour of the level set function ϕ, and with it the sought segmentation boundary Γ, need to be specified in terms of ϕ, together with an expression for their first variation. First, the sub-domains Ω_A and Ω_B, as well as the boundary $\partial\Omega$, need to be defined in terms of ϕ. This can be done using the Heaviside function defined in (5.15) and its derivative $\delta(\cdot)$, the Dirac delta function. In this case, $\mathcal{H}(\phi(\mathbf{x}))$ becomes an indicator function of the region Ω_A, while $\mathcal{H}(-\phi(\mathbf{x}))$ becomes an indicator function for Ω_B.

In practice, however, the indicator functions are rarely used, because they are discontinuous and not differentiable. Instead, they are commonly regularized by suitable differentiable approximations. One such regularization[10] redefines the Heaviside indicator function as

$$\mathcal{H}_\epsilon(z) = \begin{cases} 1, & z > \epsilon \\ 0, & z < -\epsilon, \\ \frac{1}{2}\left[1 + \frac{z}{\epsilon} + \frac{1}{\pi}\sin\left(\frac{\pi z}{\epsilon}\right)\right], & |z| \le \epsilon. \end{cases} \tag{6.6}$$

The delta function is redefined as

$$\delta_\epsilon(z) = \begin{cases} 0, & |z| > \epsilon \\ \frac{1}{2\epsilon}\left(1 + \cos(\frac{\pi z}{\epsilon})\right), & |z| \le \epsilon. \end{cases} \tag{6.7}$$

Note that both $\mathcal{H}_\epsilon$ and δ_ϵ converge to the original Heaviside and delta functions as $\epsilon \to 0$.

The *geodesic length* energy is defined in terms of ϕ as

$$E_G(\phi) = \int_\Omega \delta_\epsilon(\phi)\ g(I(\mathbf{x}))d\mathbf{x}, \tag{6.8}$$

with the geodesic function defined for example as in (3.5). Note that including the $\delta_\epsilon(\phi)$ term only integrates the function $g(\cdot)$ in those locations that are close to the zero level set of ϕ. The first variation of E_G can be shown to be[18]

$$\frac{\delta E_G}{\delta \phi} = -\delta_\epsilon(\phi)\,\mathrm{div}\left(g(I(\mathbf{x}))\cdot\frac{\nabla\phi}{|\nabla\phi|}\right). \tag{6.9}$$

As discussed in Section 3.2.3, the energies comparing the empirical probability densities of Ω_A and Ω_B can either assume a parametric model, such as a restriction to a Gaussian distribution, or be parameter-free. The Chan-Vese minimum variance energy, which is used when the Gaussian assumption is made, is defined in the level set framework as,

$$E_{CV}(\phi) = \int_\Omega |I(\mathbf{x}) - \mu_A|^2\,\mathcal{H}(\phi(\mathbf{x}))\,d\mathbf{x} + \int_\Omega |I(\mathbf{x}) - \mu_B|^2\,\mathcal{H}(-\phi(\mathbf{x}))\,d\mathbf{x}, \tag{6.10}$$

where μ_A and μ_B are the averages of $I(\mathbf{x})$ in Ω_A and Ω_B, respectively. Recall that $I(\mathbf{x})$ is not necessarily the original image, but rather a vector-valued image of its features. The above definition is consistent with such vector-valued images. The first variation of (6.10) is given by[10] as

$$\frac{\delta E_{CV}}{\delta \phi} = |I(\mathbf{x}) - \mu_A|^2 + |I(\mathbf{x}) - \mu_B|^2. \tag{6.11}$$

On the other hand, the *Bhattacharyya* energy enables non-parametric analysis of regional statistics. It uses empirical estimates of the pdfs of the samples in Ω_A and Ω_B, as shown in (3.9), which are redefined in terms of ϕ as

$$\begin{aligned} P_A(\mathbf{z}|\phi) &= \frac{\int_\Omega K(\mathbf{z} - I(\mathbf{x}))\mathcal{H}(\phi)d\mathbf{x}}{\int_\Omega \mathcal{H}(\phi)d\mathbf{x}}, \\ P_B(\mathbf{z}|\phi) &= \frac{\int_\Omega K(\mathbf{z} - I(\mathbf{x}))\mathcal{H}(-\phi)d\mathbf{x}}{\int_\Omega \mathcal{H}(-\phi)d\mathbf{x}}, \end{aligned} \tag{6.12}$$

where the smoothing *kernel* function $K(\cdot)$ can be defined, for example, as a Gaussian kernel in (4.9). The Bhattacharyya energy is given by

$$E_{bh}(\phi(\mathbf{x})) = \int_{\mathbf{z}\in\mathbb{R}^d} \sqrt{P_A(\mathbf{z}|\phi)\;P_B(\mathbf{z}|\phi)}\;d\mathbf{z}. \tag{6.13}$$

The first variation, whose full derivation is detailed in[35], is as follows:

$$\begin{aligned} \frac{\delta E_{bh}}{\delta \phi} &= \delta_\epsilon(\phi(\mathbf{x}))\cdot V(\mathbf{x}), \\ V(\mathbf{x}) &= \frac{1}{2}E_{bh}(\phi(\mathbf{x}))(A_A^{-1} + A_B^{-1}) + \frac{1}{2}\int_{\mathbf{z}\in\mathbb{R}^d} K(\mathbf{z} - I(\mathbf{x}))L(\mathbf{z}\mid\phi(\mathbf{x}))\,d\mathbf{z}, \\ L(\mathbf{z}\mid\phi(\mathbf{x})) &= \frac{1}{A_A}\sqrt{\frac{P_B(\mathbf{z}|\phi)}{P_A(\mathbf{z}|\phi)}} - \frac{1}{A_B}\sqrt{\frac{P_A(\mathbf{z}|\phi)}{P_B(\mathbf{z}|\phi)}}. \end{aligned} \tag{6.14}$$

In Chapter 7, both of these energies are tested to determine which works best in the 3D carotid ultrasound setting.

The weak geometric prior energy proposed in this work, whose background and rationale are explained in Section 5.2, requires the calculation of curvature from the level set function ϕ. In three dimensions, the mean curvature is given by the following expression:

$$\kappa(\phi) = \operatorname{div}\left(\frac{\nabla\phi}{\|\nabla\phi\|}\right) = \frac{\phi_{\mathrm{x}}^2(\phi_{\mathrm{yy}} + \phi_{\mathrm{zz}}) + \phi_{\mathrm{y}}^2(\phi_{\mathrm{xx}} + \phi_{\mathrm{zz}}) + \phi_{\mathrm{z}}^2(\phi_{\mathrm{xx}} + \phi_{\mathrm{yy}})}{\|\nabla\phi\|^3}. \tag{6.15}$$

As shown before, using this definition, one can differentiate between convex and concave portions of the level set surface. A functional $L(C)$ in (5.13), measuring the length (or surface area) of the contour, is re-defined as

$$L(\phi) = \int_\Omega \delta(\phi)\, \|\nabla\phi\|\, d\mathbf{x}, \tag{6.16}$$

and the convex length functional $L_C(C)$ in (5.14), only counting the convex portions of the overall length, is given in terms of ϕ by

$$L_C(\phi) = \int_\Omega H(\kappa(\phi))\; \delta(\phi)\, \|\nabla\phi\|\, d\mathbf{x}. \tag{6.17}$$

The convexity prior energy is defined in terms of (6.16) and (6.17) as

$$E_{conv}(\phi) = -\log\left(\frac{L_C(\phi)}{L(\phi)}\right) = \log(L(\phi)) - \log(L_C(\phi)). \tag{6.18}$$

The first step in finding the first variation of $E_{conv}(\phi)$ is to recognize that, by the rules of variational calculus and analogously with finding the derivative of a logarithm,

$$\frac{\delta E_{conv}}{\delta\phi} = \frac{\delta L(\phi)}{\delta\phi}\frac{1}{L(\phi)} - \frac{\delta L_C(\phi)}{\delta\phi}\frac{1}{L_C(\phi)}. \tag{6.19}$$

Since the length functional $L(\phi)$ is simply the geodesic length functional in (6.8) with a constant $g = 1$, its first variation is already given in (6.9). The convex length functional $L_C(\phi)$ can also be viewed as a realization of the geodesic length functional in (6.8), with $g = g_{conv}(\mathbf{x}) = \mathcal{H}(\kappa(\phi(\mathbf{x})))$. However, the geodesic function g is normally defined as "external" functions of the observed image ($g(I)$), whereas g_{conv} is an "internal" function of the level set function ϕ. Substituting g_{conv} into the first variation (6.9) ignores the dependence on ϕ, and hence ignores the dependence on the segmentation curve itself. However, this simplification is proven to perform as intended in practice, penalizing the convex portions of the segmentation boundary and minimizing their area. This simplified convexity energy is considerably simpler to express and compute than the full expression for $\frac{\delta L_C(\phi)}{\delta\phi}$.

6.2 Computational details

Many components of the segmentation algorithm explained above from a theoretical point of view can be implemented in practice in multiple ways. Straightforward implementations often suffer from drawbacks such as computational inefficiency or excessive memory requirements. Efficient algorithms, on the other hand, may approximate the theoretical results instead of computing them exactly. In a complex algorithm with many interacting "moving parts", this could lead to quite divergent segmentation results.

The strategy followed in this work is to first evaluate the behaviour of the components (i.e., different features and energies) and to draw conclusions as to their performance in the overall algorithm. At the same time, each of the concepts presented in preceding chapters has a rich array of efficient numerical implementations, and these optimizations should be examined in order to back the claim that the proposed approach is computationally tractable. As a result, alternative implementations are presented and discussed for most components. Chief among these is an accurate and efficient numerical scheme for iteratively solving PDEs semi-implicitly, a result which is applicable to both the dynamic evolution of the level set function and to the computation of the image features obtained through the use of non-linear diffusion filtering.

6.2.1 The level set method

There are two main topics in the discussion on numerical implementation of level set segmentation. Firstly, there are several different representations of the level set function ϕ. Secondly, there are two main choices of the numerical scheme for the evolution of ϕ in the energy minimization framework.

The segmentation algorithm is initialized with a contour representing an initial estimate of the true artery boundary. An initial level set function ϕ_0 must be found by computing the signed Euclidean distance transform from this contour. To evaluate this transform, one must find the shortest distance from any pixel location in the image to the given contour, which will become the zero level set when the transform is computed. One efficient algorithm for such a task is the *Fast Marching Method*, which iteratively builds the solution from a set of initial pixels, visiting each pixel only once. The algorithm is very similar to the well-known Dijkstra's shortest-path algorithm: it always visits the location with the shortest distance and updates the distances of all its unvisited neighbours[50, 41]. Another efficient way to compute the distance transform in arbitrary dimensions based on the Voronoi diagrams[34].

The most straightforward way to discretize the contour evolution equation in (6.3) is

known as the *explicit* scheme, which leads to the following update equation:

$$\phi^{k+1} = \phi^k + \tau^k \, V(\phi^k) \cdot \phi^k, \tag{6.20}$$

where τ^k represents a discrete time step following a discretization of t, k stands for the iteration number and having rewritten the right-hand side of (6.3) in matrix-vector notation, where the velocity matrix V updates the level set function sampled and collapsed into a vector ϕ^k. The discretization of the time parameter t requires placing an upper limit on the time step τ^k at each iteration k in order to limit the evolution of the contour to only the adjacent grid point. This condition is necessary to preserve stability and is known as the Courant-Friedrichs-Lewy (CFL) condition[15]:

$$\tau^k \leq \frac{c}{\|V(\phi^k)\|_\infty}, \tag{6.21}$$

where c is some constant, $0 < c \leq 1$.

As the level set function is modified at each iteration as shown in (6.20), it will deviate progressively further from being a valid signed distance function. Eventually, these accumulating errors will cause severe artifacts and the segmentation algorithm will become numerically unstable[41]. In order to maintain numerical accuracy, a re-initialization procedure must be periodically applied to the level set function ϕ. Such a procedure recomputes the distance transform in an efficient manner, while keeping the zero level set unmodified in order to preserve the segmentation boundary thus embedded into ϕ. Re-initialization algorithms aim to avoid re-computing the entire signed distance function from scratch. One such algorithm achieves this by attempting to restore the Eikonal property of the function ($\|\nabla\phi\| = 1$) by performing iterations of the following problem until steady state[55]:

$$\phi^{t+1} = \phi^t + \Delta t \; \text{sign}(\bar{\phi})(1 - |\nabla\phi^t|), \tag{6.22}$$

where $\bar{\phi}$ is the initial un-fixed level set function. In order to converge quickly to the correct solution, $\bar{\phi}$ cannot deviate too far from a valid signed distance function. Thus, the procedure must be performed every couple of iterations in order to work[41]. Another method is to run the Fast marching method for only a small number of pixels within the vicinity of the zero level set. Yet another approach attempts to avoid any ad-hoc re-initialization procedures by incorporating an energy term penalizing deviation from the Eikonal property, i.e. the deviation from $\|\nabla\phi\| = 1$[28]. The safest way to avoid approximation artifacts is to re-compute the level set function using an exact algorithm, such as [34], at each iteration, but it is also the slowest. It has been used only to ensure that using re-initialization algorithms does not significantly alter the final segmentation result.

Although the level set function is defined over the entire image domain, the only useful information it carries is contained in its zero level set Γ. The CFL condition ensures

that there is an upper limit to the rate with which ϕ is updated, and thus the contour cannot abruptly shift positions in adjacent iterations. This implies that the values of ϕ are only important in the vicinity of Γ. A popular approach at improving the computational efficiency of the level set method is to constrain all calculations to a band of pixel locations close to Γ, known as the *narrow band*[50]:

$$\mathcal{N}_b^k = \{\mathbf{x} \mid |\phi^k(\mathbf{x})| \leq b\}, \tag{6.23}$$

where b is the parameter defining the width of the band. The location of $\mathcal{N}_b$ with respect to the rest of Ω is illustrated in Figure 6.2. Only those pixels whose minimum distance to the zero level set does not exceed b are included in the narrow band; for all other pixel locations, no calculations are performed. The level set function ϕ changes with each iteration, and the narrow band changes with it.

