Segmentation-based Approaches to Pre-operative and Intra-operative Brain Ultrasound Image Registration

A

thesis submitted

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

By

Haradhan Chel

Department of Electronics and Electrical Engineering
Indian Institute of Technology Guwahati
Guwahati, India.

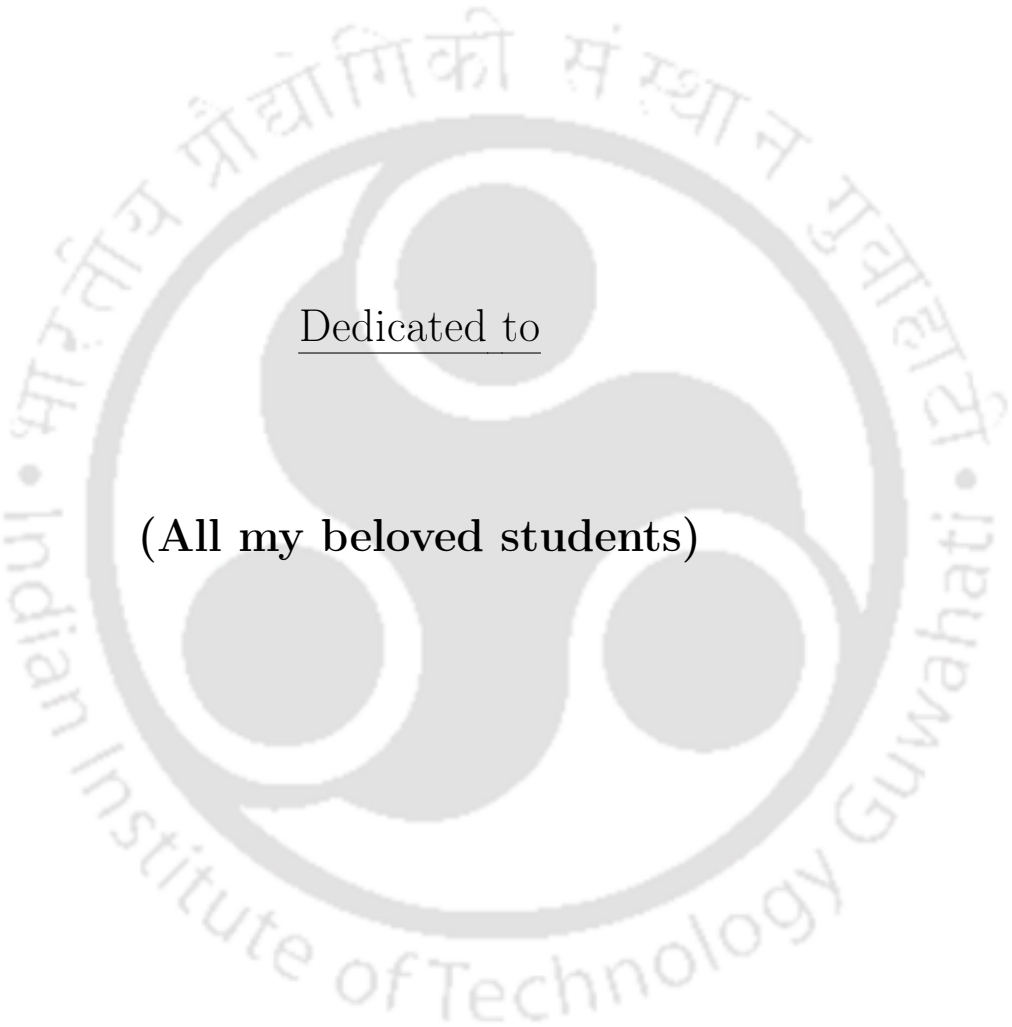
July, 2023

In memory of

My maternal uncle

Late Jagabandhu Ash, my cousins Late Asis Ash,

and Late Subhasis Ash



Dedicated to

(All my beloved students)

Certificate

This is to certify that the thesis entitled "**Segmentation-Based Approaches to Pre-operative and Intra-operative Brain Ultrasound Image Registration**", submitted by **Haradhan Chel** (146102040), a PhD student in the Department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati, for the award of the degree of **Doctor of Philosophy**, has been carried out by him under our supervision and guidance. The thesis has fulfilled all requirements as per the regulations of the Institute and in our opinion has reached the standard needed for submission. The results embodied in this thesis have not been submitted to any other university or institute for the award of any degree or diploma.

Dated:

Guwahati.

Prof. Prabin Kumar Bora

Professor,

Dept. of Electronics and Electrical Engineering,

Indian Institute of Technology Guwahati,

Assam - 781039.

Acknowledgements

This thesis is the culmination of my quest for PhD at the pinnacle of my research career as a student. At this stage I ponder over my journey and take the pleasure of expressing my gratitude to all those who contributed in various degrees to its success.

Foremost, my profound gratitude goes to my supervisors Prof. Prabin kumar Bora who had been my seemingly inexhaustible source of guidance, encouragement and valuable comments on my work. I appreciate the freedom he gave me to think deeply and explore novel ideas and most importantly the patience he have showed in reading and commenting on the manuscripts and several revisions of this dissertation. I would like to pay my sincere gratitude to Dr. Sabita Barua(Bora) for her support, without whom my health could be disastrous. Her food and sympathy acted like blessings and kept me out of physical and mental health problems always.

I take this opportunity to thank my Doctoral Committee members Prof. S. Dandpat, Prof. R. Sinha, and Prof. T.Jacob for annually reviewing my work and their incisive questions that kept me unwavering on my track.

I thank former Director of Central Institute of Technology Kokrajhar Prof. U.S. Dixit for granting me leave during my course work in IIT Guwahati. I am also thankful to Prof. Hemanta Kumar Kalita, the present Director of CITK for his support.

I express my whole-hearted appreciation to my dear friends Dr. Sumantra Chowdhury, Mridul Goswami, and Dr. Avirup Chatterjee for their support in various capacities that are too numerous to mention. I thank master Anindyadyuti(my son) for helping me in various drawings during thesis writing. I would love to pay my homage to Prof. Arunendu mondal, Prof. Amitava Nag, Dr. A. Senapati, Dr. S. maji, and Mr. Naresh paul for their tremendous help in various stages.

A sense of indebtedness is also expressed to Dr. K. K. Ramchiary, Radiologist, City Clinic and Research Centre for his valuable expertise and guidance in problem formulation and critical evaluation of my works.

My deepest sense of gratitude is reserved for my teacher Dr. Debashish Nandi, Mr N. Malakar, my parents Mr. Ramnarayan Chel and Mrs. Sandhya Chel.

Lastly, my sincere gratitude for my wife Mrs. Sunanda Chel and my son Anindyadyuti who reposed faith in my abilities when I had very little in mine

Haradhan Chel

Abstract

An image-guided neuro-navigation(NN) system is essential to aid brain surgery. A fiducial marker-based calibration is used to map the location of the surgical instruments and the diseased region to a pre-operative magnetic resonance (MR) image. The brain is made of soft tissues that floats on cerebrospinal fluid (CSF), and frequently it shifts during the surgical process. Brain-shift prevents the NN system from locating the diseased region. A brain ultrasound (BUS) imaging system is utilized to monitor the surgical procedure. Brain-shift can be modeled by registering the pre-operative brain ultrasound (pBUS) and the corresponding intra-operative brain ultrasound (iBUS) images. pBUS images are captured after removing a bone flap on the skull, and iBUS images are captured during the surgical process. For registering two images, similar regions should be present in those images. The similarity between the pBUS and the corresponding iBUS image is affected for a variety of reasons, which make the registration difficult. This thesis developed three methods to improve the similarity between pBUS and iBUS images and registered them accordingly.

The first method finds the common edge-rich regions from the registering image pair followed by the registration of those edge-rich regions. The novelty of this method lies in extracting the fan-shaped region from the composite US image and adopting a simple patch-based gradient thresholding scheme for locating the edge-rich regions.

The second method proposes a fast and fully automatic method for extracting the hyper-echoic(HE) regions from the registering image pairs. HE regions have higher signal to noise ratios, and therefore, appear similar in both the images. The segmented HE regions are registered minimizing the mean squared error (MSE) by existing optimization techniques. This method introduces thresholding of patchwise average intensity, which is fast. This method develops an edge correction method which rectifies the boundary errors and a cubic B-spline method is used for contour correction and preserving smoothness of the boundary contours.

The third approach adopts a patch-based level-set strategy for segmenting three prominent HE regions namely, the longitudinal fissure, choroid plexus, and tumor, two anechoic regions namely, the ventricles and the resection cavity. These regions appear similar in pBUS and iBUS images and hence, are used as landmarks during registration.

All the three approaches adopt patch-based processing that makes them robust to noise. During the segmentation of the HE and non-HE(NHE) regions, no complex optimization is required, which makes the process fast. Various gradient-based and

heuristic optimizations are used for minimizing MSE during registration. Experiments were conducted on selected image pairs from the RESECT and the BITE datasets. For performance evaluation and comparison, the segmented ground truth images are prepared by annotating the boundaries of different regions in coordination with an expert radiologist. For comparing registration performance, common tag-points are selected from the registering images pairs, and the improvement of mean target registration error (mTRE) after registration is analyzed. Experimental results demonstrate the superiority of the proposed segmentation based approaches.

Contents

List of Figures

List of Tables

List of Acronyms

AC	Active contour
ACWE	Active contour without edge
AIG	Absolute intensity gradient
ASCD	Asymmetric contour distance
BBO	Biogeography-based optimization
BFGS	Broyden Fletcher-Goldfarb-Shano
BLSCF	Bidirectional level-set curve flow
BUS	Brain ultrasound
BITE	Brain images of tumors for evaluation
CC	Correlation coefficient
CCL	Connected component labeling
CG	Conjugate gradient
CNS	Central nervous system
CSF	Cerebro-spinal fluid
CP	Choroid plexus
CT	Computed tomography
DBS	Deep brain stimulation
DL	Deep learning
DNN	Deep neural network
DRLSE	Distance regularized level-set evolution
DTI	Diffusion tensor imaging
ESM	Effective second order minimization
FOV	Field of view
f-MRI	Functional MRI
GA	Genetic algorithm
GAC	Geodesic active Contour
GAN	Generative adversarial network
GD	Gradient descent
GPS	Global positioning system
GVF	Gradient vector flow

HE	Hyperechoic
HGG	High grade glioma
HSI	Habitat suitability index
iMR	Intra-operative MR
iBMR	Intra-operative brain MR
ICP	Iterative closest method
iBUS	Intra-operative brain ultrasound
iUS	Intra-operative US
LBFGS	Limited memory BFGS
LCC	Local cross-correlation
LF	Longitudinal fissure
LGG	Low grade glioma
LMA	Levenberg Marquardt algorithm
LMI	Local mutual information
MI	Mutual information
MINC	Medical imaging net common data format
MR	Magnetic resonance
MSE	Mean squared error
mTRE	Mean target registration error
NCC	Normalized cross correlation
NHAIG	Normalized histogram of the AIG
NHE	Non-hyperechoic
NMI	Normalized mutual information
NN	Neuro-navigation
pBCT	Pre-operative brain computed tomography
pBMR	Pre-operative brain MR
pMR	Pre-operative magnetic resonance
PSO	Particle swarm optimization
pUS	Pre-operative ultrasound
QN	Quasi-Newton
RaPTOR	Robust patch-based correlation ratio
RC	Resection cavity
RESECT	REtrospective evaluation of cerebral tumors
RMSCD	Root mean squared contour distance
SA	Simulated annealing
SGD	Stochastic gradient descent
SIFT	Scale invariant feature transform
SIV	Suitability index variable
SLIC	Simple linear iterative clustering

SSIM	Structural similarity index
SURF	Speeded up robust feature
SSD	Sum of squared distance
ULSCF	Unidirectional level-set curve flow
US	Ultrasound
WLSE	Weighted level-Set evolution

List of Symbols

$\mathbf{V}_S$	Sensed image
$\mathbf{V}_T$	Target Image
$\mathbf{S}$	Binary map
$\mathbf{T}$	Transformation matrix
$\mathbf{\Omega}$	Coordinate of an image space
$\mathbf{\Omega}_R$	Coordinate space of the fan-beam region of a US image
$\rho(V_S, V_T)$	Correlation coefficient between V_S and V_T
$\eta(V_S, V_T)$	Correlation ratio between V_R and V_T
$J(u)$	Objective function for a transformation
$\mathbf{u}$	Solution vector in optimization
$\mathbf{H}$	Hessian matrix
$\mathbf{D}$	Gradient vector
$\mathbf{U}$	Population of the solution vector
$\mathbf{u}^i$	i^{th} member of population U
$\mathbf{u}^{lbest}$	The local-best solution in PSO
$\mathbf{u}^{gbest}$	The global-best solution in PSO
ϵ_1 and ϵ_2	The social and global coefficients of PSO
λ_j	Immigration rate at j^{th} iteration in BBO
μ_j	Emigration rate at j^{th} iteration in BBO
m_j	Mutuation Probability at j^{th} iteration in BBO
w	Width of square patch in during gross segmentation and boundary in BBO
E_b	mTRE before registration
E_a	mTRE after registration
$\phi(t, x, y)$	Level-set function at time t
∇	The gradient operator
$\|.\|$	The absolute operator
$\bigcup$	Union operation
$\bigcap$	Intersection operation
$\in$	Belongs to
$\mathbb{R}^n$	The set of all real numbers in n dimension

1

Introduction

Intra-operative imaging is an essential setup required in brain surgery. It helps to monitor the diseased region and map it to the corresponding region of the pre-operative image. Ultrasound can be used as an imaging modality after craniotomy. The registration of pre-operative and intra-operative brain US images becomes essential because of the movement of the brain during surgery. This chapter discusses aspects of brain surgery, imaging setups, the admissibility of ultrasound imaging in brain surgery, and the registration requirement. Finally, the objectives, the contributions, and the organization of the thesis are presented.

1.1 Introduction

Brain surgery is an essential means to treat a variety of conditions, such as tumors, blood clots, epilepsy, aneurysms, and Parkinson's disease. Successful neurosurgery is concerned with the diagnosis, treatment, and the post-surgical rehabilitation of patients with brain diseases. Stereotactic neurosurgery allows various surgical procedures such as the evacuation of blood clot, biopsies, drug delivery to the diseased region, surgical resection, and deep brain stimulation. These procedures are performed by making an entrance hole on the cranium [163].

When a patient meets a neurosurgeon with a disease, a brain magnetic resonance(MR) or computed tomography(CT) scan is generally advised to know the health status of the brain. If the patient is diagnosed with a disease where surgery is inevitable, the trajectory to reach the diseased region is planned based on the pre-operative brain MR (pBMR) or pre-operative brain CT (pBCT) images. Normally, surgical planning is done based on the position, size, and criticality of the affected regions which are available from the pBMR or pBCT images.

Brain surgery is difficult due to a number of reasons:

a The brain is covered by immovable cranial bones.

b The brain is made of soft tissues, so it can deform easily.

c The brain is floating into the cerebrospinal fluid (CSF), so brain-shift may happen easily.

d The brain is made of various neural tissues; damaging these tissues may impair one or more parts of the brain. For avoiding these difficulties, special surgical setups, image guided monitoring, and skilled professionals are mandatory.

Image-guided brain surgeries have became popular among the image processing community since 80s. Emergent technologies were developed for performing good neurosurgery. The requirements also became stricter in the form of reducing the surgical field, limiting damage to healthy tissues, and preserving the functionality of the brain tissues. Therefore, neuronavigation(NN) systems with pre-operative and intra-operative imaging facilities and the robotic surgery are introduced. Neurosurgery has improved in recent times because of the advances in medical imaging modalities such as ultrasound(US), CT, MR imaging, and functional-MR(f-MR) imaging. These intra-operative brain imaging modalities have helped in reducing damage

to the critical structures adjacent to the lesion [188, 209]. They have enhanced the scope of better surgical planning and performance [42, 58].