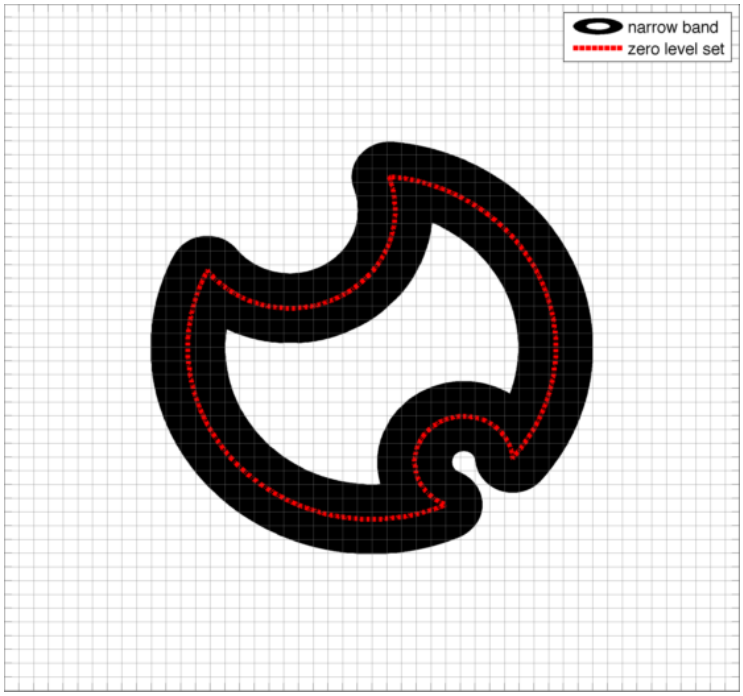


Figure 6.2: Illustration of the narrow band approach. The level set function is updated only at pixels closest to the zero level set.

Although it is considerably better than evolving the full signed distance function, the narrow band approach still carries with it a considerable computational cost, especially in 3D. In particular, the need for re-initialization of ϕ is both computationally expensive and rooted in implementation specifics rather than the energy minimization or level set theory. Several methods[51, 69, 32] attempt to improve the efficiency of contour evolution

by storing the immediate neighbourhood of the zero level set in a series of linked lists. In [69], five lists of pixel locations are maintained: the zero level set and two sets of its immediate neighbours on each side, and ϕ is evolved using only these lists[27]:

$$L_0 \leftarrow \{\mathbf{x} \mid \phi(\mathbf{x}) \in [-0.5, \quad 0.5]\} \tag{6.24}$$
$$L_1 \leftarrow \{\mathbf{x} \mid \phi(\mathbf{x}) \in (0.5, \quad 1.5]\} \tag{6.25}$$
$$L_2 \leftarrow \{\mathbf{x} \mid \phi(\mathbf{x}) \in (1.5, \quad 2.5]\} \tag{6.26}$$
$$L_{-1} \leftarrow \{\mathbf{x} \mid \phi(\mathbf{x}) \in [-1.5, \quad -0.5)\} \tag{6.27}$$
$$L_{-2} \leftarrow \{\mathbf{x} \mid \phi(\mathbf{x}) \in [-2.5, \quad -1.5)\} \tag{6.28}$$
$$\tag{6.29}$$

Contour evolution consists of evaluating the velocity force V only at the locations in the zero list, and moving the pixels whose value falls outside the range $[-0.5, 0.5]$ to one of the neighbour lists. The values of ϕ in the locations belonging to these lists are kept approximately accurate by computing a fast distance approximation at each step, since they are needed to compute the finite differencing approximations to its derivatives at the zero level set. The method maintains the accuracy of curve evolution while achieving significant reduction in computational load. Further simplifications include reducing the number of lists to three[32] (illustrated in Figure 6.3) at the cost of small inaccuracies, as well as dispensing with tracking the actual values of ϕ, only keeping track of its sign[69], leading to even faster approximate solutions. While all these methods present a viable alternative to the narrow band approach, the results presented in Chapter 7 sacrifice computational efficiency to ensure maximum accuracy, with the knowledge that improvements in speed are readily available once the choice of energies, features and priors is finalized.

Another approach at improving the computational load of level set segmentation utilizes a *semi-implicit* discretization of the gradient flow in (6.3):

$$\phi^{k+1} = \phi^k + \tau \, V(\phi^k) \cdot \phi^{k+1}, \tag{6.30}$$

where τ does not vary from iteration to iteration, in contrast with (6.20). More importantly, the solution of (6.30) now involves a solution of a linear system of equations,

$$\big(\mathrm{Id} - \tau \, V(\phi^k)\big) \cdot \phi^{k+1} = \phi^k, \tag{6.31}$$

which, as is, would require considerable effort to solve for ϕ^{k+1}. However, using the additive operator splitting (AOS) approximation technique[67, 65, 18], this problem can be transformed into that of solving tri-diagonal systems of equations,

$$\frac{1}{n} \sum_{i=1}^{n} \big(\mathrm{Id} - n\tau \, V_i(\phi^k)\big) \cdot \phi^{k+1} = \phi^k, \tag{6.32}$$

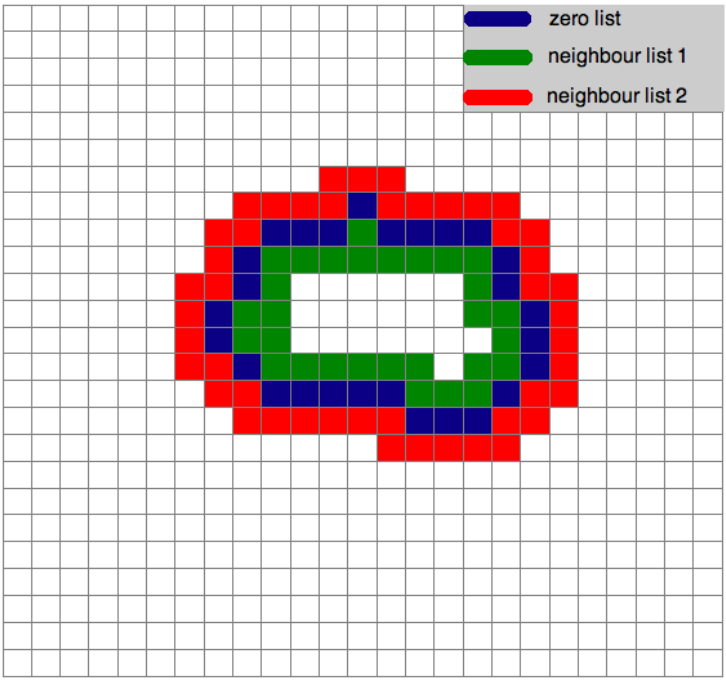


Figure 6.3: Illustration of the sparse field method using three lists. The level set function is updated only at the zero list. Contour evolution is performed by interchanging pixel locations between the three lists when the value of ϕ at those locations changes appropriately.

where n is the number of dimensions in the original image and Id is the identity matrix. To find ϕ^{k+1}, a linear system of equations must be solved. Fortunately, however, the system in (6.32) is tri-diagonal, which is efficiently computable using the Thomas algorithm. While each iteration takes longer to compute than a single iteration of the explicit scheme, this unconditionally stable technique allows a considerable speedup because much larger time steps are possible[67]. While this scheme was originally developed for diffusion filtering of images, the problem of level set contour evolution can also be considered as quasi-diffusion governed by a combination of internal and external energies, and the segmentation can be performed using the narrow band technique,[18] although, since the CFL condition no longer constrains the speed with which ϕ evolves, extra care must be taken to ensure that the narrow band is wide enough for the zero level set to always remain inside it.

6.2.2 Initialization

The accuracy of the segmentation results in the energy minimization framework is heavily initialization-dependent, especially when applied to ultrasound images. Although an initialization that is very close to the boundaries of the artery is preferred, it would involve an amount of manual work that defeats the purpose of designing a level set segmentation algorithm. However, if a small amount of manual input can significantly speed up the algorithm, as well as make the result more accurate, the tradeoff is acceptable.

The method used in the present work aims to strike a balance between the amount of manual input and the quality of segmentation. Segmentation boundary is initialized entirely inside the artery, and then expands outwards. The configuration at which the contour stabilizes depends on the other energy terms described above. In order to drive the segmentation contour to expand, a special "balloon" force needs to be introduced:

$$\frac{\delta E_b}{\delta \phi} = g(I). \tag{6.33}$$

Since $g(I) \in [0, 1]$, with 1 corresponding to edge-free regions of the image, this balloon force will be reduced at the edges, while driving the expansion of the contour over continuous regions.

An illustration of the initialization procedure is given in Figure 6.4. The top row shows two-dimensional cross-sections of the artery at four sample points, in which two markers have been manually placed: one for the internal carotid artery (ICA), and one for the external carotid artery (ECA) (in the common carotid artery before the bifurcation, the markers can be placed inside the lumen at any location, even overlapping each other). The markers are used to create two 1D "skeleton" curves, one lying entirely inside the ICA, and the other inside the ECA. The initial contour ϕ_0, whose zero level surface is shown in the bottom row and has a cylindrical shape, can then be efficiently computed using a distance transform computation method such as the fast marching method. All points on the zero level surface of ϕ_0 are approximately equidistant from their "skeleton" curve.

6.3 Computation of image features

Computational details of the image features outlined in Chapter 4 are outlined next. Some feature types, such as local statistics, are straightforward to implement in a computationally feasible manner. However, for others, such as diffusion filtering, considerable speedup can be achieved by using efficient approximations. Implementation of the proposed ultrasound decomposition algorithm is also made feasible by such apporixmations. Additionally, computation of the curvature of the level set function ϕ needs to be carefully considered.

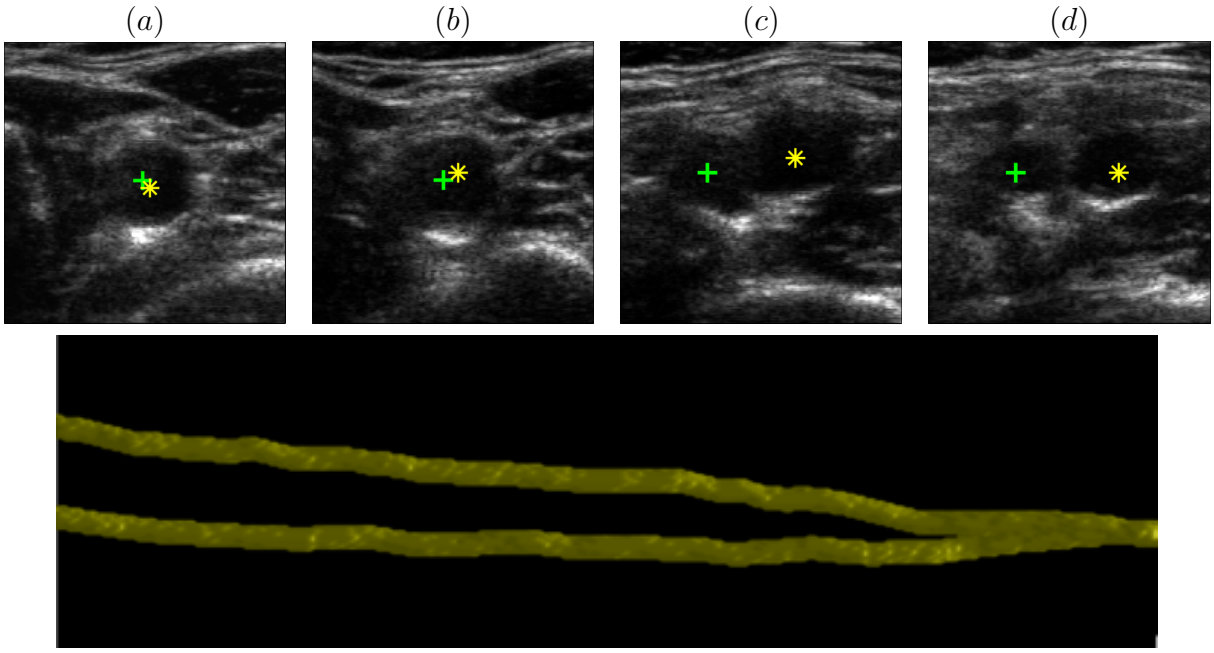


Figure 6.4: Segmentation contour initialization. Top row: The contour is manually initialized by placing markers inside the arteries. Bottom row: After cubic interpolation, the contour is initialized as a combination of two cylindrical shapes.

6.3.1 Non-linear diffusion filtering

The diffusion features of the type presented in Section 4.2.1 follow the same model as the level set function evolution. Note the equivalence of the diffusion equation in (4.12) and the first variation of the geodesic length energy in the level set formulation in (6.9). Therefore, the explicit and semi-implicit schemes are valid numerical discretizations for both the evolution of ϕ and the diffusion of the image $I(\mathbf{x})$. From the filters discussed in Section 4.2.1, the semi-implicit scheme is used as a computationally efficient approximation to the Total Variation and Speckle-Reducing Anisotropic Diffusion features, while of course being generally applicable to any spatially varying isotropic diffusion.

6.3.2 Implementation of the proposed ultrasound image decomposition algorithm

The proposed image decomposition scheme for ultrasound images, presented in Section 5.1, aims to improve the extraction of image features by first decomposing the ultrasound image into a speckle texture component and a piecewise-smooth speckle-free component.