The major constraint in a surgical procedure is the brain-shift that also includes deformation of the brain. It occurs due to several factors such as the loss of CSF, action of gravity, resections, drug administration, and applied pressure of the surgical instruments [16, 93, 70], etc. As a result, the navigation accuracy during the surgical procedure is reduced and it becomes difficult for the surgeon to rely on the spatial information [162]. Bucholz *et al.* [15] classified brain-shifts into three categories: low (0-2.9 mm), moderate (2.9-7 mm) and high (>7mm). All these movements are not linear, and hence they impart a different shape. Organs like tentorium cerebelli, corpus callosum, and falx are not prone to more shift and they are under low category. Lateral ventricles, third ventricle, and choroid plexus are in the moderate category. Tumor, blood vessels, sylvian fissure, blood clots, etc. shift more and hence they are under high category. At present, a plenty of research have been carried out by the brain image processing community to correct the brain-shift and retain the scheduled work plan of neurosurgery [8, 55, 79, 130, 179].

The improvements of medical imaging in last two decades have contributed to the major advances in neurosurgery. The combination of new imaging modalities and NN systems provides neurosurgeons with the ability to accurately monitor the surgical anatomy and locating the pathology during a surgical procedure. For instance, an optimal trajectory can be planned prior to the operation to minimize the invasiveness of the neuro-surgical procedure and prevent the perforation of functional neural tissues.

1.2 Steps of brain surgery

Brain surgery is a combination of various steps. The operation theatre for performing brain surgery must be equipped with an intra-operative imaging enabled NN system. The surgery is performed with the help of a set of pBMR or pBCT image slices and real-time intra-operative images. A 3D brain volume model is reconstructed from the pre-operative MR/CT image slices, and the position of the surgical instrument is pointed onto it by an infrared or electromagnetic tracking system. The steps followed during brain surgery are outlined here and adapted from the excellent reference in [41].

Fig. 1.1: A 3-pin skull-clamp for brain surgery [41].

Preparing the patient: First, the patient is made to lie on the operating table and given general anesthesia. After losing consciousness, the patient's head is placed in a 3-pin skull fixation device attached to the table. It makes the patient's head completely immovable during surgery as shown in the Fig. 1.1. For tracking the surgical process, the head is needed to be continuously tracked by the infrared cameras that help to correlate the 3D world coordinate system to the 3D computer brain model created from the pre-operative MRI scans. In Fig. 1.1, the three pinpoints are attached at the fixed positions that are used as references. Instruments are detected by the infrared cameras and displayed on the computer screen of the NN system. The surgical instrument contains a spherical pinpoint attached to it, that dynamically changes its position with the shift of the instrument. It is shown over the pBMR volume rendered by the computer. This tracking system functions similar to a global positioning system (GPS) to find the actual position of the tip of the surgical instrument inside the resection cavity(RC).

Skin incision: An antiseptic is used to prepare the scalp's incision region. Incisions in the skin are frequently made behind the hairline. Only a few mm broad area along the anticipated incision is shaved using a hair-sparing procedure.

Craniotomy: Craniotomy involves removing a part of the skull (cranium) and provides access to the brain. The size of the bone flap can vary depending on the location and severity of the disease. There are many sizes of craniotomy such as a key hole, a quarter-sized hole, and a burr hole [41]. Keyhole craniotomy is larger than quarter-sized; burr hole is smaller. Stereo-

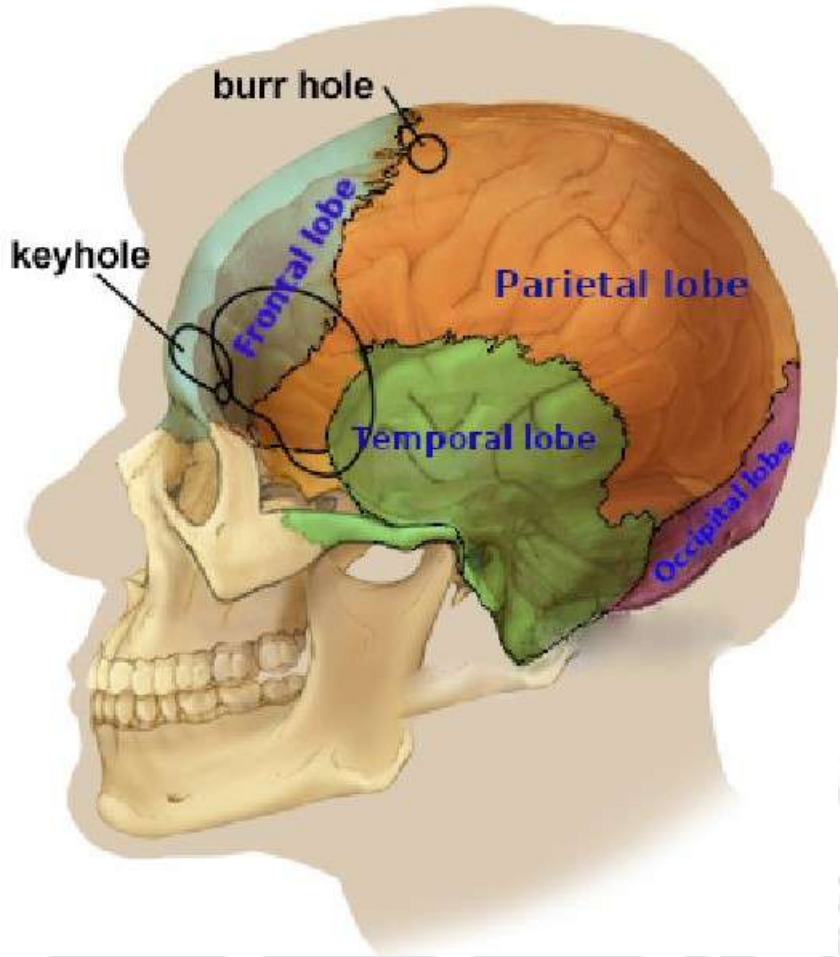


Fig. 1.2: Different lobes of head are shown. A burr hole is small circular hole whereas a key hole is a relatively bigger hole made on the cranial bone [41]

tactic framework, image-guided computer, or endoscope may be employed to successfully insert devices into these tiny holes [41]. Brain tumours, blood clots, aneurysms, severe head injuries, foreign objects, and brain edema are the most common conditions for which a craniotomy is performed

Craniotomies are named according to the area of the cranium to be removed as shown in Fig.1.2. After the completion of the surgery, the bone flap is either replaced or covered with plates and screws. For obtaining pre-operative brain US (pBUS) images, the brain can be scanned after the craniotomy. If the US probe is to fit into the same hole, the diameter of the hole is needed to be enlarged. The enlarged craniotomy is shown in Fig.1.3.a. For avoiding a large opening, another small hole is sometime made on a different location where from the lesion can be visualized better. Making this auxiliary hole is known as a mini-craniotomy and

shown in the Figure 1.3.(b)

Opening of the brain: There are three layers of meninges between the skull and the brain. The outermost thick layer is known as Dura mater, the intermediate layer is known as the arachnoid mater, and the innermost meninges is known as the pia mater. The space between the arachnoid mater and the pia mater layer holds a thin layer of CSF. A retractor (surgical instrument) may be used to gently open a passage between the brain and the skull. Magnification glasses or an operating microscope is generally used to see delicate nerves and vessels.

Repairing/resection: The brain cannot be simply moved aside to access and correct issues, because it is enclosed by the bony skull. For operations deep inside the brain, neurosurgeons employ a number of incredibly small devices. An NN system guides the whole surgical procedure. The diseased region is repaired/resected with the help of the NN system. Extreme cares are taken for the protection of the cranial nerves and arteries during the surgical process.

Closing of the craniotomy: After completion of the repairing of the diseased region, the dura is closed with sutures. The bone flap is put back in its original position and secured to the skull with titanium plates and screws. A drain may be placed under the skin for a couple of days to remove blood or fluid from the area. The muscles and skin are sutured back together. A soft adhesive dressing is placed over the incision.

1.3 Registration in an NN system

The purpose of registration is to align a pre-operative brain image volume loaded in the NN system to intra-operative 2D/3D images. 0.5 Tesla MR is generally used as an intra-operative modality which involves a huge cost limiting its use to the superspecialty hospitals only. US is a low-cost alternative to the intra-operative brain MR (iBMR) imaging. The US imaging is efficient in analyzing the soft tissues of the brain but the scope is limited because US waves can not pass through the cranial bones. However, it is possible to use US as a pre-operative modality after the craniotomy. Images captured just before the surgery provide the latest inside picture that helps the neuro-surgeon to identify the state of the lesion.

The previous section provided an idea how brain surgery is performed. There are a lot of

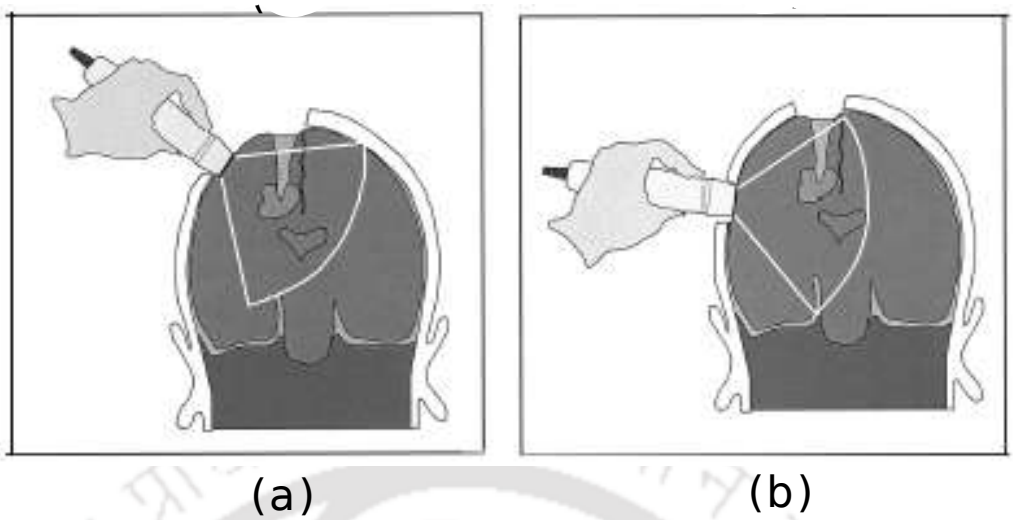


Fig. 1.3: (a) Enlarged craniotomy (b) Craniotomy connects the lesion and the mini-craniotomy is used to monitor the lesion with a US probe (the images are reproduced from [191]

challenges in successful brain surgery. The registration procedure consists of the four following steps as illustrated in Fig. 1.4 and 1.5.

1 **Fiducial-based pBMR/pBCT imaging:** Before capturing pBMR/pBCT images, some fiducial markers are placed on different locations of the head as shown in Fig. 1.4.(a). The marker positions get imprinted on the pBMR images as shown in Figure 1.4.(c).

2 **Registration between the patient coordinate system and the camera coordinate system:** There are two types of reflectors in a surgical setup: the static reflector (SR1) and the dynamic reflector (DR1) as shown in Fig.1.5. SR1 works as the reference to the infrared camera tracking system. The stylus has one or more DR1s. An infrared camera tracker as shown in Fig. 1.5 keeps track of the movement of reflector DR1. If the distances of the fiducial markers from SR1 are known, a linear transformation matrix ${}^{Camera}\mathbf{T}_{patient}$ can be obtained that registers these two coordinate systems.

3 **Registration between the camera coordinate system and the pBMR coordinate system:** After obtaining ${}^{Camera}\mathbf{T}_{patient}$, the stylus is set to the positions of the fiducial markers placed on the patient's head. When the stylus is put on a particular marker on patient's head, the corresponding DR1 can be tracked by the infrared camera. It is matched with the corresponding marker on the MR image. Combining many such markers data and

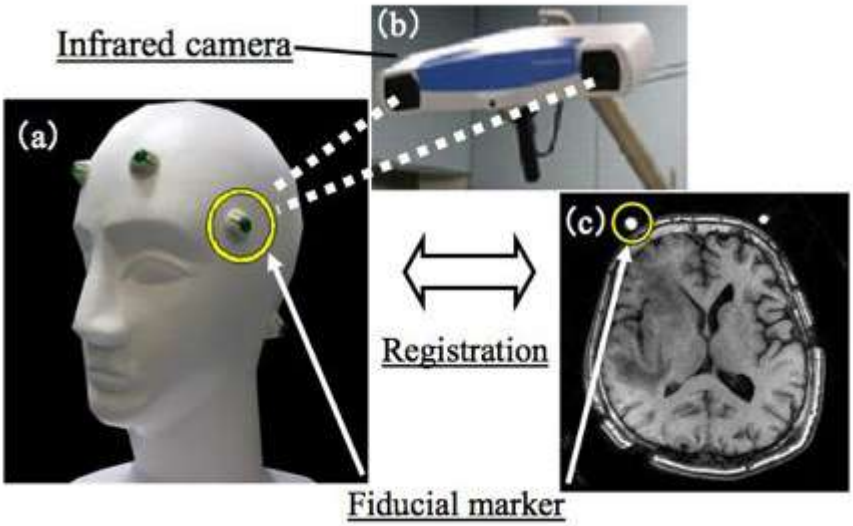


Fig. 1.4: Fiducial marker-based registration of patient and MR coordinate systems.

their corresponding DR1 positions, the transformation model ${}^{MR}\mathbf{T}_{Camera}$ can be obtained and the pBMR coordinate system gets registered with the camera coordinate system.

4 **Registration of the patient coordinate system and the rendered pre-operative imaging system:** The desired registration model ${}^{MR}\mathbf{T}_{patient}$ can be obtained by combining ${}^{Camera}\mathbf{T}_{patient}$ and ${}^{MR}\mathbf{T}_{Camera}$ as

$$ {}^{MR}\mathbf{T}^{initial}_{patient} = {}^{Camera}\mathbf{T}_{patient}\ {}^{MR}\mathbf{T}_{Camera} \tag{1.1} $$

5 **Brain-shift correction:** Both the registration matrices ${}^{Camera}\mathbf{T}_{patient}$ and ${}^{MR}\mathbf{T}_{Camera}$ depend on the positions of fiducial markers, and the position of SR1. They do not change when brain-shift occurs. No image data are required for obtaining these matrices. Brain shift changes the inside tissue locations of the brain. At this stage, the positions of DR1 will not be able to map the diseased location on the pBMR image.

Assuming the brain-shift as a linear deformation $\mathbf{T}_{BS}$, Equation (1.1) gets modified to

$$ {}^{MR}\mathbf{T}^{0}_{patient} = \mathbf{T}_{BS}\ {}^{Camera}\mathbf{T}_{patient}\ {}^{MR}\mathbf{T}_{Camera} \tag{1.2} $$

where ${}^{MR}\mathbf{T}^{0}_{patient}$ is the overall deformation model obtained after brain-shift. We need to take the help of an intra-operative modality to correct the effect of $\mathbf{T}_{BS}$ in (1.2). Figure 1.5

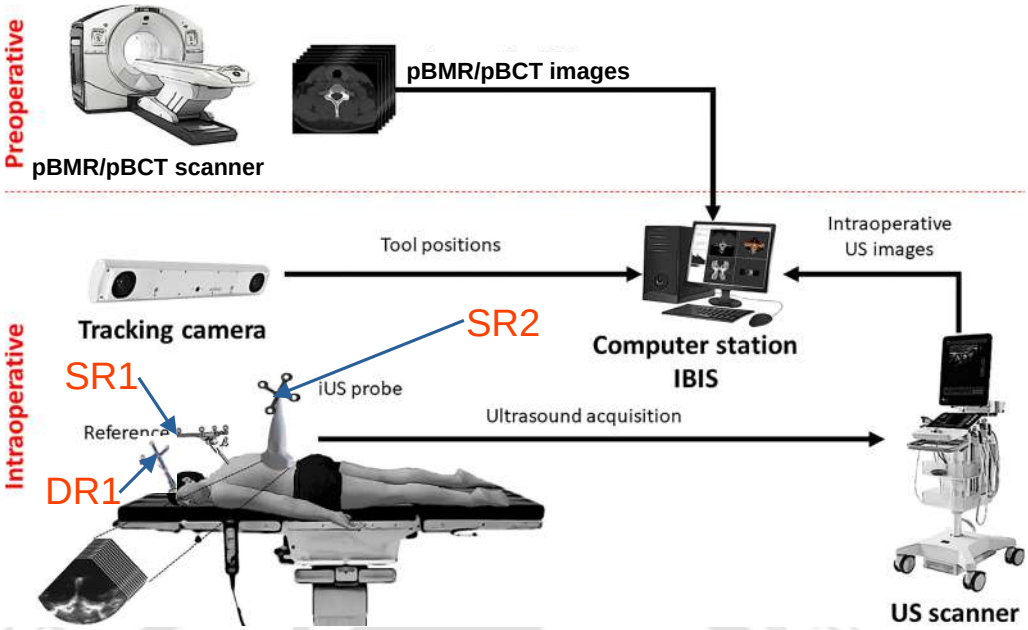


Fig. 1.5: Various components in image guided brain surgery. The figure is reproduced from [72].

shows an iBUS imaging setup that uses a US probe with a static reflector marked by SR2. First, the camera coordinate is registered with SR2. Let this transformation be denoted as ${}^{camera}\mathbf{T}_{US}$. The US probe captures an image before surgery, which is referred to as a pBUS image. If any brain-shift occurs during surgery, the alignment of the iBUS image is changed. This change is sensed by the US probe. Let the pBUS image be the reference V_R, and V_S be the corresponding iBUS image. Suppose a deformation model T is applied applied for correcting the effect of the brain-shift and registering the pBUS to the iBUS images. Then the final registration can be expressed as

$$^{MR}\mathbf{T}_{patient} = \mathbf{T}\mathbf{T}_{BS}\ {}^{Camera}\mathbf{T}_{patient}\ {}^{MR}\mathbf{T}_{Camera} \tag{1.3}$$

This thesis addresses various issues in obtaining $\mathbf{T}$ for performing registration of $\mathbf{V}_S$ and V_R.

1.4 Challenges associated with US-guided brain surgery

As discussed above, the registration process requires many calibrations to remain synchronized with the actual happenings inside the brain. However, we are interested only in performing the registration of pBUS and iBUS images. There are a few challenges in this registration from

the image processing perspective. Some of these challenges are listed below:

1 The resolution of an US image is poor and the image clarity required for performing a brain surgery is not present in a BUS image.

2 All anatomical regions are not visible in a BUS image. Some regions are visible in the pBUS image, but may not be present in iBUS image or vice versa.

3 US can not propagates through air; hence, the resection cavity is filled with saline water for better impedance matching. The iBUS image is formed with the returning echos, which is greatly affected by the changed impedance of the propagating path due to saline water and blood clotting agents present in the resection cavity (RC). The tumor region in pBUS image is hyperechoic. On the other hand, the RC in the iBUS image is anechoic and its surrounding regions are hyper-echoic. This makes the registering image pairs very dissimilar. Registering image pairs having poor similarity is challenging.

4 The intensity correspondence between the pBUS and the corresponding iBUS image greatly deteriorates due to various artifacts. For successful registration, the non-corresponding regions should be eliminated and the similar regions should be used for registration.

5 Registration during brain surgery is needed to be fast. Completing registration within small time is challenging. A simple objective function and a fast optimization algorithm have to be chosen for performing this task.

1.5 Pre-operative and intra-operative brain imaging systems

As discussed that brain surgery is guided and monitored by an intra-operative imaging enabled NN system. For the close monitoring of the diseased region, iBMR is generally preferred. MR imaging works on the phenomenon of magnetic resonance, and has been used in medical diagnostics since 1977. It provides better visualization of the soft tissues than the CT scan images. Improved soft tissue details combined with the need for intra-operative guidance led to the development of the first iBMR in 1991 at the Brigham and Women's hospital [14]. An iBMR image provides near-real-time imaging guidance and results in a better accuracy in an

intra-operative setting. It can be used to check the integrity and the dynamic changes in the lesion and surrounding tissues when microscopic visualization proves to be inadequate. It requires designated operation theatre and the use of MRI-compatible monitoring devices. All the necessary equipment in an MR-equipped operation theatre should not have ferromagnetic materials and the operation theatre staff should be trained and skilled professionals for working in a higher magnetic field. Building an iBMR imaging-equipped operation theatre involves a huge cost and maintenance burden to a small hospital and suburban government hospitals in India.

Apart from iBMR, iBUS imaging is also gaining popularity in neurosurgery [152]. US is a trustworthy modality to examine soft tissues. iBUS imaging is more beneficial than iBMR imaging because of the ease of use, less infrastructural burden, relatively low-cost and the ability to generate real-time images without cumbersome apparatuses. Because of the advancement of signal processing and image processing techniques, the imaging quality of US has improved a lot. Now it is possible to perform neurosurgery with the guidance of iBUS imaging [51]. The initial interest towards an iBUS imaging tool in neurosurgery began in 80s and continued until early 2000. The use of iBUS imaging during a neurosurgical procedure was first described by Reid in 1978 [161]. The main constraints of US is poor image quality which makes image interpretation difficult. After the advent of MR and CT imaging in the market, iBUS imaging has gradually become unpopular among the neurosurgeons. Moreover with the development of high resolution MR and CT, US became less preferable choice for the majority of the neurosurgeons [116, 132, 184]. Recently the application of iBUS images is regaining popularity [9, 57, 73, 131, 175, 155]. With the help of advanced signal processing and image processing techniques, it is now possible to delineate tumor boundaries in MR images at par with MR images. Several applications of iBUS in the spine and brain surgery found in the literature, includes lesion recognition [26, 172], guidance in surgical resection [184], vessels study [66], aspiration of the abscess [109], and posterior fossa decompression [120].

The relevance of both pBUS and iBUS imaging in brain surgery was discussed in [31, 124, 123, 122, 125, 141, 217]. During brain surgery, the pBUS and iBUS images are rendered side by side on the NN system [167, 191]. The pBUS images are captured after opening the brain and before making the RC. These images portray the latest state of the diseased region. Generally, iBUS images do not have good visibility of all the regions, and they get affected by the fluids

and blood clotting agents present in the RC [123].

iBUS imaging has shown superior performance compared to the intra-operative CT (iBCT) imaging for the localization and segmentation of brain tumors [52]. The use of US in brain tumors helps in more radical resections, thus raising the survival period of the patient. Chandler *et al.* [29] reported 21 neurosurgeries with iBUS imaging for localizing cystic regions, tumors, and surrounding edema. Pino *et al.* [152] surveyed the role of iBUS imaging in glioma surgery and found it helpful for monitoring soft tissues.

The major drawback of the US is the poor resolution of images and the non-visibility of various regions [15]. With the advent of various signal processing methods, US imaging has been improved over the decades [51]. However, registering the pBMR and the iBUS images is challenging because of many factors. Generally, the pBMR images are captured in axial, sagittal, or coronal planes of the brain, but pBUS images are captured on an arbitrary plane depending on the location of the craniotomy [163]. In such a case, only 3D volume registration is suitable because 2D slices of MR and US do not lie on the same plane. Another limitation is finding a suitable similarity measure between pBMR and iBUS images because of their different imaging natures.

1.6 Objectives of the thesis

The broad objective of the thesis is to register a pBUS image with the corresponding iBUS image. It is challenging due to various reasons like the poor resolution of US images, non-visibility of all important structures, and the poor similarity between pBUS and iBUS images. The primary objectives of this theses are listed below.

1 Edge is considered to be an important feature of an image. There are many regions in pBUS and iBUS images that have strong edges. The first objective is to extract the regions with strong edges, and to perform rigid registration of this regions using a suitable optimization algorithm.

2 The local similarity between pBUS and iBUS images is poor. Most regions in a pBUS image do not appear similar to the corresponding regions in the iBUS image except for a few hyper-echoic (HE) and anechoic regions. So the next objective is to extract the similar regions in the registering image pairs and use them for registration.

3 The registration of pBUS-iBUS image pair involves solving an optimization problem to find the optimal transformation of the iBUS image. Because of the non-convexity of the optimization problem and the presence of speckle in US data, finding a suitable optimization method is a challenging task. The third objective of the thesis is to adopt a suitable optimization method that performs fast and reliable registration of the speckle filled pBUS-iBUS image pairs.

1.7 Contributions of the thesis

In this thesis, we have developed methods for segmentation and registration of pBUS-iBUS images. The primary purpose of the segmentation methods was to remove dissimilar regions from the pBUS and iBUS images for an easy registration. The major contribution of the thesis can be summarized as follows:

1 **Registration between pre-operative and intra-operative brain ultrasound (BUS) images using patch-based gradient features and heuristic optimization techniques**
The major constraint in registration of the pBUS and the iBUS images is low similarity because of the artifacts present in the iBUS images. The speckle present in BUS images hides major features of the images. We developed a simple patch-based method to find similar edge-rich regions in pBUS and iBUS regions and register them. We have explored the performance of various gradient-based and evolutionary optimization algorithms for minimizing mean squared error (MSE) between the registering image pairs. PSO was found to be more efficient compared to the other optimization methods.

2 **A fast technique for hyper-echoic region separation from BUS images using patch-based thresholding and cubic B-spline-based contour smoothing**
HE regions comparatively appear prominent in pBUS and iBUS images, and segmentation of those regions is an effective approach to the successful registration. We introduced a method for segmentation of the HE regions from the registering image pairs. The patch-wise averaging of intensity makes the process robust to noise. A simple boundary correction step and a cubic B-spline-based contour smoothing step were applied, thus avoiding complex optimization procedures.

3 **A Fast Level-set Curve-flow Formulation by Logarithmic Step-size Control for Brain Ultrasound Image Segmentation**

We have developed a novel patch-based level-set method that is simple, does not require any optimization and can simultaneously segment hyper-echoic, hypo-echoic and anechoic regions. We segmented five regions namely longitudinal fissure (LF), choroid plexus (CP), ventricle, tumor, and RC, and presented comparative segmentation performance with other state-of-the-art methods.

4 **Semiautomatic registration of pBUS and iBUS image based on segmented HE and non-HE (NHE) regions.**

We have discussed an automatic and a semi-automatic segmentation method. Each of these methods is used as preworks to two different registration techniques. Registration using the automatic segmentation shows comparatively better performance than the edge-based method. Particle swarm optimization performed better than the Levenberg Marqardt algorithm(LMA) method and other evolutionary optimization algorithms. Comparative results showed that the proposed method performs better than other competing methods. The said method is applied to the selected pairs of pBUS and corresponding iBUS images. However, this method suffers from low accuracy in many image pairs, particularly where the similarity is low or difficult to capture. A semi-automatic registration method is developed that segments both HE and anechoic regions marked by users from both the pBUS and iBUS images. The HE and anechoic regions are marked by the user, are segmented by a patch-based level-set method. The automatic registration is performed in two stages. The first stage brings the brings the registration process close to the global solution based on the coordinates of the corresponding rectangles marked by the user. The second stage applies gradient-based limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm (LBFGS) optimization for matching the similar regions. The proposed method is faster because of the gradient-based optimization method.

1.8 Organization of the thesis

The rest of the thesis is organized as follows:

Chapter2: Brain US image registration: the state-of-the-art methodologies. This chapter presents the mathematical formulation of a generalized registration problem. We discuss three components of brain image registration, namely a) similarity measure, b) a deformation model, and c) an optimization method. The mathematical formulation of various similarity measures such as mean squared error (MSE), normalized mutual information (NMI), and correlation coefficient (CC) is provided. We have discussed various classical and heuristic optimization algorithms which are effective in medical image registration. The next part of the chapter reviews the literature on ultrasound image-guided brain surgery and the registration of pBMR-iBUS and pBUS-iBUS images.

Chapter 3: Registration of pBUS and iBUS images using edge-rich regions and heuristic optimization techniques. This chapter discusses the factors for the degradation of similarity between the pBUS and the iBUS images and proposes a method for registration of pBUS and iBUS images. The method extracts sharp-edge regions from the registering image pairs using patch-based gradient thresholding and register the image pair using these regions. We also compare the performance of the proposed method with the other state-of-the-art methods on US image registration.

Chapter 4: Registration of pBUS and iBUS images using patch-based segmentation of hyper-echoic regions. Chapter 4 consists of two parts. The first part presents a segmentation method for extracting HE regions from the registering pBUS and iBUS images. Energy minimization-based level-set segmentation methods like active contour(AC), geodesic active contour (GAC) [24], active contour without edge (ACWE) [28], distance regularized level-set evolution (DRLSE) [98], weighted level set evolution (WLSE) [90], and machine learning based method by Nitsch *et al.* [141] are compared with the proposed method.

These segmented HE regions are used as the landmarks for the registration of pBUS-iBUS image pairs. Thus, a segmentation-based registration approach to align the segmented HE regions, was proposed in the second part of this chapter. Performance comparison for various optimization methods and with other registration methods are also provided.

Chapter 5: Semi-automatic registration of pBUS and iBUS images based on fast level-set segmentation of HE and NHE regions. This chapter first explores about the

nature of echogenicity of different anatomical regions in a BUS image. A fast, patch-based, semi-automatic level-set segmentation method for segmenting HE and NHE regions is presented. This segmentation is region specific and five different regions, such as LF, CP, tumor, ventricle, and RC were segmented by this method. The performance of this method is compared with other state-of-the-art methods.

The second part of Chapter 5 proposes about a semi-automatic registration method based on the segmented regions. The proposed method registers images in two consecutive steps and performance improvement in each step is shown.

Chapter 6: Conclusions and future directions This thesis is concluded in Chapter 6 with a summary of the research work carried out in this thesis and an outline of possible research directions out of this thesis.

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

2

Brain US image registration: the state-of-the-art methodologies.

The registration of pBUS and iBUS images is necessary for successful brain surgery. There are three components in image registration: a deformation model, a suitable similarity/dissimilarity measure between registering images, and an optimization algorithm which is fast and robust to noise. This chapter provides a description of these components. A detailed literature survey is carried out highlighting the research issues.

2.1 Introduction

Image registration is a method of aligning two or more images of the same scene taken either at different times, from different viewpoints, or by different modalities. It is a spatial transformation technique; therefore, the intensity profile of the images does not change during registration. The purpose of image registration is to integrate the information from two different images of the same scene. It superimposes the pixels of a target image (V_T) on the sensed image (V_S) by aligning them into a common coordinate system. Registration involves a spatial transformation T that deforms the sensed image V_S to align it with the target image V_T.

Suppose $\Omega \subset \mathbb{R}^m$ be the spatial domain of the registering image pairs. Usually, m is 2 or 3 for medical images. Let each pixel have a coordinate $\mathbf{x} = (x_1, x_2, , x_m)^T$. Let V_S be deformed with an invertible deformation T on its coordinate vector. If T is a linear spatial transformation then

$$T(\mathbf{x}) = \begin{bmatrix} t_{11} & t_{12} & , \dots & t_{1m} \\ t_{21} & t_{22} & , \dots & t_{2m} \\ . & . & , \dots & . \\ . & . & , \dots & . \\ . & . & , \dots & . \\ t_{m1} & t_{m2} & , \dots & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_m \end{bmatrix} \tag{2.1}$$

where t_{ij} is the $(i,j)^{th}$ element of matrix T. Linear transformation is global in nature, where all pixels of an image undergo the same transformation T. For a non-linear transformation, the coordinates of each pixel is estimated using non-linear functions, which are local in nature. Therefore, the transformation of $\mathbf{x}$ may have different shifts for different pixels. Suppose $T(\mathbf{x}, \mathbf{u}) : \mathbb{R}^m \to \mathbb{R}^m$ is a nonlinear shift of $\mathbf{x}$ given by [176]

$$T(\mathbf{x}, \mathbf{u}) = [(x_1 - u_1(x), x_2 - u_2(x), ..., x_m - u_m(x)]^t \tag{2.2}$$

Image registration involves finding an optimal transformation such that $V_S(T(\mathbf{x}, \mathbf{u}))$ is as similar as possible to $V_T(\mathbf{x})$.