This is accomplished using the framework of maximum-a-posteriori statistical estimation, resulting in an energy functional (5.10). The solution to this optimization problem of two variables is obtained by alternating between optimizing convex sub-functionals,

$$E_v(v_l \mid u_l) = \|v_l + (g_l - u_l)\|_2^2 + \lambda_2 \sum_{\mathbf{x}\in\Omega} (\exp 2v_l(\mathbf{x}) - 2v_l(\mathbf{x})), \tag{6.34}$$

$$E_u(u_l \mid v_l) = \|u_l + (g_l - v_l)\|_2^2 + \lambda_1 \|u_l\|_{TV}, \tag{6.35}$$

instead of devising procedures for minimizing (5.10) directly. The solution is found by alternating steps of minimizing (6.34) and (6.35) while holding constant the intermediate estimates u_l and v_l, respectively. Figure 6.5, a pictorial overview of the proposed algorithm, shows these two stages in succession. The procedure terminates when the relative change in both components becomes negligible.

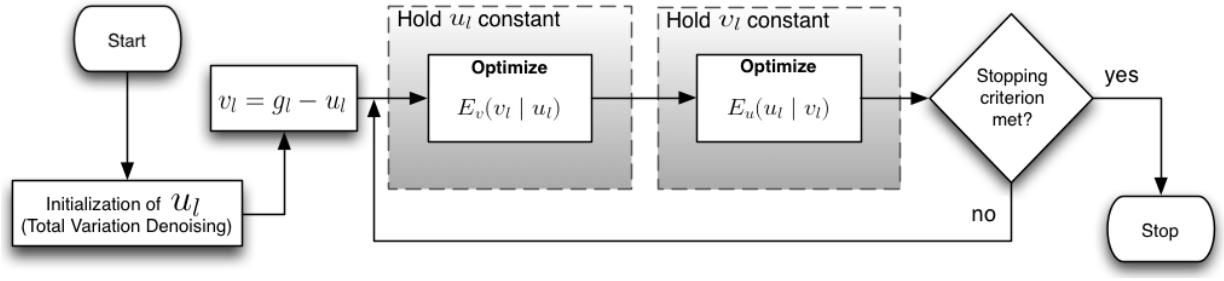


Figure 6.5: Overview of the proposed ultrasound image decomposition algorithm.

Before elaborating on the algorithms for solving these subproblems, initialization must be discussed. The convergence of this iterative algorithm can be greatly sped up by using a good initial estimate for u_l and v_l. Consequently, the initial value of u_l is defined as the minimizer of the previously discussed total variation (TV) denoising model[49],

$$u_l^{(0)} = \arg\min_{u_l} \|u_l\|_{TV} + \lambda \|g_l - u_l\|_2^2, \tag{6.36}$$

where λ is a scale parameter. The TV denoising algorithm has a number of efficient implementations (see for example [8], which was used in this work). The initial value of v_l is simply given by the difference between this result and the observed image ($v_l^{(0)} = g_l - u_l^{(0)}$). These two steps are indicated in the diagram in Figure 6.5.

Optimization of $E_u(u_l \mid v_l)$

The functional $E_u(u_l \mid v_l)$ in (6.35) describes another instance of the TV denoising problem, which can be solved in the variational framework with the use of the first variation of E_u

given by

$$\frac{\delta E_u(u_l \mid v_l)}{\delta u_l} = ((u+v)-g) - \lambda_1 \operatorname{div}\left(\frac{\nabla u}{\|\nabla u\|}\right). \tag{6.37}$$

The gradient descent algorithm can now be used to iteratively find the solution via

$$\frac{\partial u_l}{\partial t} = -\frac{\delta E_u(u_l \mid v_l)}{\delta u_l} = \lambda_1 \operatorname{div}\left(\frac{\nabla u_l}{\|\nabla u_l\|}\right) + (g_l - (u_l + v_l)), \tag{6.38}$$

where t is the artificial parameter of time, with $t = 0$ as its initial value. This problem is discretized in the same manner as other diffusion-based filtering problems, via either the explicit or the semi-implicit scheme.

Optimization of $E_v(v_l \mid u_l)$

To examine the method of minimizing (6.34), it is important to note that this functional is convex and two times differentiable in v_l, whether the latter is defined over a continuous or discrete spatial domain. Consequently, the first derivative of this function with respect to v_l is given by

$$\nabla_v E = (u_l + v_l - g_l) - \lambda_2 \left[2\exp(2v_l) - 2\right], \tag{6.39}$$

and the Hessian matrix of second-order derivatives is

$$\nabla_v^2 E = I + 4\lambda_2 \operatorname{diag}(\exp(2v_l)). \tag{6.40}$$

The Hessian in (6.40) is diagonal, and thus its inversion amounts to point-wise division of the gradient by the diagonal of the Hessian. Because of this fact, $E_v(v_l \mid u_l)$ is minimized by the Newton method. At the $(k+1)^{\text{th}\,s}$ iteration, v_l is updated according to

$$v_l^{(k+1)} = v_l^{(k)} + \beta p^{(k)}, \qquad p^{(k)} = -(\nabla_v^2 E)^{-1} \cdot \nabla_v E, \tag{6.41}$$

where β is a step size found by line search along the Newton direction $p^{(k)}$. The sequence of updates in (6.41) is continued until convergence.

6.3.3 Outlier suppression

As discussed in Section 2.2.2, an important post-processing step in ultrasound imaging is the compression of the dynamic range in order to make the observed image more suitable for both displaying it on the monitor and quantitatively processing it. This procedure can

be applied both in hardware on the analog signal and in software on its discretized and stored version.

Although the logarithm function compresses the dynamic range of the signal, in the vicinity of zero its behaviour asymptotically approaches infinity, as can be seen in Figure 5.5. As a result, the log-compressed image is contaminated with outliers in locations where the original signal's value was close to zero. The log-speckle which contaminates the ultrasound image is a Fisher-Tippett random variable (see Section 2.3.1), whose example realization, shown in Figure 2.7, illustrates these outliers. They increase the dynamic range of the signal, and thus negatively influence the discernibility of the imaged tissue for both quantitative and qualitative analysis.

A useful additional image feature can be obtained if these outliers are removed. Their suppression is performed using the technique of robust outlier shrinkage[36, 4], which is illustrated in Figure 6.6. The outliers can be isolated by subtracting from the original image (Figure 6.6-a) its median-filtered version (Figure 6.6-b). This residual (Figure 6.6-c) contains the spiky noise, from which the outliers are isolated by soft thresholding (Figure 6.6-d). The threshold may be data-driven (e.g., a function of an outlier-resistant estimate of the noise variance), such as the median absolute deviation (MAD) defined in (4.5). The resultant spikes are subtracted from the original image, which now contains the original signal together with the noise, save for the outliers (Figure 6.6-e). Crucially, the outlier shrinkage procedure alters the distribution of the noise, making it more akin to Gaussian by rejecting the left tail of the Fisher-Tippett distribution.

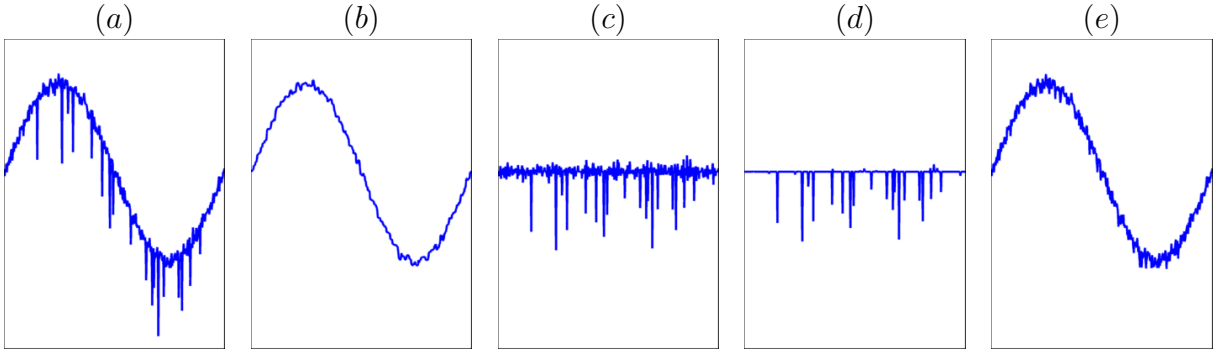


Figure 6.6: Illustration of outlier shrinkage. (a) The original signal contaminated with spiky noise; (b) Its median-filtered version; (c) The residual; (d) The residual after soft thresholding, indicating outliers; (e) Final result with outliers removed by subtraction.

6.3.4 Discretization of curvature

The evolution of the level set function heavily relies on accurate computation of the curvature of the level set function ϕ, which requires the approximation of first- and second-order partial derivatives of ϕ, as can be seen in (6.15). This approximation is done by means of various finite differencing schemes, such as central, forward and backward differences, which, if applied in the x direction on an image I are respectively denoted by

$$\begin{aligned} \frac{\partial I}{\partial x}(\mathbf{x}) &\approx 0.5(I[x+1,y,z] - I[x-1,y,z]) = \Delta_c^x I \\ \frac{\partial I}{\partial x}(\mathbf{x}) &\approx I[x+1,y,z] - I[x,y,z] = \Delta_+^x I \\ \frac{\partial I}{\partial x}(\mathbf{x}) &\approx I[x,y,z] - I[x-1,y,z] = \Delta_-^x I. \end{aligned} \tag{6.42}$$

Second order derivatives are approximated by successive application of (6.42), e.g. $I_{xx} = \Delta_-^x(\Delta_+^x I)$. These approximations can be used in various combinations in computing a numerical approximation of curvature. For example, a straightforward scheme uses central differences for all first-order derivatives, and computes the second-order derivatives as $I_{xx} = (I_x^+)_x^-$. By examining (6.42) it becomes clear that the numerical support of the curvature approximation in this case will be restricted to the immediate neighbours of the pixels. Another option is to use a more involved scheme, such as the one presented in [49], and in two dimensions given by

$$\begin{aligned} \kappa_{ROF}(\phi) = &\quad \Delta_-^x \left(\frac{\Delta_+^x \phi}{\sqrt{(\Delta_+^x \phi)^2 + (\mathrm{m}(\Delta_+^y \phi, \Delta_-^y \phi))^2}} \right) + \Delta_-^y \left(\frac{\Delta_+^y \phi}{\sqrt{(\Delta_+^y \phi)^2 + (\mathrm{m}(\Delta_+^x \phi, \Delta_-^x \phi))^2}} \right), \\ \mathrm{m}(a,b) = &\quad \left(\frac{\operatorname{sign} a + \operatorname{sign} b}{2} \right) \cdot \min(|a|, |b|). \end{aligned} \tag{6.43}$$

6.3.5 Curvature calculation for the convexity prior

The curvature discretization schemes described above all result in quantization artifacts when applied to ϕ, especially around its zero level set. The first reason for this is the error introduced by the finite differences. Secondly, the level set function is initialized using a binary mask delineating the inside and outside regions. This introduces quantization effects of its own, rendering a smooth curve jagged and irregular. As a result, the computed curvature is noisy, which, despite its suitability for the level set gradient flow, makes it unsuitable for differentiating between convex and concave regions based on curvature sign, as Figure 6.3.5 shows. In column A, the sign of directly computed curvature frequently jitters around 0, in fact producing no clear differentiation between convexity and concavity.

In Columns B and C, the level set function ϕ has been pre-smoothed with a Gaussian filter of variance σ_r^2. As a result, the sign of curvature now clearly marks concave regions of the curve.

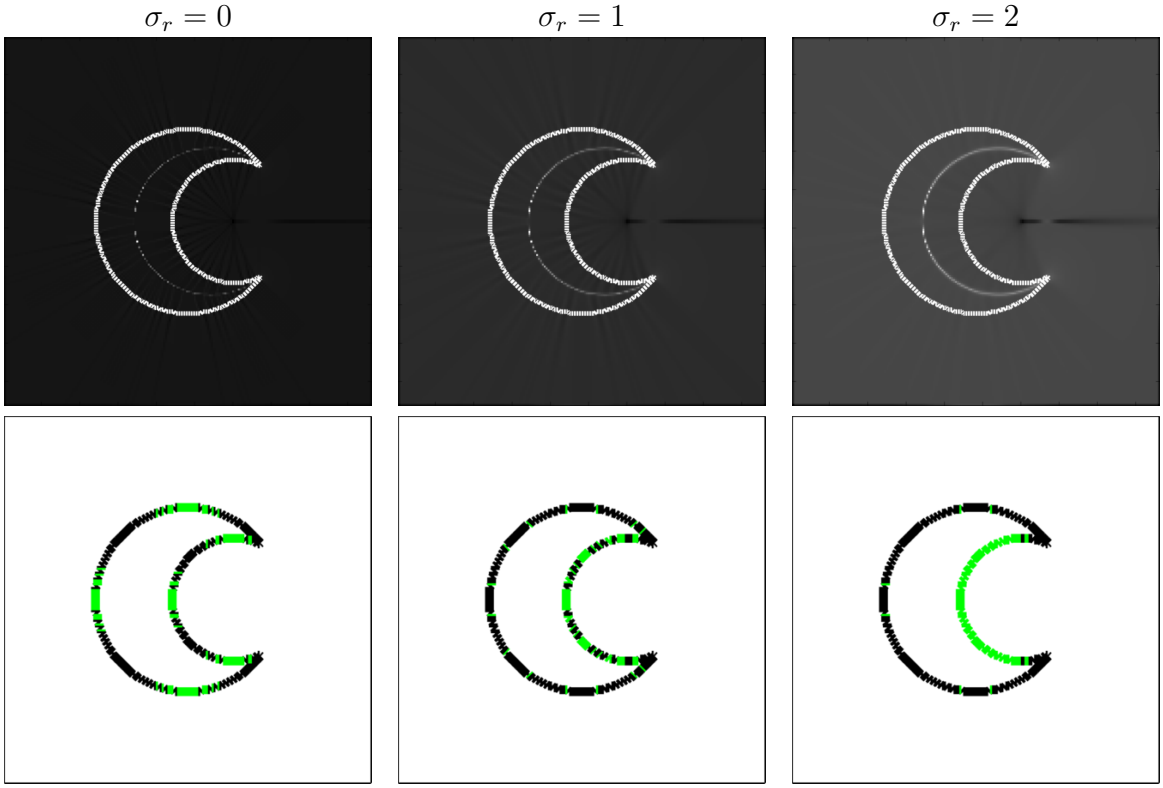


Figure 6.7: Regularization of curvature and quality of convexity detection. Top row: the curvature κ of the level set function, with the zero level set highlighted. Bottom row: sign of curvature along the zero level set; $\kappa > 0$ in dark, $\kappa < 0$ in bright.