The displacement $\mathbf{u}$ can be obtained by solving an optimization problem such as [166]

$$\min_{\mathbf{u}} J(u) = D(V_T, V_S; \mathbf{u}) + \alpha R(\mathbf{u}) \tag{2.3}$$

where J is the objective function which combines a distance measure (D), and a regularization function $R(\mathbf{u})$. α is a regularization constant. D measures the dissimilarity between transforming images $V_S(T(\mathbf{x}, \mathbf{u}))$ and V_T. α. $R(u)$ is used to tackle the illposedness of the image registration problem.

There are various challenges in obtaining the spatial transformation T in medical image registration. V_S and V_T may also be of different modalities. Obtaining T in multi-modal image registration is more difficult because of the different intensity profiles.

The registration of iBUS and pBUS images should be fast enough. This is difficult to achieve because of the large size of data. For reducing the complexity, feature-based registration appears to be preferred in medical image registration.

2.2 Components of BUS image registration

The main components of BUS image registration are: a) a deformation model, b) objective function, and c)an optimization method. The deformation model is characterized by a transformation $\mathbf{T}$ may be either rigid or non-rigid depending on the nature of deformation in iBUS image during surgery. The objective function measures the similarity/dissimilarity between the target and the transformation of the sensed image. An optimization algorithm finds the optimum deformation model that ensures the optimum of the objective function.

Comparing two large images pixel-by-pixel is computationally expensive. Sometime, the objective function is measured on matching features of the registering image pairs. Features are the sparse representation of an image, which are extracted from the image pairs using some rule-based feature-extraction algorithms. A deformation model is applied to the features extracted from the sensed image and the best-possible match with the features of the target image is obtained through an optimization method. Scale invariant feature transform (SIFT) [215], speeded up robust features (SURF) [25], and Harris corner detector [34] are among the preferable feature extraction schemes for medical image registration.

A brief description of the components of image registration is presented below:

2.3 Deformation model

The spatial transformation **T** of the sensed image to the target image is a deformation model. It can be either rigid or non-rigid. A rigid transform includes relative translation, rotation, scaling, and affine transformation. For a rigid planner model, a pixel (x_1, y_1) from image V_S is transformed as the pixel (x_2, y_2) in V_T. The transformation is given by [185]

$$\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \tag{2.4}$$

where a, b, c, d, e and f are the transformation parameters. The nature of the rigid transformation is global, and it does not change the shape of the object. On the other hand, non-rigid deformation can account for differential local deformations, and can be characterized by a deformation vector field. Generally, the amount of non-rigid deformation of the brain is small. In comparison to that the rotation and translation of human brain is significant, because the brain floats on CSF. We model the brain-shift of iBUS images as a rigid transformation.

2.4 Objective function

The objective function includes a similarity/dissimilarity measure and a regularization term as shown in (2.3). There are various similarity metrics available in this context, such as mutual information (MI) [198, 217], correlation coefficient (CC) [168], correlation ratio (CR) [169]. The sum of squared differences (SSD) is also a popular measure used in image registration [63, 160]. The selection of a measure also depends on modalities chosen for the sensed and the target images. The intensity profiles of registering images depend on the inline physics of the modalities. Thus, modeling the relationship between them requires knowledge of both the underlying anatomy and the image formation processes. Though the registration of the pBUS and the iBUS images is not multi-modal, the anatomical regions in the pBUS and iBUS images differ due to various artifacts. For such limitations, all similarity measures are not suitable

for BUS image registration. A detailed study of various similarity measures used in US image registration are provided in [168], [164]. A brief description and mathematical formulation of these measures are given below.

2.4.1 Similarity/dissimilarity measure

Suppose $\mathbf{V}_S$ and $\mathbf{V}_T$ are to be registered by relating their common anatomical regions. Assume that a transformation $\mathbf{T}(., \mathbf{u})$ is applied on V_S. Defining $\mathbf{A} = \mathbf{V}_T(\mathbf{x})$ and $\mathbf{B} = \mathbf{V}_S(T(\mathbf{x},\mathbf{u}))$, the formulations of various similarity/dissimilarity measures are described below.

2.4.1.1 Mean squared error

Mean squared error(MSE) is the most popular dissimilarity measure. It is defined as

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (A(i,j) - B(i,j))^2 \tag{2.5}$$

where each image is of size $M \times N$. This measure is simple and popular to be used in mono-modal image registration [13].

2.4.1.2 Mutual information (MI)

MI is a common similarity measure. It finds a statistical relationship between image intensities. It is used for both mono-modal and multi-modal image registrations [111, 154].

Let $p(i,j)$ be the joint probability distribution of $\mathbf{A}$ and $\mathbf{B}$, and $p(i)$ and $p(j)$ be the corresponding marginal distributions. The mutual information $I(\mathbf{A}, \mathbf{B})$ can be described as

$$I(\mathbf{A}, \mathbf{B}) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} p(i,j) log \left(\frac{p(i,j)}{p(i)p(j)} \right) \tag{2.6}$$

MI does not make any assumption about the nature of the relation between the registering images. It treats the intensity in a pure qualitative way, without considering any assumption of proximity in the intensity space. Secondly, estimating the $p(i,j)$ values is computationally expensive.

Klein *et al.* addressed the concept of local MI (LMI) [91]. They applied computation at different spatial locations and averaged the results to obtain LMI. Here the spatial neighborhood needs to be large enough to contain enough samples for reliable MI estimation. With the increase of neighborhood size, the computation time of LMI linearly increases, that slows down the registration process. This similarity measure is widely used for mono-modal and multi-modal image registration [198, 76, 100]. Huang *et al.* [81] applied the maximization of MI for registering cardiac US-MR images. Shekhar *et al.* [180] applied a 3D volume registration combining motion information in prestress and poststress cardiac ultrasound images with MI as the similarity measure. Other applications of MI can be found in [213, 92, 126, 96, 145]

2.4.1.3 Correlation coefficient (CC)

CC measures the similarity between a pair of images and is the ratio of covariance of the image pairs to the product of their individual standard deviations. Thus, the CC denoted by $\rho(\mathbf{A}, \mathbf{B})$ can be written as

$$\rho(\mathbf{A}, \mathbf{B}) = \frac{Cov(\mathbf{A}, \mathbf{B})}{\sqrt{Var(\mathbf{A})}\sqrt{Var(\mathbf{B})}} \tag{2.7}$$

A higher value of ρ indicates higher similarity between A and B. It is generally useful for matching images of the same modalities [168]. Covariance assumes a linear relationship between the intensities of the corresponding structures in the two images. Thus, the greater the CC, the better is the registered image. There are wide applications of CC in monomodal and multimodal medical image registration [7, 119, 5]. Rivaz *et al.* [166] applied CC for measuring the similarity of pBUS and iBUS image volumes. Zhou and Rivaz [217] registered pBUS and iBUS images using the CC value of the corresponding regions of the two images. CC is simple to measure, and therefore can be applied for real time applications.

2.4.1.4 Correlation ratio (CR)

Assuming a more restrictive functional relationship that maps the intensity in one image to the other, Roche *et al.* [169] proposed CR for multimodal image registration. Let $E(\mathbf{B}|\mathbf{A})$ be the conditional expectation of $\mathbf{B}$ given $\mathbf{A}$. Then the CR defined by $\eta(B|A)$ is defined as

$$\eta\left(\mathbf{B}|\mathbf{A}\right) = 1 - \frac{Var\left(\mathbf{B} - E(\mathbf{B}|\mathbf{A})\right)}{Var(\mathbf{B})} \tag{2.8}$$

where $var(.)$ is the variance operator. Unlike MI, it captures the functional dependence between **A** and **B**. Its value varies between 0 and 1. The value of η close to 0 implies no functional relationship between **A** and **B** while $\eta = 1$ implies a perfect functional relationship between them. Roche *et al.* applied this measure for many multimodal image registration works [150, 167, 170, 119]. Hermosillo *et al.* [78] proposed a framework for local similarity measures, by combining CR and MI. This combination enhances the local correspondences between the registering image pairs. Zvitia *et al.* [220] considered CR as the similarity measure to register sets of fibers extracted from the white matter in brain MR images.

2.4.1.5 Robust patch based correlation ratio (RaPTOR)

In many registration problems, the statistical natures of the registering image pair are different, and therefore, the same region appears to be different in two registering images. Rivaz *et al.* [164] proposed a measure named as *Robust patch based correlation ratio* (RaPTOR). It estimates local CR to achieve robustness against the large spatial intensity inhomogeneity. For reducing computational complexity, two steps are considered. First, the dynamic range of intensities is divided into a few bins and $E(B|A)$s are estimated based on those bin values. For reducing the complexity, CR is calculated on randomly selected rectangular patches and their average is considered as RaPTOR. Suppose, N_p rectangular patches $\Omega_i, 1 = 1, 2, ..., N_p$, are chosen randomly from the registering images. If $\eta(\mathbf{B}|\mathbf{A} : \Omega_i))$ is the CR for Ω_i, RaPTOR can be expressed as

$$RaPTOR(\mathbf{A}, \mathbf{B}) = \frac{1}{N_p} \sum_{i=1}^{N_p} (1 - \eta\,(\mathbf{B}|\mathbf{A}{:}\Omega_i)) \tag{2.9}$$

It is fast because it works on small size patches and reduces the overall computational complexity. This measure has been used successfully in US-MR registration in [164].

2.4.2 Regularization term

To tackle the illposedness of the image registration problem, a regularization term is added to the objective function. Among various regularization terms, the smoothness parameter of the deformation field [164, 165, 186] is widely used. This regularization controls the spurious growth of the deformation parameters. For preserving smoothness of the deformation field, bending energy of the deformation field T was used as a regularizer by Machado *et al.*[107].

Zimmer *et al.* [219] used the average difference of the Laplacian of the the target and sensed images as the regularizer.

2.5 Optimization methods

Optimization is a crucial step in brain US image registration. It is usually iterative and computationally expensive. The time required for completing all the steps involved in a NN system, should be fast. Most of the similarity measures are highly affected by noise. Hence an optimization procedure should be fast and noise-resistant. There are two broad approaches to optimization: classical and heuristic. Below is a brief summary of different types optimizations applied in BUS image registration.

2.5.1 Classical optimization

Optimization methods in image registration essentially maximize or minimize an objective function that depends on the two images and the parameters of deformation model. There is a large number of derivative-based optimization methods, which can be used for image registration [47]. Some of them are (1) gradient descent method (GD) [38], (2) stochastic gradient descent method (SGD) [164, 91], (3) Newton's method, (4) conjugate gradient method (CG) [199, 181], (5) Levenberg-Marquardt algorithm (LMA), and (6) Broyden-Fletcher-Goldfarb-Shanno (BFGS) [102, 6] method. Image registration is formulated as a minimization of a dissimilarity measure between the pair of images. The optimization algorithm sets iterative search directions for obtaining the minimum. The direction of search usually applies the first-order and/or second-order specific information about the objective function. These methods do not guarantee to reach the global minima. Generally, all the derivative-based approaches require an initial guess of the solution. When multiple minima are present, only the nearest minimum can be obtained by these methods. This is a drawback of the derivative-based approach.

The optimization of the objective function begins with Taylor's series. First, we write the multi-dimensional error hyper-surface with the Taylor series expansion. The second-order approximation of the Taylor's series is written as,

$$J(\mathbf{u}+\mathbf{h}) = J(\mathbf{u}) + \mathbf{D}^{\mathbf{T}}\mathbf{h} + \mathbf{h}^{\mathbf{T}}\mathbf{H}\mathbf{h} \tag{2.10}$$

where the $\mathbf{D} = \left[\frac{\partial J}{\partial u_1} \frac{\partial J}{\partial u_2} ... \frac{\partial J}{\partial u_m}\right]^T$ is the gradient vector, and

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 J}{\partial u_1} & \frac{\partial^2 J}{\partial u_1 \partial u_2} \cdots & \frac{\partial^2 J}{\partial u_1 \partial u_m} \\ \frac{\partial^2 J}{\partial u_2 \partial u_1} & \frac{\partial^2 J}{\partial^2 u_2} \cdots & \frac{\partial^2 J}{\partial u_2 \partial u_m} \\ . & . & . \\ . & . & . \\ \frac{\partial^2 J}{\partial u_m \partial u_1} & \frac{\partial^2 J}{\partial u_m \partial u_2} \cdots & \frac{\partial^2 J}{\partial^2 u_m} \end{bmatrix}$$

is the Hessian matrix. $\mathbf{D}$ and $\mathbf{H}$ play important roles in derivative-based optimization. The following algorithms are considered.

Gradient descent: The search direction in GD is determined by the gradient of the objective function. Let $\mathbf{u}_n$ be the solution at n^{th} iteration and $\mathbf{D}_n$ be the gradient vector at $\mathbf{u}_n$. The gradient descent solution at $(n+1)^{th}$ iteration is as follows,

$$\mathbf{u}_{n+1} = \mathbf{u}_n - \zeta \mathbf{D}_n \tag{2.11}$$

where ζ is step length parameter. During each iteration, the objective function moves in a direction that reduces the difference by moving the solution variable towards its negative gradient. Huang *et al.*[81] applied GD for maximizing mutual information for registering real-time 3D-US volumes with the 3D-MR volumes of the heart for the assistance of cardiac-surgery monitoring. The rate of convergence depends on ζ chosen in GD. An increase in ζ increases the speed of convergence, but there is also a chance of oscillation from the optimum solution. Therefore choosing the correct ζ is difficult in this method.

Newton's method: GD considers J as a piece-wise linear function and hence introduces error and often results in slow convergence. Newton's method is a second-order method that is relatively faster. This method finds the iterative solution in the following manner

$$\mathbf{u}_{n+1} = \mathbf{u}_n - \zeta \mathbf{H}_n^{-1} \mathbf{D}_n \tag{2.12}$$

Newton's method has two drawbacks. The estimation and inversion of $\mathbf{H}_n$ is computationally expensive. Secondly, for a minimization problem, $\mathbf{H}_n$ should be positive definite, which can not

be guaranteed during the iterative steps.

Conjugate gradient (CG) method: To overcome the drawback of the GD and the Newton's method, the CG method was developed. In this method, first search direction is $-\mathbf{D}_1$. At each iteration, the search direction is the component of $\mathbf{D}_n$ which is conjugate to the earlier search directions. The conjugacy of the directions S_i and $\mathbf{S}_j$ is given by $\mathbf{S}_i^T\mathbf{H}\mathbf{S}_j = 0$.

Gauss-Newton method: Gauss-Newton's method approximates $\mathbf{H}$=$\mathbf{D}\mathbf{D}^T$, which avoids the computation of second-order elements of $\mathbf{H}$. CG reduces the computational burden involved in the inversion of $\mathbf{H}$, but it cannot solve the problem of the singularity or the negative definiteness of $\mathbf{H}$. On the other hand, Gauss-Newton's method does not require the computation of second-order directional derivatives but requires the inversion of approximated $\mathbf{H}$.

Levenberg Marquardt Algorithm: LMA provides a solution to both of these problems by changing the diagonal elements of $\mathbf{H}_n$. The iterative solution can be found by the following equation.

$$\mathbf{u}_{n+1} = \mathbf{u}_n - \zeta \left(\mathbf{H}_n + \lambda.diag(\mathbf{H}_n)\right)^{-1} \mathbf{D}_n \tag{2.13}$$

where λ is a scalar parameter, known as the scale factor and $diag(\mathbf{H}_n)$ is the diagonal matrix formed by the diagonal elements of $\mathbf{H}_n$. LMA resolves the problem of zero or negative definiteness of $\mathbf{H}$, but the complexity to calculate the Hessian can not be avoided.

BFGS: BFGS is a quasi-Newton (QN) method which resolves the issue of negative definiteness of the Hessian matrix, does not compute second derivative and simplifies the computing of search directions. It also reduces the complexity of calculating the elements of Hessian matrix. This method updates H_n as

$$\mathbf{H}_{n+1} = \mathbf{H}_n + \frac{\mathbf{y}\mathbf{y}^T}{\mathbf{y}^T\mathbf{s}} - \frac{\mathbf{H}_n\mathbf{s}\mathbf{s}^T\mathbf{H}_n}{\mathbf{s}^T\mathbf{H}_n\mathbf{s}} \tag{2.14}$$

where $\mathbf{s} = \mathbf{u}_{n+1} - \mathbf{u}_n$ and y=$\mathbf{D}_{n+1} - \mathbf{D}_n$. The iterative inversion of the $\mathbf{H}$ is computed by the following equation:

$$\mathbf{H}_{n+1}^{-1} = \left(\mathbb{I} - \frac{\mathbf{s}\mathbf{y}^T}{\mathbf{y}^T\mathbf{s}}\right)\mathbf{H}_n^{-1}\left(\mathbb{I} - \frac{\mathbf{y}\mathbf{s}^T}{\mathbf{y}^T\mathbf{s}}\right) + \frac{\mathbf{s}\mathbf{s}^T}{\mathbf{y}^T\mathbf{s}} \tag{2.15}$$

where $\mathbb{I}$ is an identity matrix of the order of $\mathbf{H}_n$. For maintaining positive definiteness of $\mathbf{H}_n$, $\mathbf{y}^T\mathbf{s} > 0$ is needed to be satisfied.

Limited memory BFGS: Limited-memory-BFGS(LBFGS) uses a small amount of computer memory to approximate BFGS. While the original BFGS stores a dense $N \times N$ approximation to the inverse Hessian (N being the number of variables in the problem), LBFGS only stores m no of pairs of $\{s_i, y_i\}_{i=1}^{m}$ that implicitly represent the approximation. LBFGS uses an estimate of the inverse Hessian matrix to guide its search through variable space. When the $(m+1)^{th}$ pair is computed, the oldest pair is discarded. The LBFGS approach is particularly well suited for optimization problems with numerous variables because of its consequent linear memory consumption.

This method finds the descent direction using a different way than ordinary BFGS method. In place of the recursion of the Hessian matrix, the descent direction $(\mathbf{H}_n^{-1}(m)\mathbf{D}_n)$ is obtained from the stored $(m+1)$ sets of data. $\mathbf{H}_n^{-1}$ is updated using m pairs of $\{\mathbf{s}_i, \mathbf{y}_i\}_{i=1}^{m}$ is denoted as $\mathbf{H}_n^{-1}(m)$. This approximation allows the computation of the product $\mathbf{H}_n^{-1}(m)\mathbf{D}_n$. In LBFGS the initial matrix $\mathbf{H}_n^{-1}$ is chosen as the identity matrix scaled by the quantity

$$\gamma_n = \frac{\mathbf{s}_{n-1}^T \mathbf{y}_{n-1}}{\mathbf{y}_{n-1}^T \mathbf{y}_{n-1}} \tag{2.16}$$

γ_n is the scaling parameter which plays an important role in the update of the Hessian and the search direction. It restricts the abrupt change in the values of the elements of the Hessian matrix especially when the curvature is small. Detailed description and global convergence analysis of LBFGS algorithm can be found in [133].

2.6 Heuristic optimizations

The classical optimization techniques often fail to find a solution to the optimization problem with a large number of variables. Either they prematurely converges to a local optimum, or the solution is not obtained within a reasonable time. Alternative approaches that are capable of determining good quality approximations to the exact solutions by the method of searching, are known as heuristic methods. A meta-heuristic (MH) is a non-deterministic

strategy to find solution of advanced and complex optimization problems using a heuristic search process. It explores a wide search space using a heuristic methods in order to find an optimal solution. MHs are among the most successful techniques in various complex and computationally hard combinatorial and numerical optimization problems.They can be seen as a general algorithmic frameworks that require relatively few modifications to be adapted to tackle a specific problem. There are various approaches to image registration, that do not use any derivative-based optimization. These methods use simple heuristic approaches and can be implemented using parallel processing architectures. Some of the popular MHs are simulated annealing (SA)[113] [54], genetic algorithm (GA)[76] [194], particle swarm optimization (PSO) [216] [100] [196], differential evolution [205] [187], gray-wolf optimization [49] [127], coral reef optimization [11] [10], and biogeography based optimization (BBO) [36]. Some of them are applied in medical image registration and briefly presented below.

2.6.0.1 Genetic Algorithm (GA)

GA is a meta-heuristic algorithm that mimics the processes observed in natural evolution [47, 46]. This algorithm consists of 4 steps namely: i) fitness evaluation, ii) selection, iii) crossover, and iv) mutation. All these four operations are inherently stochastic. Both crossover and mutation are applied to a variable with certain probabilities. The algorithm starts with a set of random solutions $\mathbf{u}$, known as population. Each element in this set is named as a chromosome. The value of the objective function $J(\mathbf{u})$ for each solution is named as their fitness. The population evolves in each iteration by two steps referred as crossover and mutation. The crossover is a binary operation, where the genetic materials of a pair of chromosomes get exchanged for generating new offspring. It is a binary operation between two solutions where a few bits are mutually exchanged for exploring more points within the search space. For performing crossover, the solutions are first ranked parameters is determined based on fitness values. The fittest individuals in the population are randomly selected for crossover operations. After the crossover operation, chromosomes undergo minor random variation at certain positions, known as mutation. Generally for a solution, some bits are flipped based on their small mutation probabilities. The steps of GA is described in Algorithm- 1. This crossover operation involves two chromosomes and generates two new chromosomes. For example, in Fig.2.1, Chromosome 1 and Chromosome 2 exchange their elements between two positions and

Algorithm 1 Genetic Algorithm

1: Initialization: population size=N, Percentage of crossover= r, mutation probability=p.
2: Define stopping criteria
3: Generate a random population of size N
4: Repeat if stopping criteria does not meet
5: Find fitness of all N individual solutions
6: Order based on fitness and select best N_r chromosomes
7: Chose a pair of solutions randomly and apply crossover.
8: Apply mutation on each variable in a solution with a probability p.
9: Update the new solutions

produce Offspring 1 and Offspring 2 that appear as new solutions.

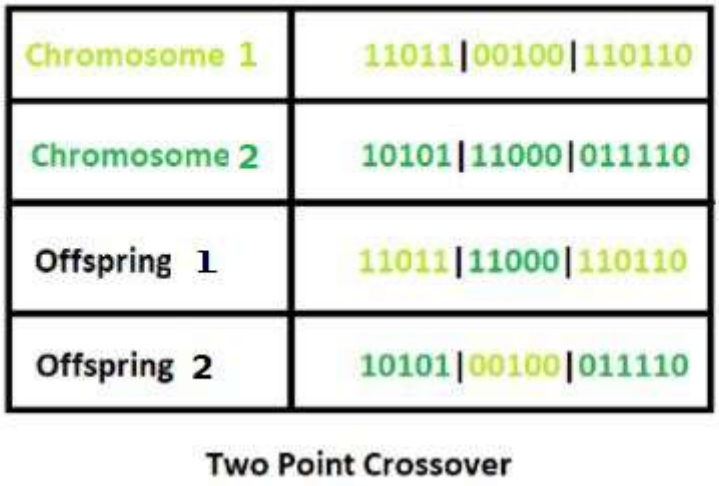


Fig. 2.1: Example of 2 point crossover

GA has been applied in medical ultrasound image registration [53].

2.6.0.2 Simulated Annealing

Simulated annealing mimics controlled the heating and cooling of a material to modify its physical properties. The energy state of a physical system is defined as the objective function $J(\mathbf{u})$. The goal is to drift the system from its arbitrary initial state to a state having the minimum energy. The algorithm starts with an initial solution $\mathbf{u} = \mathbf{u}_0$, and it considers a

few neighboring points $\mathbf{u}'$. If a neighbor lowers the objective function, then it is converted to $\mathbf{u} = \mathbf{u}'$. Otherwise, it is accepted with a probability

$$p(J(\mathbf{u}), J(\mathbf{u}'), \Theta) = exp^{\frac{(J(\mathbf{u})-J(\mathbf{u}')}{\mathbb{K}.\Theta}} \tag{2.17}$$

where $\mathbb{K}$ is a constant and Θ is the temperature which is a time-dependent function. It follows a cooling schedule for Θ over the iterations. Generally Θ is reduced linearly with its value zero at the last iteration. In this way, the search is performed over a broad region of the search space containing good solutions. The details of the algorithm can be found in [37].

2.6.0.3 Particle swarm optimization (PSO)

Kennedy and Ebarhart [89] proposed PSO, a population-based algorithm that mimics the food searching method of a swarm of particles. The rank of particles can be prepared based on the fitness of all particles in the population. In each iteration, it finds the local best $\mathbf{u}^{lbest}$ and the best solution until the present iteration is $\mathbf{u}^{gbest}$. The position of a particle represents a solution vector and the objective function is assumed to be the amount of food. Through the movement of the particles, the algorithm examines new solutions. If better solutions are found, $\mathbf{u}^{lbest}$ and $\mathbf{u}^{gbest}$ are updated.

In algorithm form, the search process is initiated by a fixed number of variables. The position of each particle represents a solution vector of length m over the entire search space $\mathbb{R}^m$. The positional change in every iteration is termed as the velocity of a particle. The algorithm uses a velocity up-gradation followed by a position up-gradation equation.

Let $\mathbf{U}$ be a set of k variables representing k particles and $\mathbf{u}^j \epsilon \mathbb{R}^m$ be its j^{th} element. $\mathbf{U}$ is given by

$$\mathbf{U} = \{\mathbf{u}^1, \ \mathbf{u}^2, \ \mathbf{u}^3, \ ..., \ \mathbf{u}^j, \ ..., \ \mathbf{u}^k\} \tag{2.18}$$

At instant n, $\mathbf{u}^j(n)$ is a function of $\mathbf{u}^j(n-1)$ and velocity of that particle $\mathbf{v}^j(n)$

$$\mathbf{u}^j(n) = f(\mathbf{u}^j(n-1), \mathbf{v}^j(n)) \tag{2.19}$$

$\mathbf{u}^j(n)$ and $\mathbf{v}^j(n)$ can be written as

$$\mathbf{u}^j(n) = \left[\ u_1^j(n), \ \ u_2^j(n), \ \ u_3^j(n), \ \ \ldots \ \ u_m^j(n) \ \right] \tag{2.20}$$

and,

$$\mathbf{v}^j(n) = \left[v_1^j(n), v_2^j(n), v_3^j(n), ..., v_m^j(n)\right] \tag{2.21}$$

The velocity up-gradation equation at n^{th} iteration is expressed as,

$$\mathbf{v}_i^j(n) = \xi v_i^j(n-1) + p_1\alpha_i(n)(u_i^{lbest}(n) - u_i(n-1)) + p_2\beta_i(n)(u_i^{gbest}(n) - u_i^j(n-1)) \tag{2.22}$$

and the position upgradation is given by

$$u_i^j(n) = u_i^j(n-1) + v_i^j(n) \tag{2.23}$$

where ξ is the inertia parameter which is generally kept nearly equal to 1, p_1 and p_2 are constants, and $\alpha_i(n)$ and $\beta_i(n)$ are two random variables each uniformly distributed between 0 and 1, $\mathbf{u}^{lbest}(n)$ is a tracking variable which tracks the best out of k particle positions at n^{th} time step and $\mathbf{u}^{gbest}(n)$ is a tracking variable which tracks the best particle out of all past and present iterations. The local best and the global best at any iteration can be written as.

$$\mathbf{u}^{lbest}(n) = \underset{j\in\{1,2,...,k\}}{argmin}\ \left(J(\mathbf{u}^j(n)\right) \tag{2.24}$$

and

$$\mathbf{u}^{gbest}(n) = \begin{cases} \mathbf{u}^{lbest}(n) & \text{if } J\left(\mathbf{u}^{gbest}(n-1)\right) > J(\mathbf{u}^{lbest}(n)) \\ \mathbf{u}^{gbest}(n-1) & \text{otherwise.} \end{cases} \tag{2.25}$$

Equation (2.22) shows that the velocity of each particle has two types of motivations: $\mathbf{u}^{lbest}$ guides the social motivation and $\mathbf{u}^{gbest}$ shows the cognitive motivation which a particle acquires from the longer past history and remembering the best events in the past.

2.6.0.4 Biogeography based optimization

Biogeography-based optimization (BBO) algorithm uses a stochastic population as candidate solution and iteratively improves it for a given value of the fitness function. BBO belongs to the class of metaheuristics with many variations of the originally proposed method by Simon [183]. This method does not require any prior assumptions about the problem, and therefore can be applied to a wide class of engineering problems. BBO follows the phenomenon of the geographical distribution of biological organisms. Mathematical models of biogeography describe

speciation (the evolution of new species), the migration of species (emigration and immigration) between islands, and the extinction of species. An island that is resourceful and friendly to live in is said to have a high habitat suitability index (HSI). The features that determine HSI are called suitability index variables (SIVs). SIVs are the independent variables whereas HSI is dependent on SIVs.

Immigration and emigration of different species are continuous processes in biogeography. Immigration increases and emigration decreases on an island with the increase in the HSI. This phenomenon reverses when the HSI of an island is low.

Suppose $\mathbf{U} = \{\mathbf{u}^1, \mathbf{u}^2, ..., \mathbf{u}^j, ..., \mathbf{u}^k\}$ is the set of sorted populations with respect to their decreasing HSI. Each variable in $\mathbf{U}$ are considered as a habitat or candidate solution. Each candidate solution consists of m independent variables or features u_i^j [104]. $J(\mathbf{u}^j)$ is the HSI of j^{th} habitat. A habitat u^j faces both emigration and immigration with an emigration rate (μ_j) and immigration rate (λ_j), which can be estimated by the following equations

$$\mu_j = \frac{\mu_0 j}{k} \tag{2.26}$$

$$\lambda_j = \lambda_0 \left(1 - \frac{j}{k}\right) \tag{2.27}$$

with λ_0 and μ_0 as the maximum immigration and emigration rates respectively. They are chosen in such a manner that $\lambda_j + \mu_j$ is a constant less than or equal to 1 [183]. If an independent variable in a habitat is chosen for immigration or emigration based on their respective probabilities, they exchange their values with other habitats. It is not likely that all independent variables migrate from one habitat to the other; it is carefully controlled by changing the values of μ_0 and λ_0.

Catastrophic disasters have the potential to significantly alter a natural habitat's HSI. They can also lead to a species count deviating from its equilibrium value. Therefore, a habitat's HSI may alter quickly as a result of circumstances which force to change some features of a species. This process is known as mutation. In BBO, some of the independent variables in a candidate solution is replaced by a random number based on its mutation probability. The chance of mutation is more in non-elite members, and a few species get mutated based on the mutation

probability. The probability of a species to be mutated is determined by

$$m_j = m_0 \left(1 - \frac{P_j}{P_{max}}\right) \tag{2.28}$$

where m_0 is the maximum mutation probability for the j^{th} habitat, P_{max} is the maximum possible species count and P_j is the species count of the habitat. The species count is analogous to the HSI of a habitat. If the i^{th} position of the j^{th} candidate solution is selected with mutation probability m_j, the element is replaced by a random number.

The algorithm workflow of BBO is described in Algorithm 2

Algorithm 2 Biogeography-based algorithm

1: Initialize BBO parameters: Maximum species count k, number of iterations, maximum mutation rate m_0, maximum emigration rate μ_0, and maximum immigration rate λ_0.
2: Generate initial random population of $k \times m$, where k is the number of habitats/islands and m is the dimension of the variables.
3: Evaluate HSI(cost function) of every habitat.
4: Sort the population from the best to the worst in terms of HSI.
5: Map HSI to the number of species.
6: Calculate μ_j and λ_j as per the Equation (2.26) and (2.27) respectively.
7: Identify the habitat and its variable at i^{th} position, which will be emigrating based on a randomly generated number $rand_i > \mu_j$.
8: Identify the immigrating solution based on a randomly generated number $rand_j > \lambda_j$.
9: Copy the variable from the emigrating variable to the same location i of the immigrating variable.
10: Follow step 7-10 for all habitats
11: Apply mutation to the non-elite members of the population based on the probability of mutation m_j in Equation (2.28).
12: Evaluate all the new habitats
13: If a new habitat is found to be better than the old habitats, replace the worst member of the old habitats.
14: Goto Step 3 until stopping criteria are satisfied.