Chapter 7

Results

The results presented in this chapter are grouped into four sections. First, the process of selecting image features for use in the regional segmentation framework is explained. Next, the proposed ultrasound decomposition scheme is evaluated visually and quantitatively. The detection of image edges is then discussed by illustrating the key tradeoffs involved in their calculation. Finally, the segmentation approach is tested in its entirety on real 3D ultrasound data sets, with a detailed illustration of the various options on one of these images, and typical results for three others.

The analysis presented below involves three-dimensional images and various image features. However, for demonstration purposes, all figures of these 3D objects used in this chapter are two-dimensional slices.

7.1 Image features

Examples of image features are presented next, with an accompanying algorithm for selecting from them a small subset of features most useful for the purpose of region-based segmentation. Examples of components obtained from the proposed decomposition algorithm are shown, and the statistical discriminability of the piecewise-smooth component is compared to an equivalent total variation denoising result.

7.1.1 Feature selection

The selection of image features that best discriminate between the artery and background was performed using the concept of Bhattacharyya distance between empirical densities, defined in (3.10). Since the features are ultimately used by regional segmentation energies,

choosing the Bhattacharyya metric on the basis of this distance seems justified. It is more convenient to rewrite the metric as a Bhattacharyya similarity,

$$B_{ij} = B(P_i(\mathbf{z}), P_j(\mathbf{z})) = \int_{\mathbf{z}\in\mathbb{R}^d} \sqrt{P_i(\mathbf{z})P_j(\mathbf{z})}\ d\mathbf{z}. \tag{7.1}$$

Figure 7.1 shows a representative sample of the image features discussed in Chapter 4, with the description of each feature in Table 7.1. This set of examples demonstrates a number of interesting properties mentioned previously:

- D1 shows the smearing of edges that results from applying the Gaussian filter
- A2 and B2 show the result of local entropy with and without logarithm, demonstrating the invariance under the logarithmic transformation
- C3 and D3 demonstrate the effect of performing edge-enhancing anisotropic diffusion
- A4, B4 and C4 show the output of the bilateral filter with different degree of smoothing in the range dimension
- D4 shows the output of the bilateral filter applied on an CEAD edge-enhanced image instead of the original image

Table 7.1: Ultrasound image features. Annotation of Figure 7.1

	A	B	C	D
1	Original image	Logarithm	Logarithm with outlier suppression	Gaussian filter, $\sigma = 2$
2	Local entropy	Local log-entropy	Local variance	3x3 Median filter
3	Total variation, $\lambda = 5$	Total variation on logarithm image, $\lambda = 5$	CEAD[1], more smoothing	CEAD, less smoothing
4	Bilateral filter, σ_r small	Bilateral filter, σ_r medium	Bilateral filter, σ_r large	CEAD, followed by bilateral filter σ_r small
5	SRAD with log, 10 iterations	SRAD with log, 20 iterations	SRAD, 10 iterations	SRAD, 40 iterations

The feature selection procedure was performed based on manual delineations of a number of 2D slices extracted from multiple 3D carotid ultrasound volumes. Although using such manual delineations is imperfect due to incompleteness, this option is preferable to selecting features by qualitatively assessing the degree to which they discriminate between the artery and background. These delineations provide a measure of "ground truth" and

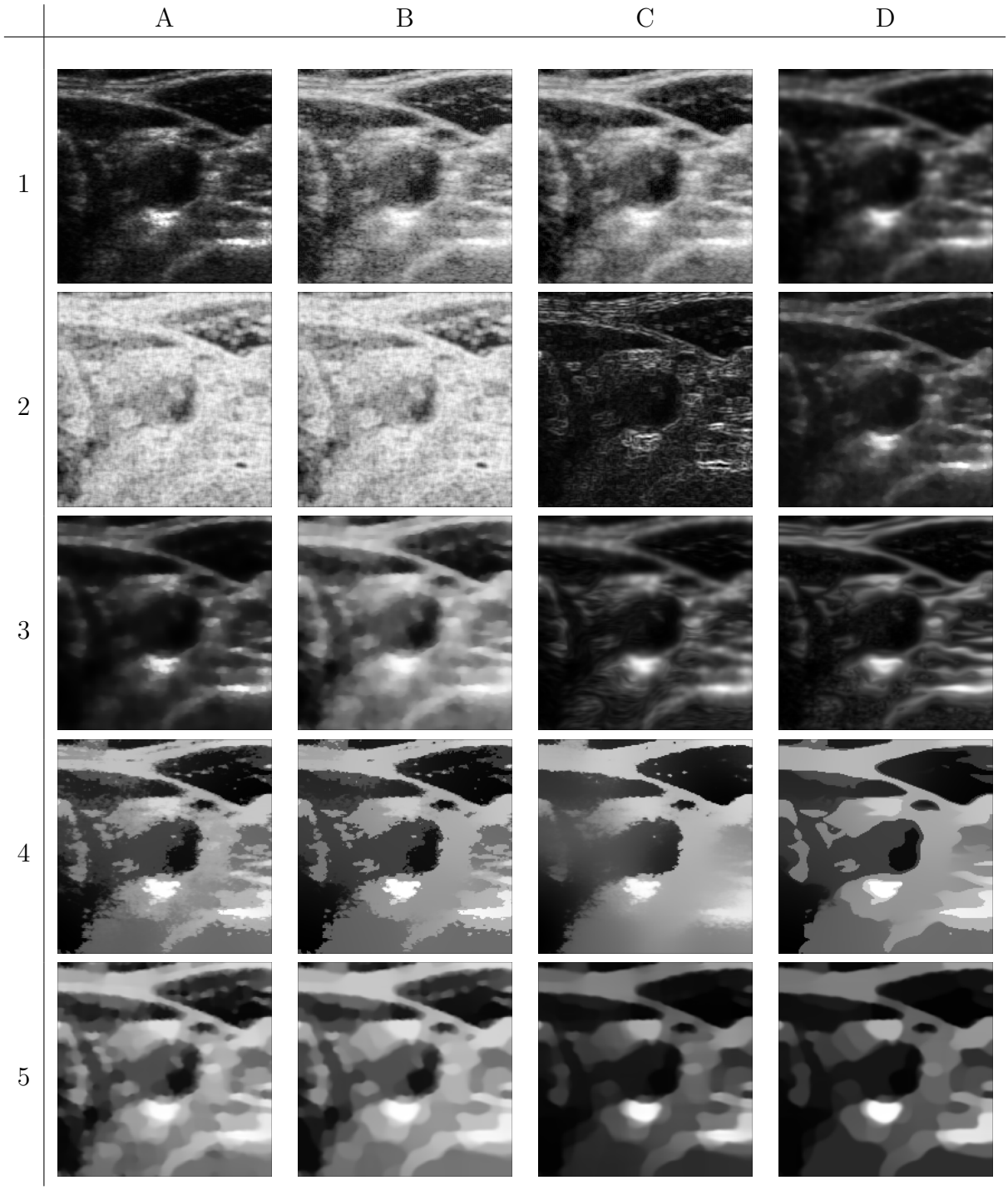


Figure 7.1: Examples of ultrasound image features. See annotation in Table 7.1.

allow one to differentiate between empirical densities inside and outside the artery, $P_A(\mathbf{z})$ and $P_B(\mathbf{z})$, averaged across all 2D slices.

The search strategy for selecting the features is based on the backwards elimination[31] approach. The procedure starts with all candidate features $\mathbf{z} = \{z_i\}_{i=1}^{d}$. At each iteration, a pair of features that are most similar to each other via the Bhattacharyya metric in (7.1) are chosen, defined by the combined similarity

$$S_{ij} = S(z_i, z_j) = B(P_A(z_i), P_A(z_j)) + B(P_B(z_i), P_B(z_j))). \tag{7.2}$$

One of these two features z_i and z_j is redundant, so it is removed. The feature that ends up removed from $\mathbf{z}$ is the one that has the smaller artery-background similarity $B(P_A(z_k), P_B(z_k))$. The algorithm proceeds until the number of features has reached the desired number, in this case 3. The search algorithm is shown below:

$\mathbf{z} \leftarrow$ feature-extraction(I)
while size$(\mathbf{z}) > 3$ **do**
  $\{z_1, z_2\} = \underset{z_i, z_j \in \mathbf{z}}{\arg\max}\ S(z_i, z_j)$
  if $B(P_A(z_2), P_B(z_2)) > B(P_A(z_1), P_B(z_1))$ **then**
    $\mathbf{z} \leftarrow \mathbf{z} \backslash \{z_2\}$
  else
    $\mathbf{z} \leftarrow \mathbf{z} \backslash \{z_1\}$
  end if
end while

As a result of this procedure, the three most "dissimilar" features are selected; their empirical densities inside and outside the artery are shown in Subplots B,C and D of Figure 7.2, along with the densities obtained using the original image shown in Subplot A for comparison. It can be visually confirmed that the densities P_A and P_B are more dissimilar in the three selected features than in the original intensity image.

The remainder of this chapter uses Bilateral filter with CEAD, Local log-entropy and log-compression with suppressed outliers as the three features with which to segment the carotid artery. Although using more features may ostensibly improve the accuracy of segmentation, the computational burden increases accordingly, with three being an acceptable tradeoff. Although the algorithm did not select the proposed ultrasound decomposition features, discussed in the next section, they were the last to be eliminated.

7.1.2 Ultrasound decomposition

The proposed ultrasound decomposition framework, designed to produce piecewise continuous and speckle components (resp. u and v), is illustrated using two ultrasound images.

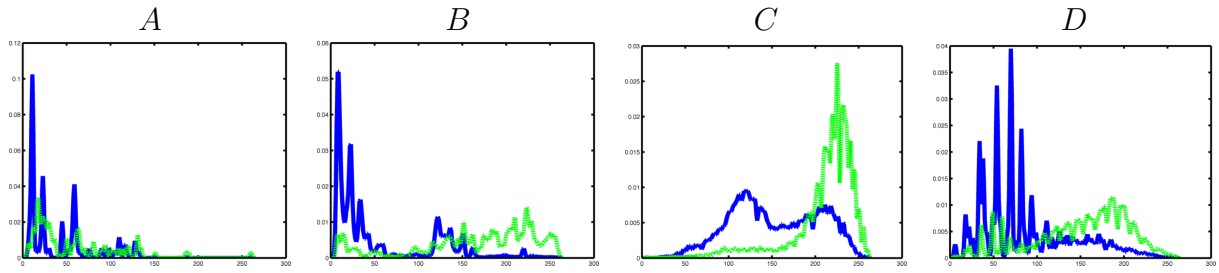


Figure 7.2: Empirical densities P_A (dark solid line) and P_B (light dotted line). A: Original image I; B: Bilateral filter with CEAD; C: Local log-entropy; D: log-compression with suppressed outliers.

First, Figure 7.3, analysing the decomposition of an image of the thyroid gland, shows the initial components (Subplots E & F), as well as the final result (Subplots G & H). Since, as discussed in Section 6.3.2, the component u is found through the use of TV denoising, it is instructive to compare the final decomposition result (Subplots G & H) with a similar result of standalone TV denoising (Subplots C & D). The thyroid gland, which appears as the dark object occupying most of the right half of the image, is more homogenous in the u component in Subplot G than it is in the TV denoising result in Subplot C. The surrounding tissue is similarly more homogenous in the decomposition result.

The same analysis is performed on a two-dimensional cross-section slice of the 3D carotid ultrasound volume, shown in Figure 7.4. As in the case of thyroid, the piecewise-smooth component u in Subplot G exhibits more homogeneity both inside and outside the artery, whose convex cross-section is seen in the centre of the image.

In both figures, note the effect of performing the decomposition using regularization terms for both the smooth and the speckle components. Subplots E and F show the initial components, obtained through TV denoising as discussed in Section 6.3.2. The weight of the TV term, λ_1, is unchanged. Thus, the differences between Subplots E and G (as well as between F and H) show the effect of adding the speckle component to the energy functional. While TV denoising by itself succeeds in removing most speckle, the decomposition technique results in an image which is more regularized while maintaining sharp edges.

In order to quantitatively confirm the latter observation, empirical densities inside and outside the artery were compared, using the same manual delineation technique as in the preceding section. The results are shown in Figure 7.5 and confirm the observation: the Bhattacharyya similarity measure is smaller for the decomposition result than for standalone TV denoising result[2].