2.7 State-of-the-art of BUS image registration

The literature on US-guided brain image registration is divided into two categories: a) registration of pBMR and iBUS images; b) registration of pBUS and iBUS images. The state-of-the-art researches in each categories are described in the following sub-sections.

2.7.1 Registration between pBMR and iBUS images

Trobaugh *et al.*[190] were the first to propose employing US imaging as an intra-operative modality in conjunction with pBMR or pBCT scans. A tracking device was used to monitor the location and orientation of the US transducer. The iBUS images captured by that NN sytem are compared with the pBMR or pBCT images. Gronningsaeter *et al.*[71] described the capability of ultrasound in brain surgery along with the results on eight patients. They employed colour flow imaging to localize the blood vessels, arteries, veins in the brain. Commeau *et al.*[43] reported several aspects of the US-assisted brain surgery including brain-shifting, target volume rigidity, patient rigidity, picture distortion, tracking device inaccuracy, and so on. The literature also describes several applications of brain surgery employing a NN system. Based on clinical investigation, Hammoud et al. [73] verified the applicability of US imaging as an intra-operative modality. They demonstrated that iBUS image is not only helpful in localizing and defining the margins of different types of brain tumors, it also accurately determines the extent of resection. Unsgaard et al. [192] reported a comprehensive clinical review on a variety of topics, including probe location, open tumor resection, closed tumor resection, surgical setup, and so on. Publications about the registration approach between the intra-operative ultrasound(iUS) and pre-operative MR (pMR) images can be found after 2000. Maintz and Viergever's survey work [115] on medical image registration, focuses on the registration of MR-CT, MR-MR, CT-MR, and CT-CT images, as well as registration of stereotactic and either MR/CT images.

Pagoulatos *et al.* [148] first attempted the registration of US-MR multimodal images on a phantom object. The object was scanned for generating 3D MR and 2D US images. Magnetic sensors detect the movement of an object inside the phantom and each sensor works as a trans-receiver system. The transmitter remains stationary, and the receiver is mounted on the floating object (e.g., a surgical tool or a US probe) to be tracked. The transmitter space is registered with both MR and US image coordinate systems resulting two rigid transformation matrices. These two matrices are combined to obtain the final transformation matrix for registering the MR image plane to the US image plane. Roche *et al.* [170] proposed first automatic rigid registration of 3D-US and pre-operative MR images. The principal focus of the method is to capture inter-modal image similarity using CR as the similarity measure. The novelty of this similarity measure lies on both the intensity and its gradient to incorporate multivariate information from MR data. They formulated a generalized form of the CR, and automatic

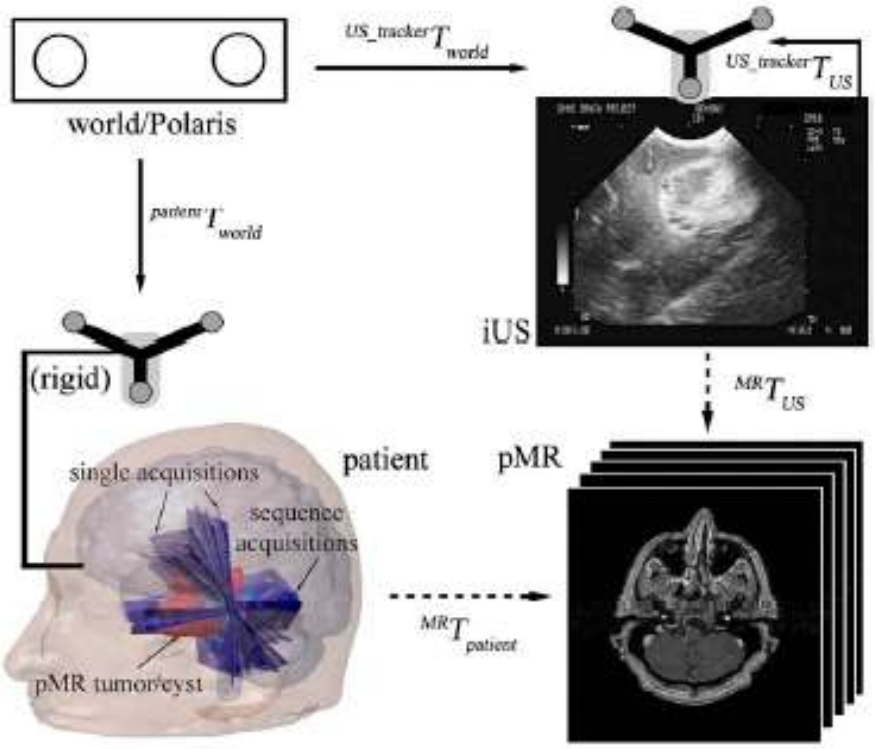


Fig. 2.2: Various coordinate systems used in image registration are depicted in the diagram. Transformations determined from registration are indicated by dashed arrows and are subject to change during the re-registration process. The transformations defined through calibration are indicated by solid arrows, which are fixed. Inverse transformation obtained through matrix inversion is denoted by a reverse arrow. (This image is reproduced from the article of Ji *et a.l*[84]

registration was accomplished by maximizing CR.

Rasmussen *et al.* [159] developed and demonstrated a rigid registration method for automatic brain-shift correction of a pre-operative 3D MR image using an intra-operative 3D US image. It also provided a clinical guideline for performing effective surgical planning using fMRI, the diffusion tensor imaging (DTI), and 3D US imaging.

Ji *et al.* [84] performed a multistage fiducial-marker-based registration of pBMR and iBUS images. The US images are initially transformed into a stationary MR imaging space. The registration scheme is shown in Fig.2.2, which was taken from [84]. It depicts an image-guided neuro-surgery (IGNS) system, that captures 3D-world coordinates using a Polaris tracking camera system. The IGNS system is initially loaded with the pBMR images. Normalized mutual information (NMI) was used as a similarity measure which was maximized using the Powel's optimization [151, 170] method.

Nigris *et al.* [139] employed the LMI, which was proposed by Karacali [85] previously. They

also came up with the adaptive local mutual information (ALMI) method for registering US and MR images. ALMI saturates smoothly and as a result no oscillation happens near the optimum. When optimization approaches the maximum, it allows the energy function to stabilize easily.

Nigris *et al.* [45] proposed multimodal rigid registration of pBMR and iBUS images. They presented a novel similarity metric which considers the angle between the local intensity gradient vectors. Initially, the images are regenerated into multiple scales, ranging from the coarsest to the finest. The gradient orientations for all points in each scale are found and rendered as a gradient image. The gradient image is resampled based on a threshold value of the gradient. The threshold is determined by choosing a proportion of the total pixels in a neighborhood. The transformation matrix is calculated by maximizing the sum of the cosine of inter-gradient angle between two different images.

Farnia *et al.* [59] proposed a registration approach on MR-MR and US-MR modalities. The subject was an artificial brain phantom. Brain lesion and landmarks were created artificially by inflatable catheters. Plastic tubes were used for representing the blood vessels. The phantom volume was scanned with both MR and US scanners. For matching of the landmarks in MR and US images, the coherent point drift (CPD) [13] method was adopted. CPD is a robust probabilistic multidimensional approach to both rigid and nonrigid point set registrations. The alignment of two sets of points is treated as a probability density estimation. One set represents the centroids of a Gaussian mixture model (GMM), and the other represents the data points. This approach maximizes the likelihood function iteratively to fit the GMM centroids. This method outperforms the iterative closest point (ICP) method [13]. The ICP algorithm is a method for match point sets that use a least square error minimization approach. Geometric data, parametric curves and surfaces, implicit curves, and surfaces can all be matched using this method. The complete data collection is first transformed into a point-set placed at the vertices of triangles that represent the surface. The ICP method finds the best match between the points in the target and the sensed images during registration.

Arbel *et al.*[4] introduced the concept of a pseudo image of MR slices, which has similar statistics like an US image. For an easier match, the MR image is converted into a pseudo-US image. A peudo-US image is prepared from an MR image by various signal processing steps to match its statistic similar to a US image. The registration of pre-operative pseudo-US image and intra-operative US image, they used a package named Automatic Non-linear Image Matching

and Anatomical Labeling (ANIMAL) that combines segmentation and a registration algorithm. During segmentation, the entire image volume is divided into approximately 100 anatomical structures. Experts performed this segmentation manually. The intensity gradient of US image was calculated in the radial direction, assuming the US probe-tip as the origin. An intensity remapping operation was performed on the segmented structures and on the boundaries to attain a pseudo-acoustic density simulation that is comparable to the intensity profiles of intra-operative US brain images. Mercier *et al.* [125] proposed an innovative approach to convert the segmented MR image into a pseudo-US image for reducing the registration of US and MR images. The performance of this approach was proven better than the method proposed in [4]. CR between the US and pseudo-US images was used as the similarity measure. Both of these registrations was performed using MINC tool in *minctracc* software. Authors reported that CR performs better than NMI as a similarity metric for US-pseudo US registration.

Rivaz *et al.* [165] suggested a strategy for accomplishing automatic MR-US nonrigid registration. Robust patch-based correlation ratio (RAPTOR) was introduced, which is less sensitive to high levels of spatial intensity inhomogeneity. This study makes a number of contributions: a) measurement of patch-based local similarity, b) developing a strategy that removes non-corresponding regions between the registering images, c) performing direct registration of pBMR and post-operative BUS images, and d) matching anatomical landmarks between pBMR and post-operative BUS images for post-registration validation. Christensen *et al* proposed inverse consistent registration methods [38] which perform transformation on both the target and the sensed image; which are referred to as the forward (T_1) and inverse transformation (T_2) respectively. They eliminate bias by computing T_1 and T_2 and penalizing the difference between $T_1 \circ T_2^{-1}$ and the identity transformation. A symmetric and inverse consistent approach, similar to [77], was used in that study. The iterative approach of Chen *et al.* [35] was applied for inverting the transformation matrix. The full forward and backward deformations may then be calculated as $T_1(0.5) \circ T_2(0.5)^{-1}$ and $T_2(0.5) \circ T_1(0.5)^{-1}$, where 0.5 corresponds to half of the deformation field.

2.8 Registration between pBUS and iBUS images

The literature on pBUS-iBUS image registration is limited to a few methods only. The research in this field got a new height after publication of the BITE database by Mercier and her team at Montreal Neurological Institute in 2012 [123]. It contains pBUS, iBUS, pBMR, and iBMR volumes of 14 patients. This dataset was developed for the pupose of the segmentation of brain tumor and monomodal or multimodal registration of brain images. Later Xiao *et al.* [202] published a new dataset *REtroSpective Evaluation of Cerebral Tumors* (RESECT) in 2017 for pBUS-iBUS images. This dataset also contains 3D pBUS, iBUS, pBMR, and iBMR images of 23 patients. The current researches in this field use volume images from these datasets.

Mercier *et al.* [122] first developed a software tool named *minctracc* for registering pBUS and iBUS image volumes. This software is capable of the registration of the pre- and post-operative ultrasound images. It estimates the best linear spatial transformation required to register two 3D volumes. The flexibility of choosing different similarity measures like CC, MI, NMI, etc. made it popular. Developers of this tool reported that CC was the best similarity measure for performing registration of pBUS and iBUS images.
Another complete work on the registration of pBUS and iBUS images was published by Zhou and Rivaz in 2016 [217]. The contribution of the work was twofold:

a) Removing the non-corresponding regions outlier regions from the registering image pairs. They estimated the gradient of the NCC at all pixels. It was shown that the orientation of the gradients in the non-corresponding regions is random in nature.

b) Adopting a cubic B-spline-based nonrigid registration scheme for the PBUS-IBUS image registration. For reducing the computational complexities, and to capture the local properties, they adopted a patch-based approach. Both images are divided into non-overlapping rectangular patches. For formulating the minimization problem, -NCC^2s of all the patches were summed up to get the value of the objective function. The patches significantly different from one another, were classified as outliers, and were not included in the similarity calculations.

Apart from the above methods, a few recent publications are available. Machado *et al.* [108] highlighted the difficulty in registration because of the poor similarity between the registering pBUS and iBUS image pair. They extracted 3D-SIFT features from the registering pBUS and

iBUS images and matched them using an exhaustive search method.

Canalini *et al.* [20] proposed a segmentation-based approach for registering pBUS and iBUS volume images. They trained a deep learning (DL) network 3D-U-Net [40] using annotated data of major HE structures of the brain. Registration was performed based on the information available in the segmented images. In an another work, Canalini *et al.* [21] highlighted the issue of dissimilarity between the RC region in iBUS image and the tumor region in pBUS image. The tumor region in the pBUS image is absent in the iBUS image. They segmented the RC region using 3D-U-net [40] and excluded it from the measurement of the normalized gradient field between the registering image volumes. LBFGS optimization was applied for minimizing the distance measure.

Recently, Canalini *et al.* [19] improved their earlier work [20] by incorporating a distance function for the similarity measure. The segmentation of the HE regions obtained by the 3D-U-Net can be considered as a 2-class binary map where the inside region is labeled by binary 1 and the outside as 0. The Euclidean distance transform is applied on the inside region of the binary map. LBFGS optimization was applied to minimize the normalized gradient field between the registering images.

Despite these various methodologies and cutting-edge techniques, this discipline still has a long way to go in terms of better similarity or dissimilarity measures, post-processing, optimization, and so on. The current state of this field is focused on developing a suitable similarity and dissimilarity measure, with gradient-based optimizations being used to speed up the iterative steps. There is room to enhance this problem by applying better deformation models and optimization approaches, which could result in a significant improvement over the current state-of-the-art results.

2.9 Conclusions

This chapter discussed the components, namely the deformation model, different similarity metrics, and optimization methods used in the BUS image registration. It described various imaging modalities used for monitoring brain surgery. This chapter provided a detailed literature survey on difficulties related to the monitoring of brain surgery using imaging modalities. The review included the state-of-the-art methods for the registration of pre-operative and

intra-operative brain images including the pBUS and iBUS images.

3

Registration of pBUS and iBUS images using edge-rich regions and heuristic optimization techniques.

Registration of pBUS and iBUS images is essential during US-guided brain surgery. The similarity between the registering image pairs is desired for successful registration. Because of speckle and various artifacts, many regions in the pBUS image do not match well with the corresponding regions in the iBUS image. In this chapter, the similarity between pBUS and iBUS images is improved by using the regions with prominent edges. Patchwise intensity gradients are thresholded to isolate the edge-rich regions of the registering image pairs. Registration is performed using classical and heuristic-based optimization methods, and their results are compared. It is found that the registration of edge-rich regions using particle swarm optimization results in good registration performance.

3.1 Introduction

It is difficult to avoid the deterioration in registration for brain-shift during brain surgery. When brain-shift occurs, the NN system maps the diseased region into a wrong position on the pBUS image stacks [124]. To avoid such misalignments, a suitable registration method is required, which corrects the brain-shift and highlights the surgical device on the pBUS images at the correct positions.

The advantage of US-US registration over the MR-US registration is that the intensity profiles of pBUS and iBUS images are similar. However, using the US in brain surgery suffers from a few limitations such as

(i) **Poor visibility of regions-** The primary limitation of the BUS images is the poor visibility of the organs/regions. Only a few regions like sulcus, falx, choroid plexus, ventricles, haemorrhage, cysts, and tumorous tissues are partially visible in the BUS images[123]. It is also difficult to identify grade-I and grade-II gliomas in the BUS images.

(ii) **Artifacts present in the iBUS image-** The other crucial problem of the iBUS image is the presence of various artefacts that deteriorate the similarity measure between the pBUS and the iBUS image. If the dissimilarity between the corresponding pBUS and iBUS images is significant, the non-corresponding regions are excluded from the registering image pairs [217]. The correspondence between the pBUS and the corresponding iBUS images degrades due to the following reasons:

- **The presence of saline water in the RC:** The RC is filled with saline water for two purposes [123]: a) US does not propagate through the air, and water helps it to propagate, and b) water does not allow the tissues surrounding RC to displace much. But the presence of water has some adverse effects also. Saline water in the RC does not reflect US waves, whereas the tumor tissues in pBUS images appear hyper-echoic. Saline water changes the acoustic impedance of the transmitting and receiving paths of the US wave. Hence, the same region in pBUS and iBUS images appears different. Fig. 3.1.(a1) shows a HE pBUS image with a tumor region. After removal of the tumor, the region is filled with saline water and appear as black in Fig. 3.1.(b1). It increases dissimilarity between the registering image pairs.

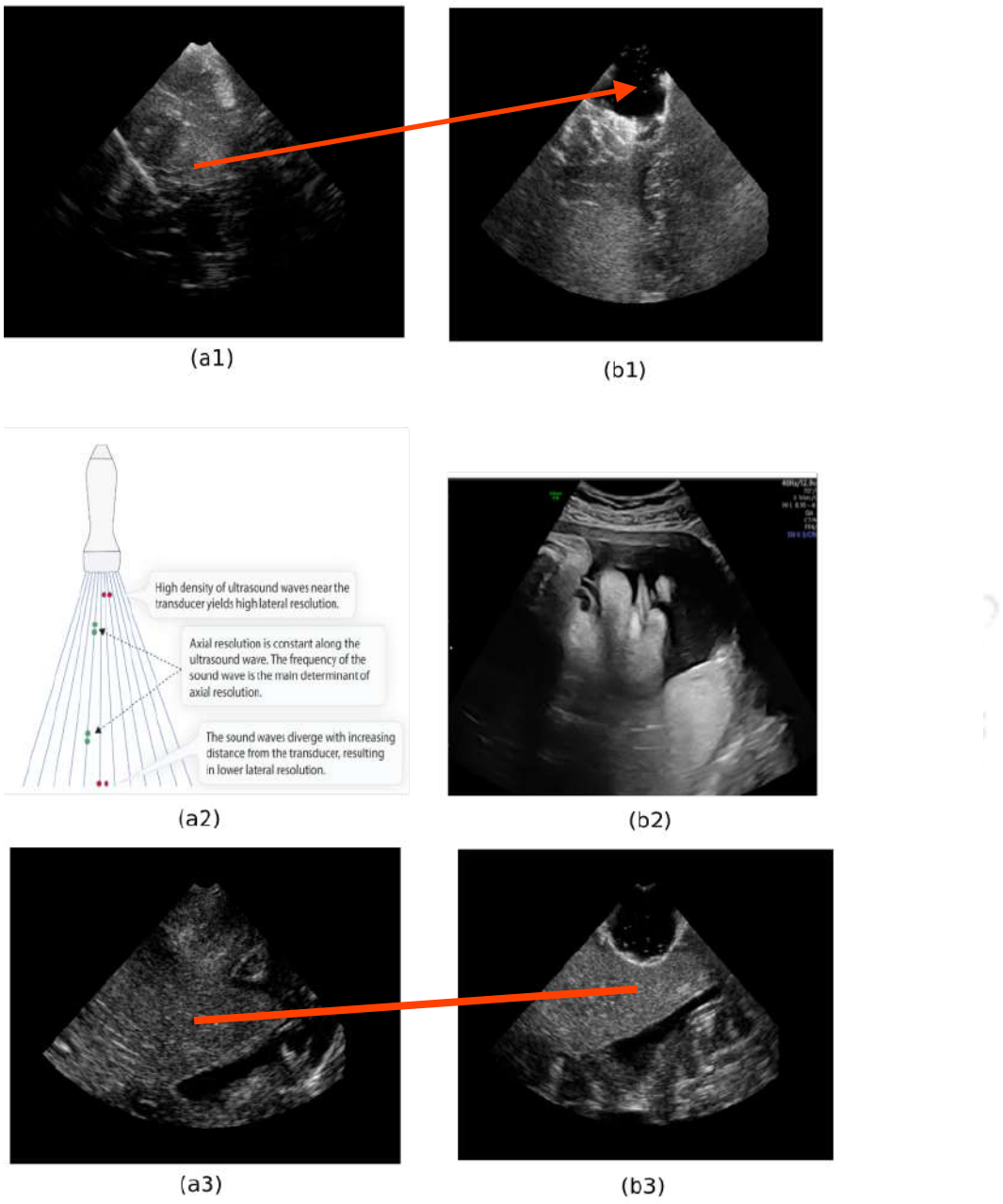


Fig. 3.1: (a1) shows a HE tumor region, which appears anechoic after surgery in the intra-operative image as shown in (b1). (a2) shows the issue of low resolution at a higher scan depth in a US image. (b2) shows higher resolution at the lower depth and low lateral resolution at the higher depth of a US image. (a3) the region below the tumor has lower intensity, whereas in (b3), brightness is more due to the scattering of light from blood clotting agent in the RC

- **Changes in the scan depth:** The scan depth of the US probe is often changed for better observation of the position of the diseased region. The intensity of the reflected wave decreases, and visual information gets reduced with the increased depth. For scanning with higher depth, a low-frequency US sensor is used. Similarly, for shallow screening, a high-frequency US probe is needed. If both the pBUS and iBUS apply different frequencies, their appearance changes; hence, their similarity deteriorates. Fig. 3.1.(a2) shows the mechanism of radiating US waves by a polar probe. US beam propagates as a steered beam along the radial directions. The distance between two such paths increases at higher depth. The fan-shaped image is reconstructed in a rectangular grid using interpolation. It results lower lateral resolution at a higher depth. Fig. 3.1.(b2) shows that the visibility and fine details at a lower depth is better than the objects present in the higher depth.
- **Presence of blood clotting agents in the RC:** For stopping bleeding in the RC, various blood clotting agents are used. This results in major artifacts in the pBUS images. Fig. 3.1.(a3) shows a pBUS image that consists a HE region. The same region is marked by red line, appears more hyper-echoic due to blood clotting agents in the resection cavity.

It was discussed in Section 1.3 that both pBUS and pBMR imaging can be used along with iBUS imaging in an NN system. The main challenge in MR-US registration lies in choosing a suitable similarity measure, because the intensity profiles in these modalities are different. Among various similarity measures, the correlation ratio(CR) [170, 165], the normalized cross-correlation (NCC)[165], the local phase difference and the robust patch-based correlation ratio (RAPTOR)[164] have been used.

For obtaining accurate registration, adopting a suitable deformation model is essential. If the shape deformation of the brain is not significant, it is assumed to be rigid. However, if we consider the local shape changes due to the applied pressure by the surgical instruments, the shrinkage in volume by the leakage of the CSF, etc., the deformation becomes nonrigid. In this work, we assume that the local shape deformation is insignificant; hence, the nature of deformation is a combination of translation, rotation, and affine transformation.

This chapter proposes a new method for the registration of pBUS-iBUS images. First we have extracted the fan shaped part of the BUS image from the rectangular image. For

improving the similarity between the pBUS and iBUS images, the regions with prominent edges are extracted. We argue that the edge-rich regions of an image contains more information, which may be exploited for registration. The organization of the rest of the chapter is as follows. Existing approaches are described in Section 3.1.1. The proposed method is presented in Section 3.2. The results are shown in Section 3.3 and the conclusion is drawn in Section 3.5.

3.1.1 Existing approaches and their limitations

It has already been stated that pBUS-iBUS image registration is easier than iBUS-pBMR image registration because selecting an appropriate similarity metric between US-US images is easier than US-MR images. However, only a few studies on pBUS-iBUS registration have been reported till date [21, 108, 122, 165]. The lack of BUS data is a key limitation in this research. To the best of our knowledge, only two datasets namely BITE [123] and RESECT [202] are available for BUS image segmentation and registration. Mercier *et al.*, a research team from Montreal Neurological Institute published the BITE database [123] in 2012. It consists of images of 14 brain tumor patients of different age groups. Both pre-operative and post-operative MR and US image slices of the same patients are available in this dataset. The primary objective of this dataset was to help this community in brain tumor segmentation and brain image registration. Similarly, Xiao *et al.* published RESECT dataset [202] in 2017. This dataset consists of both pre- and post-operative brain US images of 23 patients. Both these datasets contain 3D pBUS and iBUS images. Each 3D image consists of 200-400 2D image slices, among which we have chosen our experimental images judiciously with the help of radiologist.

The concept of iBUS-based NN system was first described by Mercier *et al.* in [124]. They published another article [122] on 3D US-US volume registration using pBUS and iBUS images of the BITE database by optimizing normalized mutual information (NMI) between the pair of registering images. For the validation of the results, common tag points are provided in the datasets, which are the corresponding points in the pBUS and iBUS images. The performance of this registration technique was assessed by comparing those 3D tag points. This method resulted in an average registration error of 3.2 mm between the set of tag points. Another method known as REgistration of Ultra-SOUND volumes (RESOUND) was proposed by Rivaz and Collins [165]. This method does not directly register the brain pBUS with the iBUS images.

In the first step, the pBMR volume is registered with the pBUS image volume. In the second stage, the pBUS volume was registered with iBUS volume. Both these steps find transformation models that ultimately get combined to obtain the resulting deformation model. RESOUND [165] applies NCC as the similarity measure. Machado *et al.* [108] developed a 3D US-US registration technique that adopted a feature matching strategy. 3D-SIFT features extracted from the image pair were matched by an automated process. A rigid registration model was formulated and the model was obtained by minimizing the average Euclidean distance between the feature points. The performance of the registration was validated by comparing *mean target registration error* (mTRE) of the landmarks in both the images. Let, $\mathbf{x}_i$ $i = 1, 2, ..., K$ be an elemental landmark in the pBUS image and $\mathbf{x}'_i$ be its corresponding landmarks on iBUS image. The mTRE as stated in [108], can be written as

$$mTRE = \frac{\Sigma_{i=1}^{K} ||T(\mathbf{x}_i) - \mathbf{x}'_i||}{K} \tag{3.1}$$

Machado *et al.* [108] significantly improved the mTRE compared to the method of Mercier [122] and RESOUND [165]. Image registration using matched features requires the same key points in both the registering images. Generally, some important and unique points like ridges, corners, etc. are identified as features. But the noise and various artefacts in those images make their appearances different. In such cases, obtaining the same key points in the image pair is difficult and registration fails. On the other hand, the segmentation-based approach matches common areas in place of matching key points. Recently, Nitsch *et al.* [141] proposed a segmentation-based approach to the registration of iBUS to the pBUS images. They find one or more common HE anatomical regions from both the images, and perform registration to align those regions. A superpixel is a tiny cluster of pixels having similar intensity confined within a small area. The concept of superpixel was introduced by Achanta *et al.* [1] which was later used by many researchers [2, 141, 67, 112]. It is assumed that the pixels within a super-pixel have almost same intensity and are closed to the centroid of the cluster. Before segmentation, a rectangular region is needed to be marked by the user for focusing on the region of interest. The identified region is then divided into multiple tiny superpixels. Before training, labeling of those superpixels is done manually by the experts. The random forest classifier is used for an initial segmentation of all those superpixels. The initial classification is further improved by a couple of post-processing steps. After obtaining the final segmentation of the common anatomical

regions, these segmented images are registered using MeVis Image Registration Toolkit [141]. Local cross-correlation (LCC) was used as the similarity having a window size of $3 \times 3 \times 3$. This method showed better efficiency for registering brain US images compared to its predecessor methods. But, this method has a limitation in that it needs exhaustive training over a large dataset which is difficult for brain US images because of non-availability of labeled data. The post-processing steps after classification requires many parameters to tune which are needed to be performed carefully by trial and error. The performance of the super-pixel-based method depends on the number of super-pixels considered initially for the entire image [1, 177]. The segmentation gets improved with the increase in the number of super-pixels, that results into higher complexity.

As discussed, optimization plays a critical role in image registration. The optimization algorithms in US image registration include GD, SGD, LBFGS and QN methods. A few evolutionary optimizations like PSO [31, 174] and SA [113] were also applied. Efstathiou *et al.* [56] showed a detailed comparison of the various gradient-based optimization methods on ultrasound image registration. They concluded that, LBFGS performs as the best method among various optimization methods. Canalini *et al.* [20] performed optimization of SSE for registering pBUS and iBUS images. Compared to the gradient-based methods, the evolutionary optimization methods perform better search of global optima. These methods do not require the gradient information, and thus they do not converge into a local minimum in the error hyper-surface.

Ultrasound images suffer from poor resolution and as a result, the visibility of all regions is not good. It can been observed that the regions between the pBUS and iBUS images having good visual similarity possess sharp edges. In this work, we extract edge-rich regions by applying a threshold of the absolute patch-based gradient in both images. For making the method robust to speckle, a patch-based approach is adopted. The detailed description of the method is given in the next section.

3.2 Proposed Registration Method

The registration problem is essentially an optimization problem to obtain a suitable T for which the cost function J will be the minimum in (2.3). This work presents a rigid registration

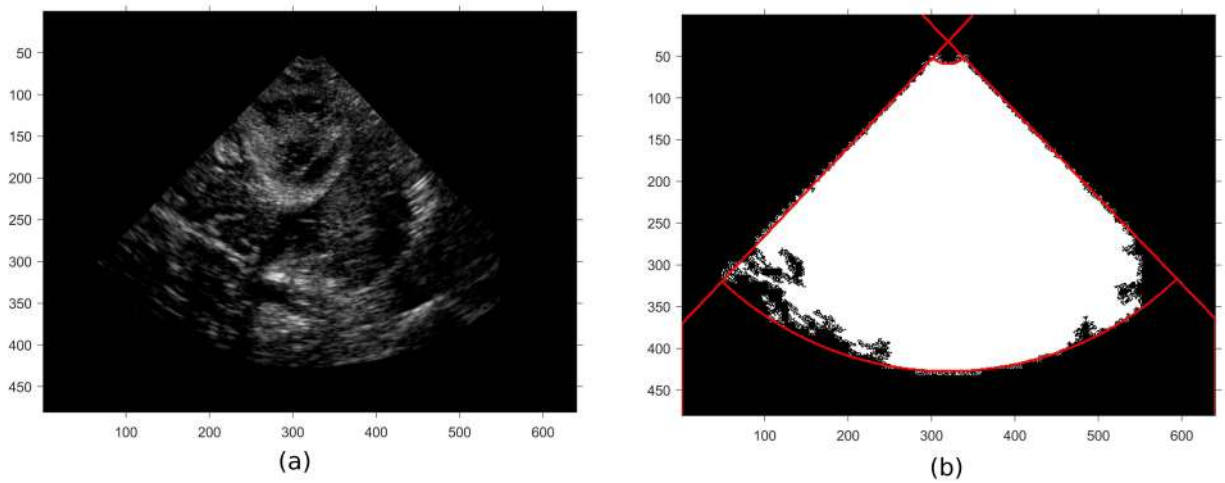


Fig. 3.2: (a) Original brain US image. (b) The white region represents the true positive area of the FOV, and the black region is false negative area within the FOV. The red coloured boundary separates both the region which is obtained by regression. The black region beyond the angular section is the redundant area

method, which considers translation, rotation, and affine transformation as the nature of relative mapping. Among the various types of US probes, the convex and phased array US probes are commonly used to examine brain tissues [23]. The field of view (FOV) region in those images is fan-shaped and placed within a rectangular grid. Hence, there are some redundant part if the image, which is treated as background, and shown in black in Fig. 3.2. The images formed by a linear US probes are rectangular. The registration is needed to be done based on the information available within the FOV. Hence extraction of the FOV is important. The proposed method of extracting the FOV is described below.