[2]Note that these empirical densities (Figure 7.5) are obtained only from Subplots 7.4C and 7.4G, not

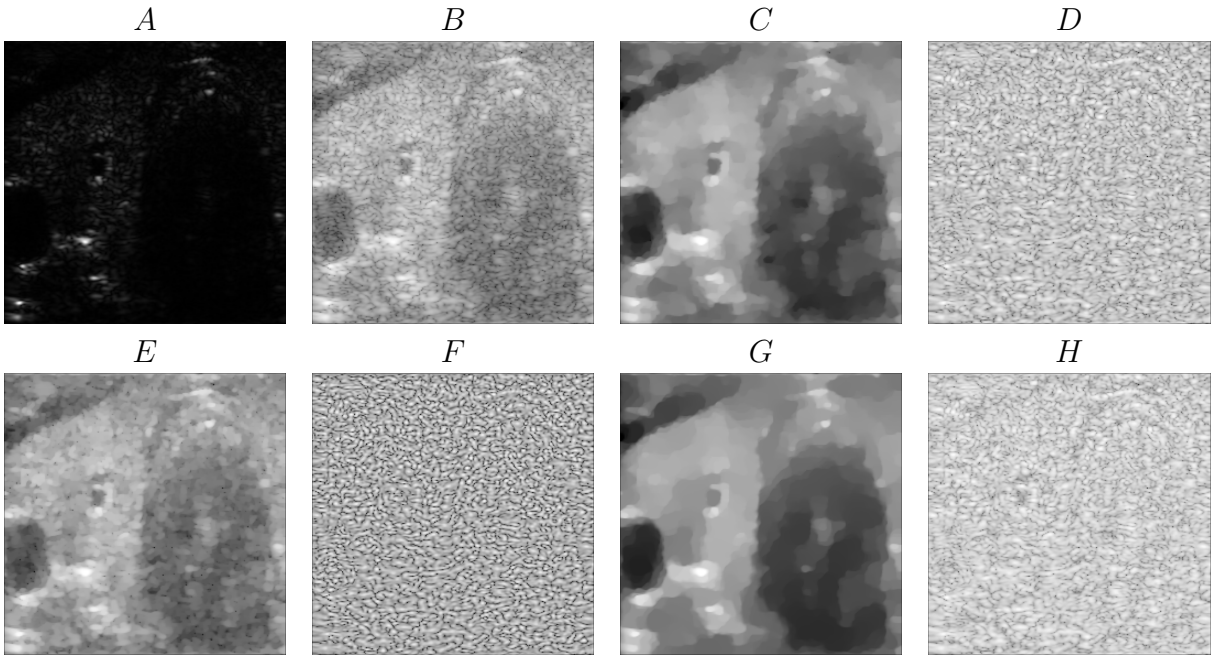


Figure 7.3: Decomposition of a thyroid ultrasound image. (A) The original envelope image; (B) Log-compressed image; (C,D) TV denoising result and residual, $\lambda = 2.8$; (E,F) Initialization of the decomposition algorithm (TV denoising, $\lambda = 1$); (G,H) The final decomposition result, $\lambda_1 = 1$, $\lambda_2 = 0.2$.

In conclusion, the demonstration of the proposed ultrasound decomposition algorithm shows that the resulting components have the intended properties, and thus can be considered alongside previously devised features in the feature selection procedure. Moreover, the improved tissue selectivity of the piecewise smooth component can be seen visually and confirmed by the reduction in the Bhattacharyya similarity measure, when contrasted with a comparable TV denoising result.

7.2 Edge detection

As described in Section 3.2.3, the geodesic length energy encourages the segmentation boundary to overlap as much as possible with the edges present in the image. Many ways of defining the edge indicator function $g(I)$ have been proposed. Conceptually, this function performs the same role as the scalar diffusivity function for non-linear diffusion

from multiple slices of several volumes as in Section 7.1.1. This allows direct comparison between Figures 7.4 and 7.5.

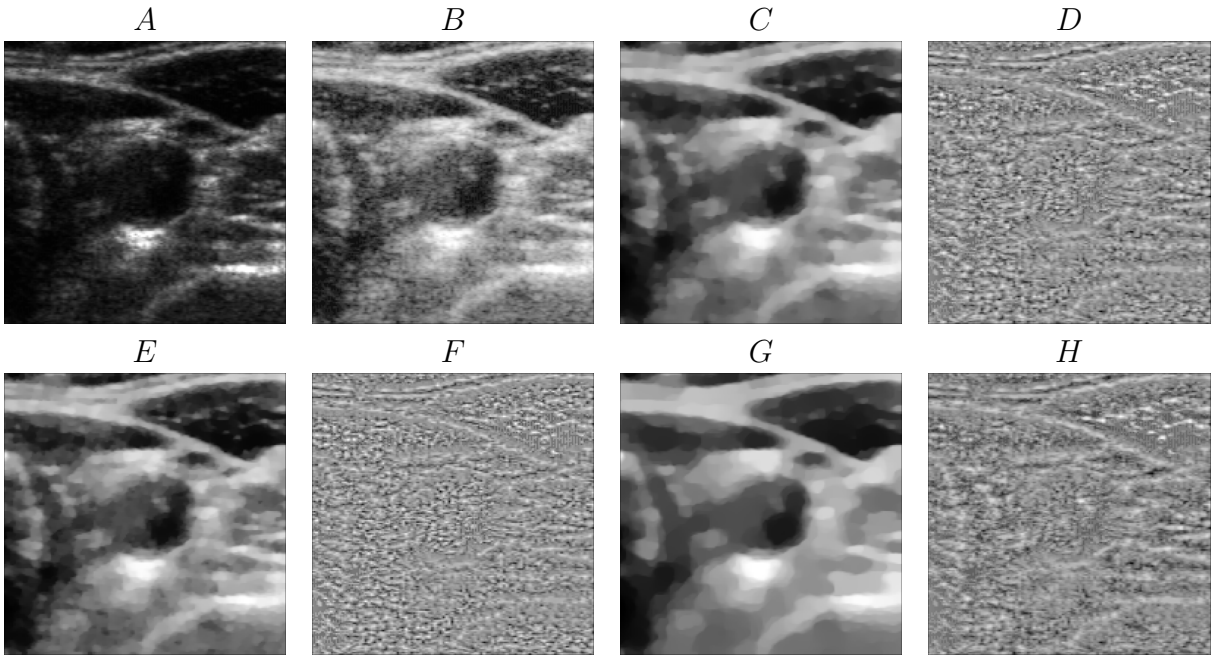


Figure 7.4: Decomposition of a carotid ultrasound image. (A) The original envelope image; (B) Log-compressed image; (C,D) TV denoising result and residual, $\lambda = 0.67$; (E,F) Initialization of the decomposition algorithm (TV denoising, $\lambda = 0.4$); (G,H) The final decomposition result, $\lambda_1 = 0.4$, $\lambda_2 = 0.5$.

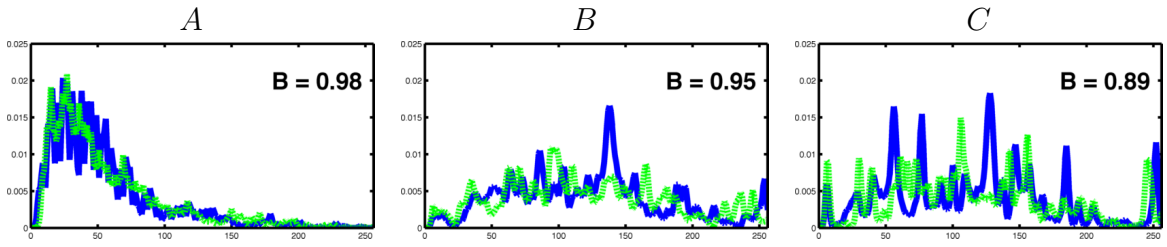


Figure 7.5: Empirical densities P_A (dark solid line) and P_B (light dotted line) and their Bhattacharyya similarity values. (A) Original image I; (B) TV denoising result depicted in Figure 7.4C; (C) final decomposition result depicted in Figure 7.4G.

filtering, discussed in Section 4.2.1, in the cases when the desired output of the filter is a smoothed image with preserved edges. For example, one can define $g(\cdot)$ as

$$g(Y) = \frac{1}{1 + \alpha Y^{\beta}}, \tag{7.3}$$

where parameters α and β control the smoothness of the resulting edges and can be set to 0.1 and 2 respectively if the input Y is normalized. However, computing edge-detector

functions directly on the image produces noisy results, as evidenced by Subplot B of Figures 7.6 and 7.7. The calculation of edges used in this work proceeds by first pre-processing the image to make it piecewise smooth and emphasize strong edges, and then applying (7.3) on the result. In Figure 7.6 and 7.7, this is shown for respectively the bilateral filter on its own (Subplot C) and in combination with CEAD (Subplot D). The resulting edge indicator is a lot less noisy, especially inside the artery. Since the segmentation contour is initialized inside the artery and expands outwards, this is a desirable property. Ideally, the first strong edge that the contour encounters as it expands corresponds to the arterial wall. If significant plaque is present, this may not be the case, and this region may be misclassified as background. In this case, the outcome of segmentation depends strongly on initialization: initializing the contour to overlap with both the lumen and plaque improves the resulting delineation.

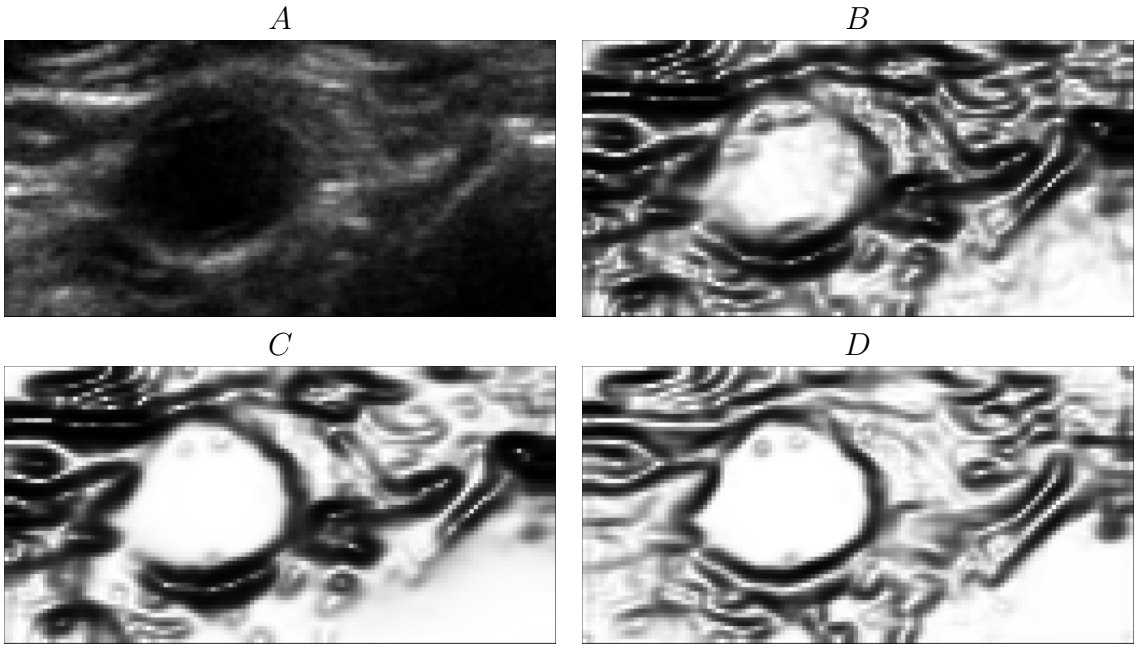


Figure 7.6: Extraction of edges from pre-processed carotid ultrasound volumes (example 1). (A) The original 2D cross-section; (B) Application of (7.3) on the original image; (C) Application of (7.3) on the output of a bilateral filter; (D) Application of (7.3) on the output of a bilateral filter followed by CEAD.

Although the results shown in this section and used in the overall segmentation algorithm employ the bilateral filter, any other edge-preserving filter may be used instead, including the proposed ultrasound decomposition approach. Regardless of the method used, this pre-processing step is intended to preserve strong edges and remove weak ones. Applying such a filter before extracting edges results in better segmentation performance.

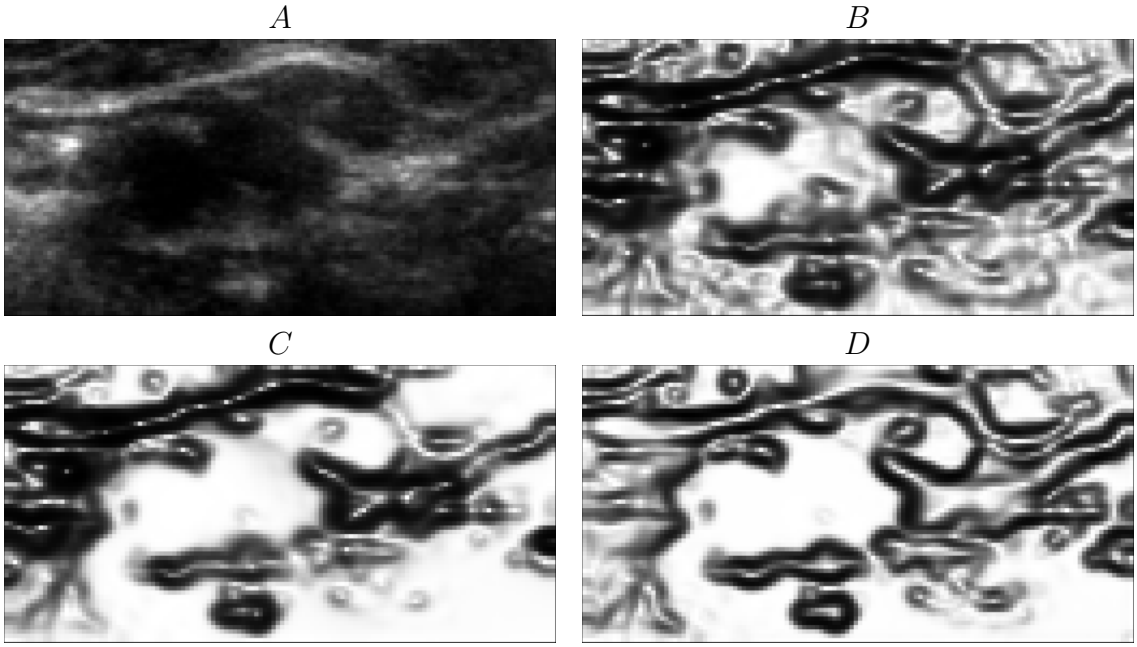


Figure 7.7: Extraction of edges from pre-processed carotid ultrasound volumes (example 2). (A) The original 2D cross-section; (B) Application of (7.3) on the original image; (C) Application of (7.3) on the output of a bilateral filter; (D) Application of (7.3) on the output of a bilateral filter followed by CEAD.

7.3 Segmentation of carotid ultrasound images

The proposed segmentation algorithm was tested on a set of clinical 3D carotid ultrasound images. Manual delineations of these images by medical experts were also available, but restricted to the bifurcation area, both due to the fact that it is the most important part of the image from a clinical standpoint, and because of the laborious nature of this delineation process. In total, about ten two-dimensional delineations were performed in each volume. The proposed algorithm is compared to these delineations both visually and in terms of quantitative measures.

The manual delineations were stored in a form of a collection of N points $\mathcal{M} = \{\mathbf{x}^i \in \mathcal{R}^3\}_{i=1}^N$, with each of these points lying on the outer boundary of the carotid artery. In order to quantitatively assess the correspondence between the automatic segmentation boundary and these points, their distances to the segmentation result were used. Since the signed distance function ϕ already embeds this information, computing these distances amounts to simply evaluating ϕ at every point in $\mathcal{M}$. In this work, the degree to which the automatically obtained segmentation boundary deviates from $\mathcal{M}$ was quantified by three

measures. First, one can simply add the signed distances:

$$MD(\phi) = \sum_{\mathbf{y} \in \mathcal{M}} \phi(\mathbf{y}), \tag{7.4}$$

which will indicate how *biased* the zero level set of ϕ is. Using the definition of ϕ in 6.4, a positive value of "mean deviation" $MD(\phi)$ means that the zero level set of ϕ under-segments the artery, and conversely negative $MD(\phi)$ implies over-segmentation. Another useful measure of error is a "mean absolute deviation" defined as the sum of the absolute distances,

$$MAD(\phi) = \sum_{\mathbf{y} \in \mathcal{M}} |\phi(\mathbf{y})|, \tag{7.5}$$

which will be a measure of overall correspondence between $\mathcal{M}$ and ϕ. Finally, it is useful to look at the "maximum absolute deviation",

$$MaxAD(\phi) = \max_{\mathbf{y} \in \mathcal{M}} |\phi(\mathbf{y})|. \tag{7.6}$$

In this thesis, a variety of possible segmentation approaches have been proposed. In particular, a multitude of image features were presented, and several were selected due to their usefulness in discerning the region of interest from the background. Thus, in the experimental part of the thesis, significance of proper feature selection is demonstrated by comparing the segmentation result with one which is obtained using the original image intensity only. Additionally, the importance of incorporating prior information for carotid segmentation has been advocated. Accordingly, in the experiments presented in this work, the effectiveness of the proposed weak convexity priors is demonstrated.

In order to establish the importance of proper feature selection, a comparison of the quality of segmentation was made using a set of three features found using the method of Section 7.1.1 versus using the image intensity by itself. The comparison is presented in Figure 7.8. In order to aid the visual assessment of the segmentation contour's accuracy, the original ultrasound images of five different cross-sections are shown in the leftmost subplot of each row. Rows 1 and 2 also show the manual delineation of that particular 2D slice in the middle subplot. The rightmost subplots show the automatic segmentation results. Using the three previously selected features results in more accurate results than using only intensity. The intensity-only result has a worse correspondence with the visual extent of the artery in Row 3. In Row 2, it can be seen that the bifurcation occurs in a different location. Finally, looking at the Row 1, it can be seen that intensity-based result under-segments the artery. This can be confirmed by examining the median-deviation (MD) values in Table 7.2 for volume V1. The visual and quantitative comparison provides justification for extracting and employing the three features in the region-based energy embedded in the segmentation algorithm.

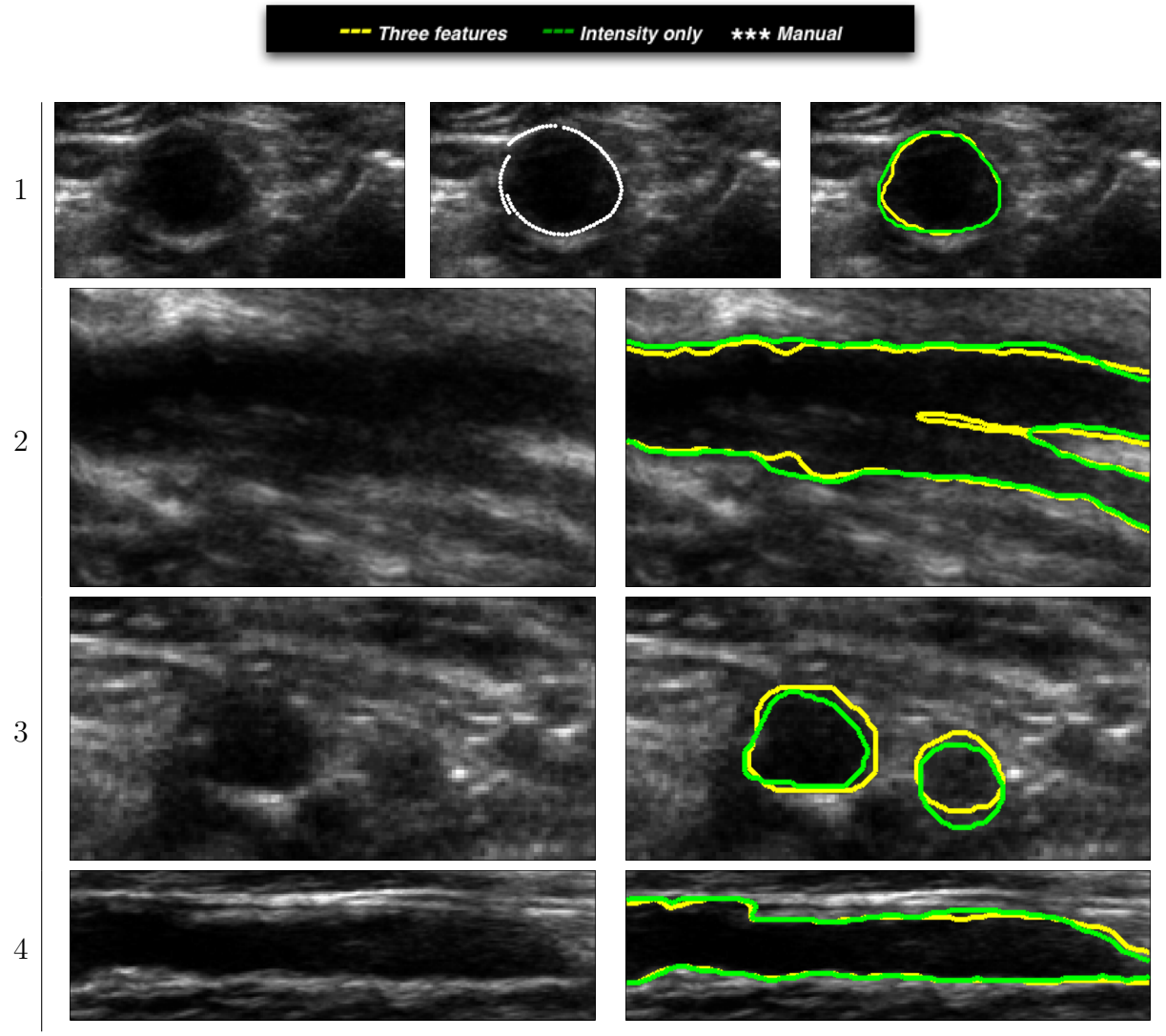


Figure 7.8: The effect of using multiple features from Sec. 7.1.1 on segmentation quality.

The comparison of the algorithm's performance with and without the weak convexity prior energy is shown in Figure 7.9. Rows 1 and 2 show both the manual delineation (middle column) and the automatic segmentation contour with and without the convexity prior energy (rightmost column). Imposing convexity regularization succeeds in fixing some non-convex artifacts in the segmentation boundary. In Row 2, it can be seen that the manual delineation indicates the presence of two arteries, while the automatic segmentation results in a single contour. However, if one examines the original image it can be seen that the arteries are starting to bifurcate but have not separated completely in this location. In

this case, one of the concavities is anatomically correct, but smoothed over by the convexity prior. However, as Row 3 shows, when the two branches of the common carotid artery have clearly separated, they are correctly shown as two separate objects. Moreover, the use of convexity prior in this case results in a more visually accurate segmentation or the right branch of the artery (unfortunately no manual segmentation is available for that portion of the volume). Row 4 shows that in those locations where tissue boundaries are strong and smooth, segmentation results with and without the convexity prior match closely.

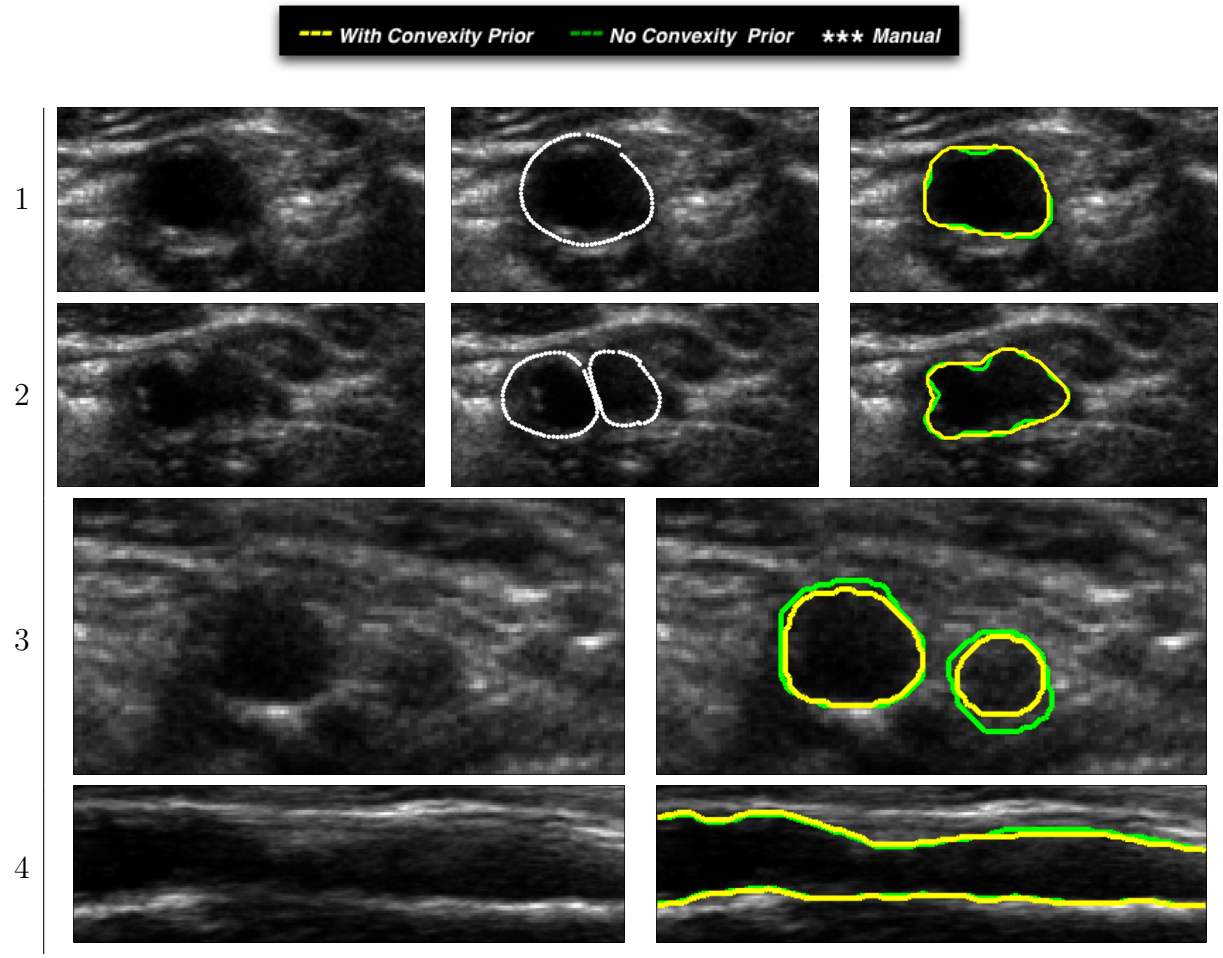


Figure 7.9: The effect of using the weak convexity prior on segmentation quality.

Figure 7.10 shows a rendering of the 3D surface identified as the carotid artery inside a carotid ultrasound volume. It can thus be visually confirmed that the proposed 3D segmentation algorithm converges to an anatomically plausible configuration.

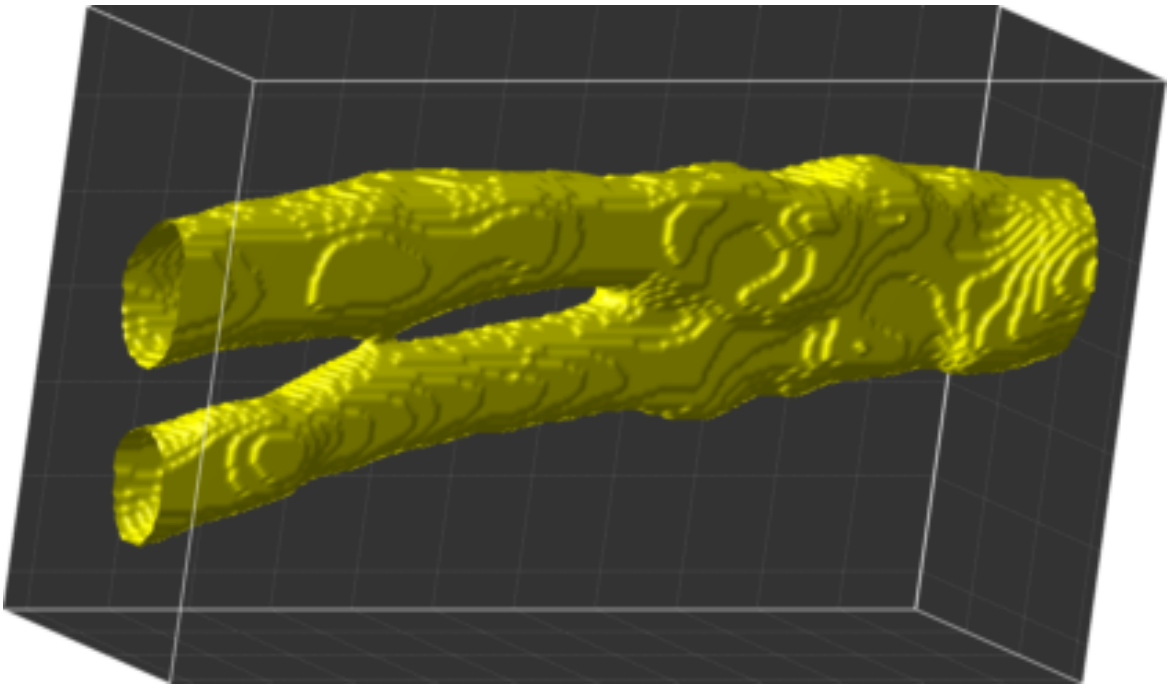


Figure 7.10: Three-dimensional rendering of a typical segmentation boundary.

Table 7.2: Correspondence of manual and automatic segmentation, volume V1

	Full	No Prior	Intensity Only
MD	0.22	**0.09**	0.96
MAD	**2.41**	2.52	2.47
MaxAD	**12.5**	12.8	13.5

Segmentation results using the above configuration for two different carotid ultrasound images are shown in Figures 7.11 and 7.12. These two images show the progression of atherosclerosis in the carotid artery of the same patient over time. The classification of pixels is generally visually good. The most notable problem with these results is the failure to identify two separate branches. Aside from that, the outer boundary of the artery is identified quite closely. In Table 7.3, the error measures for these segmentation results are shown under rows V2 and V3; they are consistent with each other, and with the first examined volume.

In conclusion, automatic segmentation results are presented in this section, which indicate the viability of the chosen approach in labelling regions of interest in carotid ultrasound images in an automated way. These results bear close correspondence with manual delineations done by clinical experts, both visually and in terms of shape mismatch error

Table 7.3: Correspondence of manual and automatic segmentation, volumes V2 and V3.

Volume	MD	MAD	MaxAD
V2	0.2	2.37	14.8
V3	-0.15	2.39	10.2

measures. Some problems exist with the automatically obtained segmentation boundaries, such as a tendency to incorrectly identify the two branch arteries as a single object when the differentiation between them is difficult to determine due to their proximity or image artifacts such as shadowing and obscure boundaries. Despite such shortcomings, the results are promising and demonstrate the viability of the chosen approach in labelling regions of interest in carotid ultrasound images in an automated way.

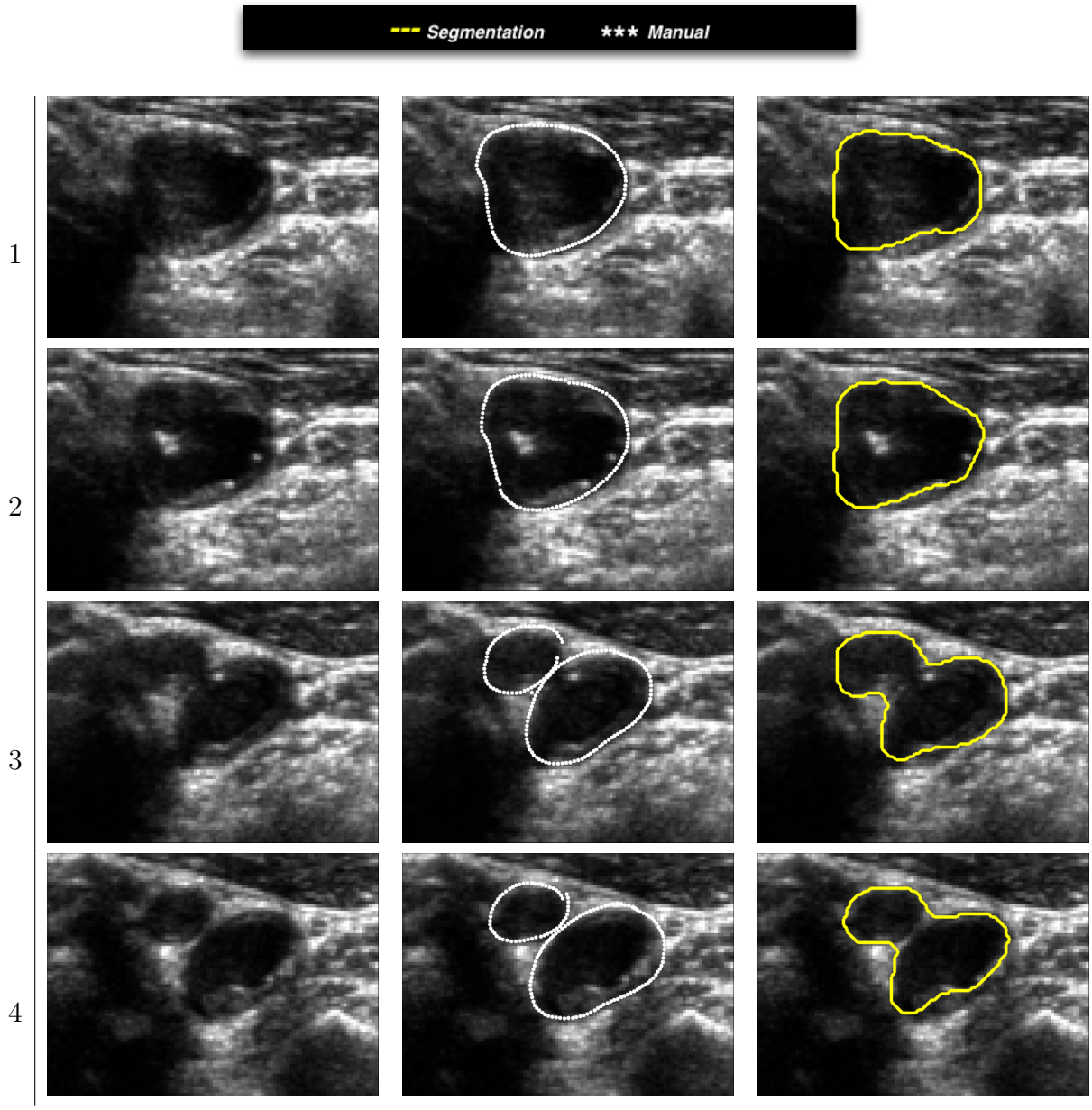


Figure 7.11: 2D cross-sections of a segmentation result, carotid volume 2.

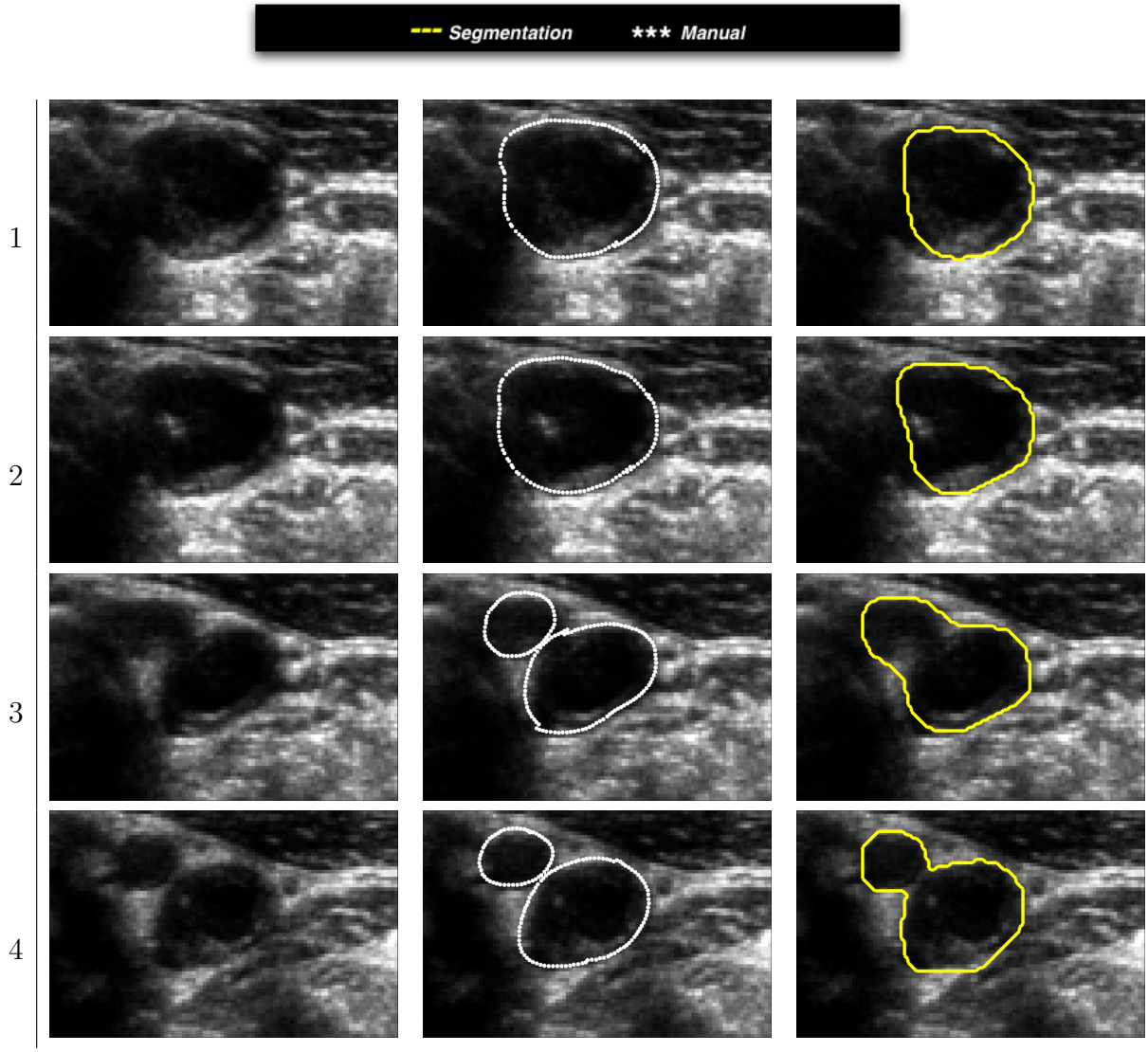


Figure 7.12: 2D cross-sections of a segmentation result, carotid volume 3.

Chapter 8

Conclusions and future work

In this work, a theoretically and numerically sound framework for segmentation of 3D carotid ultrasound images was presented and evaluated. The goal of this work is to extract from the image two regions of interest, the artery lumen and the atherosclerotic plaque, particularly around the bifurcation region susceptible to this disease. The developed algorithm is intended to track the approximate location of the outer wall of the carotid artery, separating both the lumen and plaque from the background. Subsequent measurements of the extent and morphology of the plaque can be used for diagnostic purposes, as well as in drug trials.

Different sources of information are combined in an active contours energy minimization framework. The segmentation boundary is encouraged to overlap as much as possible with the image edges via a geodesic length energy, and using an edge detector which indicates strong interfaces within the image while filtering out noisy edges. Furthermore, the statistical dissimilarity between the region of interest and its background is exploited using two alternative energy terms, differing in modelling assumptions about the nature of the underlying pdf's. This analysis of regional statistics is performed on a number of features derived from the original intensity image.

One of the contributions of this work is a novel set of image features, obtained by performing a decomposition of the ultrasound image into a piecewise continuous component representing the underlying tissue and a texture component containing the speckle pattern which is characteristic of ultrasonic imaging. The decomposition scheme is derived in a MAP statistical framework and efficiently solved using the alternating minimization technique, employing total variation denoising as one of its principal building blocks. The resulting features compare favourably to results of total variation denoising, both qualitatively and quantitatively.

In addition to information derived from the observed image, a-priori knowledge about the anatomical plausibility of carotid artery shapes is incorporated using a novel weak

shape prior which encourages the segmentation boundary to be as convex as possible. One of the chief advantages of using the latter energy is that it does not impose restrictive constraints on the segmentation process, acting as a flexible regularizer.

Although promising results are shown in Chapter 7, further work is necessary before this approach is suitable for clinical use. The analysis is performed on a limited number of examples, and thus it would be beneficial to determine the efficacy of the algorithm on more data sets. Furthermore, it is necessary to further segment the artery into the lumen and plaque regions and to compare that delineation with their manually obtained counterpart.

A critical stage of the proposed algorithm is initialization of the level set function. Although the procedure presented in Section 6.2.2 is intuitive, requires little manual input from the user and produced good results, other initializers may improve the performance. For example, instead of the cylindrical shape, the initial contour may be obtained by thresholding, region growing or another quick and simple segmentation algorithm, and then refined using the algorithm presented in this work.

One important limitation of the ultrasound decomposition technique presented in Section 5.1 is the assumption that the speckle noise present in the envelope ultrasound image is statistically independent. This assumption considerably simplifies the decomposition algorithm and makes it computationally viable. However, considerable correlation exists between neighbouring pixels in ultrasound images[36]. One way to account for this discrepancy is to de-correlate the ultrasound image prior to decomposition. However, this requires access to the original RF signal, which was not available for this study. This data can, however, be readily acquired from an ultrasound scanner and the noise present in it de-correlated, using for example the method outlined in[36].

In conclusion, despite many improvements which can be devised to the proposed algorithm, it represents a promising and viable approach for the segmentation of 3D carotid ultrasound volumes. The performance varies from one data set to another, due to considerable variability in the scale, resolution and discernibility of the acquired data. Given the quality of the available data, however, the successful segmentation results bear strong visual and quantitative correspondence to the location of the carotid artery's outer wall.

Bibliography

[1] Federal Standard 1037C. Telecommunications: Glossary of Telecommunication Terms. National Communication System Technology & Standards Division, 1996. 12

[2] T. Binder, M. Sussner, D. Moertl, T. Strohmer, H. Baumgartner, G. Maurer, and G. Porenta. Artificial neural networks and spatial temporal contour linking for automated endocardial contour detection on echocardiograms: A novel approach to determine left ventricular contractile function. *Ultrasound Med. Biol.*, 25(7):1069–1076, 1999. 22, 23

[3] A. Blake and M. Isard. *Active contours*. Springer London, 1998. 29, 47

[4] A.G. Bruce, D.L. Donoho, H.Y. Gao, and R.D. Martin. Denoising and robust nonlinear wavelet analysis. In *Proc. SPIE*, volume 2242, pages 325–336. SPIE, 1994. 61

[5] A. Buades, B. Coll, and J.M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Model. Simul.*, 4(2):490–530, 2006. 33

[6] L.R. Caplan. Diagnosis and treatment of ischemic stroke. *J. Am. Med. Ass.*, 266(17):2413–2418, 1991. 2

[7] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *Int. J. Comput. Vis.*, 22(1):61–79, 1997. 25

[8] A. Chambolle and P.L. Lions. Image recovery via total variation minimization and related problems. *Numer. Math*, 76:167–188, 1997. 59

[9] T. Chan and J. Shen. *Image Processing and Analysis: Variational, PDE, Wavelet, And Stochastic Methods*. SIAM, Philadelphia, PA, USA, 2005. 40

[10] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. Image Process.*, 10(2):266–277, 2001. 25, 26, 34, 49, 50

[11] D. Cremers, M. Rousson, and R. Deriche. A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. *Int. J. Comput. Vis.*, 72(2):195–215, 2007. 25, 26, 29

[12] A. Fenster, D.B. Downey, and H.N. Cardinal. Three-dimensional ultrasound imaging. *Phys. Med. Biol.*, 46:R67, 2001. 15

[13] A. Fenster, A. Landry, D.B. Downey, R.A. Hegele, and J.D. Spence. 3D ultrasound imaging of the carotid arteries. *Curr. Drug Targets Cardiovasc. Haematol. Disord.*, 4(2):161–175, 2004. 14, 15, 16

[14] R.A.H. Fernandes da Rocha. *Image segmentation and reconstruction of 3D surfaces from carotid ultrasound images.* PhD thesis, Universidade do Porto, 2007. 21, 22

[15] D. Freedman and T. Zhang. Active contours for tracking distributions. *IEEE Trans. Image Process.*, 13(4):518–526, 2004. 53

[16] A. Gee, R. Prager, G. Treece, and L. Berman. Engineering a freehand 3D ultrasound system. *Pattern Recognit. Lett.*, 24(4-5):757–777, 2003. 14, 15, 16

[17] M. Giles and P. Rothwell. Medical treatment for carotid stenosis. In J. Gillard, M. Graves, T. Hatsukami, and C. Yuan, editors, *Carotid Disease*, pages 59–71. Cambridge University Press, 2006. 3

[18] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky. Fast geodesic active contours. *IEEE Trans. Image Process.*, 10(10):1467–1475, 2001. 25, 50, 55, 56

[19] M. Graves, J. U-King-Im, and J. Gillard. Magnetic resonance angiography of carotid artery. In J. Gillard, M. Graves, T. Hatsukami, and C. Yuan, editors, *Carotid Disease*, pages 140–157. Cambridge University Press, 2006. 4

[20] C. Guy and D. Ffytche. *An introduction to the principles of medical imaging.* Imperial College Press, 2005. 5, 6, 7, 47

[21] A. Hamou, S. Osman, and M. El-Sakka. Carotid ultrasound segmentation using DP active contours. *Proc. Image Anal. Recog.*, pages 961–971, 2007. 41

[22] S. Hu, E.A. Hoffman, J.M. Reinhardt, et al. Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images. *IEEE Trans. Med. Imaging*, 20(6):490–498, 2001. 21

[23] J.A. Jensen. *Estimation of blood velocities using ultrasound: A signal processing approach.* Cambridge University Press, 1996. 11

[24] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. Comput. Vis.*, 1(4):321–331, 1988. 29

[25] J. Kittler and J. Illingworth. Minimum error thresholding. *Pattern Recognit.*, 19(1):41–48, 1986. 21, 22

[26] A. Krasinski, B. Chiu, J.D. Spence, A. Fenster, and G. Parraga. Three-dimensional Ultrasound Quantification of Intensive Statin Treatment of Carotid Atherosclerosis. *Arthroscopy*, 35(11):1763–1772, 2009. 3, 16, 18

[27] Shawn Lankton. Sparse Field Methods. Technical report, 2009. Available at `http://www.shawnlankton.com/wp-content/uploads/articles/lankton-sfm-TR-2009.pdf`. 55

[28] C. Li, C. Xu, C. Gui, and M.D. Fox. Level set evolution without re-initialization: A new variational formulation. In *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pages 430–436. Citeseer, 2005. 53

[29] C.H. Li and C.K. Lee. Minimum cross entropy thresholding. *Pattern Recognit.*, 26(4):617–625, 1993. 21, 22

[30] S. Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer, 2009. 39, 40

[31] H. Liu and H. Motoda. *Feature selection for knowledge discovery and data mining*. Springer, 1998. 67

[32] J. Malcolm, Y. Rathi, A. Yezzi, and A. Tannenbaum. Fast approximate surface evolution in arbitrary dimension. In *SPIE Conference Series*, volume 6914, page 149, 2008. 54, 55

[33] P. Martin, P. Réfrégier, F. Goudail, and F. Guérault. Influence of the Noise Model on Level Set Active Contour Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(6):803, 2004. 26

[34] C.R. Maurer Jr, R. Qi, and V. Raghavan. A Linear Time Algorithm for Computing Exact Euclidean Distance Transforms of Binary Images in Arbitrary Dimensions. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(2):265, 2003. 52, 53

[35] O. Michailovich, Y. Rathi, and A. Tannenbaum. Image segmentation using active contours driven by the Bhattacharyya gradient flow. *IEEE Trans. Image Process.*, 16(11):2787–2801, 2007. 27, 50

[36] O. Michailovich and A. Tannenbaum. De-speckling of ultrasound images. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 53(1):64–78, Jan 2006. 13, 38, 40, 61, 81

[37] S.G. Moursi and M.R. El-Sakka. Active contours initialization for ultrasound carotid artery images. In *Proc. IEEE/ACS International Conference on Computer Systems and Applications*, pages 629–636. IEEE Computer Society, 2008. 41

[38] P. Nederkoorn, C. Majoie, and J. Stam. Computed tomographic angiography of carotid artery stenosis. In J. Gillard, M. Graves, T. Hatsukami, and C. Yuan, editors, *Carotid Disease*, pages 158–165. Cambridge University Press, 2006. 3, 4

[39] J.A. Noble and D. Boukerroui. Ultrasound image segmentation: a survey. *IEEE Trans. Med. Imaging*, 25(8):987–1010, 2006. 11, 25, 47

[40] C. J. O'Donnell. Epidemiology of carotid artery atherosclerosis. In J. Gillard, M. Graves, T. Hatsukami, and C. Yuan, editors, *Carotid Disease*, pages 22–34. Cambridge University Press, 2006. 2, 3

[41] S. Osher and R.P. Fedkiw. *Level set methods and dynamic implicit surfaces*. Springer Verlag, 2003. 52, 53

[42] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11:285–296, 1975. 22

[43] Palmieri, M. "Ultrasound beamforming and image formation". `http://http://dukemil.bme.duke.edu/rsna07/RC632B.pdf`, September 2010. 9

[44] S. Paris and F. Durand. A fast approximation of the bilateral filter using a signal processing approach. *Int. J. Comput. Vis.*, 81(1):24–52, 2009. 33

[45] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(7):629–639, 1990. 34

[46] B. Preim and D. Bartz. *Visualization in medicine: theory, algorithms, and applications*. Morgan Kaufmann Pub, 2007. 21

[47] J. Revkin and D. Lester. Vascular imaging and the clinical development of new pharmaceuticals. In J. Gillard, M. Graves, T. Hatsukami, and C. Yuan, editors, *Carotid Disease*, pages 418–429. Cambridge University Press, 2006. 3

[48] G. Rodriguez-Granillo and P. Serruys. Assessment of carotid plaque with intravascular ultrasound. In J. Gillard, M. Graves, T. Hatsukami, and C. Yuan, editors, *Carotid Disease*, pages 223–234. Cambridge University Press, 2006. 14

[49] L.I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4):259–268, 1992. 35, 40, 41, 59, 62

[50] J.A. Sethian. *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science.* Cambridge University Press, 1999. 47, 48, 52, 54

[51] Y. Shi and W. Karl. A Fast Level Set Method Without Solving PDEs. In *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, pages 97–100, 2005. 54

[52] K.K. Shung, M.B. Smith, and B.M.W. Tsui. *Principles of medical imaging.* Academic Press, 1992. 6, 8, 9, 10

[53] L. Soler, H. Delingette, G. Malandain, J. Montagnat, N. Ayache, C. Koehl, O. Dourthe, B. Malassagne, M. Smith, D. Mutter, et al. Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery. *Computer Aided Surg.*, 6(3):131–142, 2001. 21

[54] Q. Sun, J.A. Hossack, J. Tang, and S.T. Acton. Speckle reducing anisotropic diffusion for 3D ultrasound images. *Comput. Med. Imaging Graph.*, 28(8):461–470, 2004. 34, 35

[55] M. Sussman, P. Smereka, and S. Osher. A level set approach for computing solutions to incompressible two-phase flow. *J. Comput. Phys.*, 114(1):146–159, 1994. 53

[56] M. Takahashi, M. Ashtari, Z. Papp, M. Patel, J. Goldstein, W.M. Maguire, T. Eacobacci, A. Khan, and PG Herman. CT angiography of carotid bifurcation: artifacts and pitfalls in shaded-surface display. *AJR Am. J. Roentgenol.*, 168(3):813, 1997. 21

[57] P. Tanuseputro, D.G. Manuel, M. Leung, K. Nguyen, H. Johansen, et al. Risk factors for cardiovascular disease in Canada. *Can. J. Cardiol.*, 19(11):1249–1260, 2003. 1

[58] T. Thom, N. Haase, W. Rosamond, V.J. Howard, J. Rumsfeld, T. Manolio, Z.J. Zheng, K. Flegal, C. O'Donnell, S. Kittner, et al. Heart disease and stroke statistics–2006 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*, 113(6):e85, 2006. 1

[59] A. Tsai, A. Yezzi Jr, 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 Trans. Med. Imaging*, 22(2):137, 2003. 28, 29, 41

[60] M. Tur, K.C. Chin, and J.W. Goodman. When is speckle noise multiplicative? *Appl. Opt.*, 21(7):1157–1159, 1982. 13

[61] University of Zagreb. "Selected Chapters of Geometry and Mathematics – Introduction". `http://www.grad.hr/itproject_math/Links/sonja/gausseng/introduction/introduction.html`, September 2010. 42

[62] VascularWeb. "Carotid Artery Disease, Stroke, Transient Ischemic Attacks". http://www.vascularweb.org/vascularhealth/Pages/CarotidArteryDisease,Stroke,TransientIschemicAttacks(TIAs).aspx, September 2010. 2

[63] R. Virmani, A. Burke, E. Ladich, and F.D. Kolodgie. Pathology of carotid artery atherosclerotic disease. In J. Gillard, M. Graves, T. Hatsukami, and C. Yuan, editors, *Carotid Disease*, pages 1–21. Cambridge University Press, 2006. 1, 2

[64] R.F. Wagner, S.W. Smith, J.M. Sandrik, and H. Lopez. Statistics of speckle in ultrasound B-scans. *IEEE Trans. Son. Ultrason.*, 30(3):156–163, 1983. 12

[65] J. Weickert. *Anisotropic diffusion in image processing*. B.G. Teubner, 1998. 33, 35, 55

[66] J. Weickert. Coherence-enhancing diffusion filtering. *Int. J. Comput. Vis.*, 31(2):111–127, 1999. 35, 36

[67] J. Weickert, B.M. ter Haar Romeny, and M.A. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Trans. Image Process.*, 7(3):398–410, 1998. 34, 55, 56

[68] B. Weiss. Fast median and bilateral filtering. In *ACM Trans. Graph.*, page 519. ACM, 2006. 33

[69] R.T. Whitaker. A level-set approach to 3D reconstruction from range data. *Int. J. Comput. Vis.*, 29(3):231, 1998. 54, 55

[70] Wikipedia. "Convex Set — Wikipedia, The Free Encyclopedia". http://en.wikipedia.org/w/index.php?title=Convex_set&oldid=382037946, September 2010. 43

[71] Wikipedia. "Level set method — Wikipedia, The Free Encyclopedia". http://en.wikipedia.org/w/index.php?title=Level_set_method&oldid=383137982, September 2010. 48

[72] R.S. Xu, O. Michailovich, I. Solovey, and M. Salama. A probability tracking approach to segmentation of ultrasound prostate images using weak shape priors. In *Proc. SPIE Med. Imaging*, volume 7623, page 76234E, San Diego, CA, 2010. 28

[73] P. Yang and O.A. Basir. Adaptive weighted median filter using local entropy for ultrasonic image denoising. In *Proc. 3rd Int. Symposium Image Signal Process. Anal.*, volume 2, 2003. 32

[74] Y. Yu and S.T. Acton. Speckle reducing anisotropic diffusion. *IEEE Trans. Image Process.*, 11(11):1260–1270, 2002. 34, 35

[75] Y. Zimmer, S. Akselrod, and R. Tepper. The distribution of the local entropy in ultrasound images. *Ultrasound Med. Biol.*, 22(4):431–439, 1996. 32