3.2.1 Extraction of FOV

The appearance the FOVs in the images produced by a convex or phased-array probe are fan-shaped. Extraction of this region is important because of various reasons that will be discussed later. A fan-shaped FOV can be described as an angular cross section of two concentric circles. The background area has an intensity mean 0. If we perform a zero intensity thresholding, the binary map roughly separates the FOV from the background. For example, if the pBUS image shown in Fig. 3.2.(a) is zero thresholded, the corresponding binary map can be seen in Fig. 3.2.(b). There is a large number of false negative pixels within the FOV, that distorts the

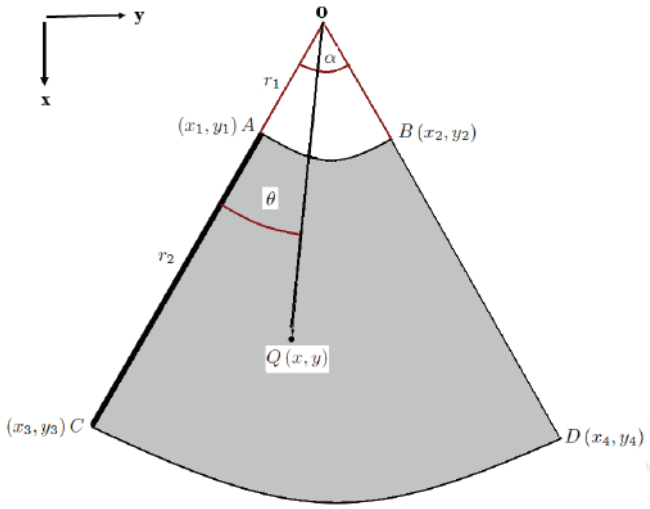


Fig. 3.3: Fan-beam representation of an US image.

shape of the FOV region.

For resolving the false segmentation, four boundaries of the FOV region are mathematically modeled as an angular section between two concentric circles namely the *inner* circle of radius r_1 and the *outer* circle of radius r_2 with $r_2 > r_1$. The FOV area is bounded by two straight lines and two circular arcs of radius r_1 and r_2 respectively. Fig. 3.3 shows a schematic diagram of the FOV region. This region is bounded by two straight lines AC and BD, and two circular arcs AB and CD. The boundary obtained after zero thresholding is shown in Fig. 3.2.(b) which does not match perfectly with an angular cross-section. From the derived binary map, it is needed to find the actual angular cross-section as shown by red boundaries in Fig. 3.2.(b). We apply a curve fitting approach for obtaining these boundaries. The straight lines AC and BD are obtained using a curve fit using linear regression technique [30]. Let, the equations of AC and BD be $y = m_1 x + c_1$ and $y = m_2 x + c_2$ respectively. The slopes m_1, m_2, and the constant parameters c_1 and c_2 can be obtained through linear regression technique. The intersection point $O\ (x_0, y_0)$ of AC and BD can be obtained by the following formula

$$\begin{bmatrix} y_0 \\ x_0 \end{bmatrix} = \begin{bmatrix} 1 & -m_1 \\ 1 & -m_2 \end{bmatrix}^{-1} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{3.2}$$

and, the angle α between AC and BD is given by,

$$\alpha = tan^{-1}\left(\frac{m_1 - m_2}{1 + m_1 m_2}\right) \tag{3.3}$$

Once (x_0, y_0) is known, the radii r_1 and r_2 are obtained by a non-linear regression technique by fitting two concentric circles with center at (x_0, y_0). Fig.3.2.(b) shows the white region as the inside region. Within the FOV area, there are large number of pixels that are false-negative. The angular section redefine the region, which solves the false-negative cases. The image plane can now be described in the polar coordinate system with O as the origin and OA as the reference direction. Consider a point $Q\,(x, y)$ in the image plane. Its polar coordinate r and θ are given by,

$$r = \sqrt{(x_o - x)^2 + (y_o - y)^2}$$

and

$$\theta = tan^{-1}\frac{m_1(x - x_o) - (y - y_o)}{(x - x_o) + m_1(y - y_o)} \tag{3.4}$$

For an arbitrary point $Q\,(x, y) \in F$, the following conditions are satisfied:

$$Radius\ condition : (r_1 \leq r \leq r_2) \tag{3.5}$$

$$Angle\ condition : (0 \leq \theta \leq \alpha)\ and\ x > x_0 \tag{3.6}$$

The above model is applicable for the images using the phased array or the convex probes and are vertically aligned. If the images are not vertically aligned, appropriate rotation with respect to the axis of the fan-shape region is required. The images formed by a linear probe appear rectangular and they are modeled as a quadrilateral. All four sides of it can be reconstructed after zero thresholding and using the linear regression method discussed above.

After completion of this step, the FOV of the pBUS image is assumed to be the reference, and that of the iBUS image is assumed to undergo a deformation T. J in Equation 2.3 is computed and optimized for the best set of registration parameters. Before applying T, both the FOVs align with each other. But, when the sensed image undergoes a transformation, the

FOV of the iBUS image shifts, and thus only a partial region overlaps with the FOV of the pBUS image. It forms two sub-regions of the FOV: a) the common region X^{common} between the FOVs and b) the union of the rest of the regions of the FOVs. The method of extraction of the common region is described as follows.

3.2.2 Extraction of the Common region from the FOV

Let V_R and V_S represent the pBUS and iBUS images respectively. An example of pBUS and iBUS images can be seen in Fig. 3.4.(a1) and Fig. 3.4.(b1) respectively. For obtaining the FOV, a threshold operation is carried out on all the pixels, and the non-zero intensity pixels are considered inside the FOV. For example, after the application of zero thresholding on the image in Fig.3.4.(b1), it produces a FOV as shown in Fig.3.4.(b2). Many tiny red regions are seen in image because many pixels within the actual FOV have zero intensity. A morphological flood fill operation is applied for filling those tiny regions. The region X^c represents the RC, which is absent in V_S. The FOVs of two registering images after hole filling are shown in Fig.3.4.(a3) and (b3). X^C region should be excluded from the similarity calculation. There are some tiny HE regions present within X^c due to the presence of air bubbles in the RC. These tiny regions are marked using connected component labeling. The smaller regions having less than 100 pixels within them are eliminated from the binary map. Fig.3.4.(a4) shows the FOV of the iBUS image excluding RC. The curve fitting approach is then followed to obtain the find the final FOVs. as described in the previous subsection.

Suppose the FOV of iBUS image undergoes a rigid transformation ($\mathbf{T}$), the coordinate space $\mathbf{X}$ gets shifted to $\mathbf{TX}$. This spatial transformation produces some regions that do not overlap with the coordinate space of the FOV of pBUS image as shown in Fig.3.4.(b4). The correlation between the registering images is negligible in these regions. The common region, where the deformed image and the reference image overlap, may consist of similar pixels of the two images. Thus, the coordinate space $\mathbf{X}^{common}$ of that region can be defined as

$$\mathbf{X}^{common} = \mathbf{X} \cap T(\mathbf{X}) \tag{3.7}$$

Apart from the reason of the mismatch of FOVs due to T, the region $\mathbf{X}^c$ of the RC also causes dissimilarity. The RC region appears to be completely anechoic but some hyper-echoic tiny

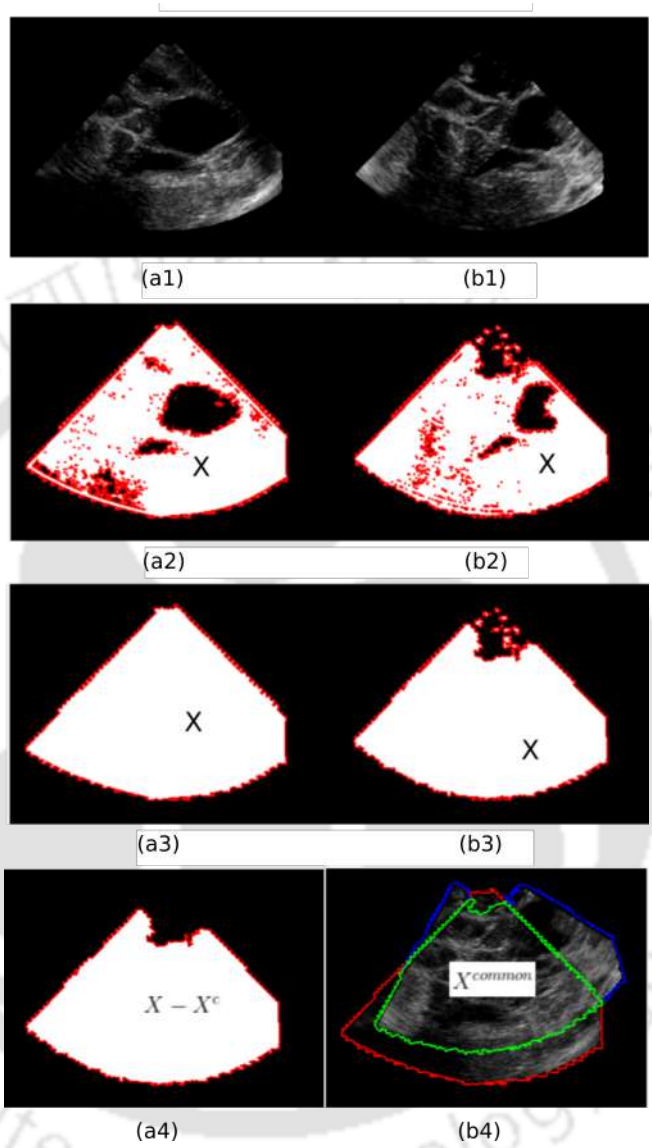


Fig. 3.4: (a1-b1) show the original pBUS and iBUS image pair respectively. X^c is the resected tumor region that is absent in the iBUS image. (a2-b2) represent the binary map of FOVs after zero intensity thresholding. (a3-b3) show the FOVs after morphological hole filling operation on the binary maps in (a2-b2). (a4) shows the FOV excluding the RC. (a5) shows the reference and the moving image in a single frame. X^{common} is the common area that intersects between two FOVs. This region is shown by a green boundary.

regions sometimes appear due to the air bubbles present as in Fig.3.4.(b3). Because the nature of the intensity in X^c is different in these images, this region should also be excluded. These bubbles are small in size and can be identified using connected component analysis and reject smaller regions applying thresholding on the areas of the disconnected regions [32]. Out of all these disconnected regions, the biggest region represents the FOV excluding the RC which can be seen in Fig.3.4.(a4).

Hence, after the removal of X^c from the common region, (3.7) is modified as,

$$\mathbf{X}^{common} = \mathbf{X} \cap T(\mathbf{X} - \mathbf{X}^c) \tag{3.8}$$

The next step for extracting the edge-rich regions because they posses more similarities compared to the other regions. The registration model adopted here is global and rigid and they are performed based on the edge-rich regions only. To align the target image closely with the reference image, the estimation of the six optimum parameters as given in (3.10) is required. Suppose an image $\mathbf{V}_S$ is transformed to $\mathbf{V}_T$ by the transformation matrix T. The transformed image is given by

$$\mathbf{V}_T(\mathbf{X}) = \mathbf{V}_S(\mathbf{TX}) \tag{3.9}$$

where

$$\mathbf{T} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \tag{3.10}$$

$\mathbf{T}$ includes translation, rotation, and affine transformation. The block diagram of the proposed approach is shown in Fig.3.5.

In most cases, the registration of the pBUS and iBUS images is difficult because of the poor similarity between the image pairs. Thus, it requires a preprocessing step on $\mathbf{V}_S$ and $\mathbf{V}_T$ for extracting the edge-rich regions $\widehat{\mathbf{V}}_S$ and $\widehat{\mathbf{V}}_T$ respectively. The *mean squared error*(MSE) between the registering image pairs is taken as the dissimilarity metric. Thus the cost function J between $\widehat{\mathbf{V}}_S$ and $\widehat{\mathbf{V}}_T$ can be written as

$$J = \frac{1}{|\mathbf{X}^{common}|} \sum_{\mathbf{X} \in \mathbf{X}^{common}} (\widehat{\mathbf{V}}_S(\mathbf{TX}) - \widehat{\mathbf{V}}_T(\mathbf{X}))^2 \tag{3.11}$$

where J is a function of the parameters a, b, c, d, e and f. Thus, the optimization problem can

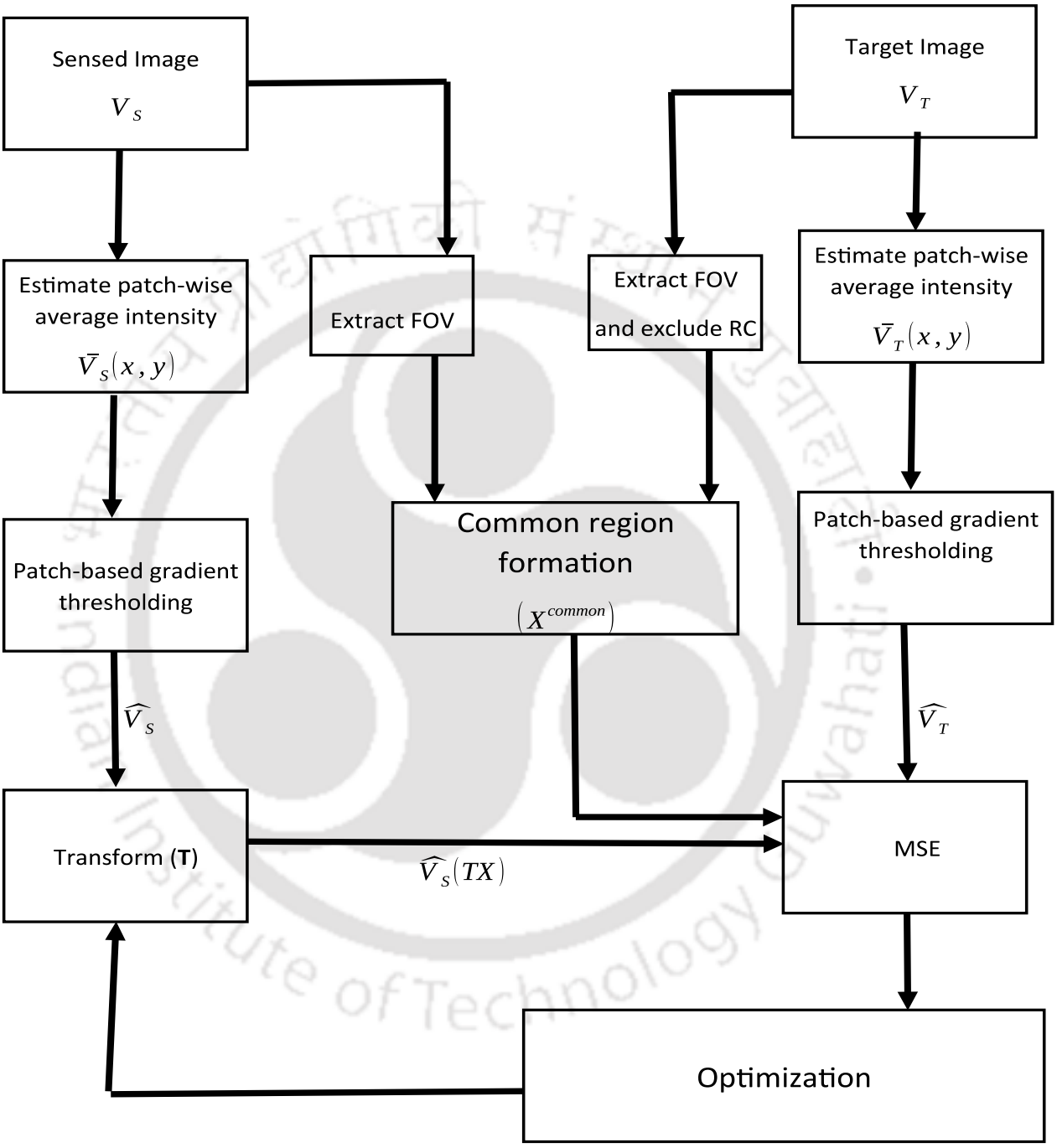


Fig. 3.5: Block diagram of the proposed method

be formulated as,

$$\min_{a,b,c,d,e,f} J \tag{3.12}$$

Canalini *et al.* [20] applied LBFGS optimization for minimizing the sum of squared error(SSE) between the registering image pairs. For exploring the possibility of a better optimization method, we have applied LBFGS, and a few popular heuristic optimization methods such as PSO, BBO, and SA. Thus, the novelty of the proposed method is therefore two-fold, i) extraction of the FOV and $\mathbf{X}^{common}$, and ii) adopting a patch-based gradient threshold scheme that is able to extract the edge-rich regions of the registering image pairs.

3.2.3 Gradient thresholding

It has been observed that many regions in an iBUS image do not have sharp edges, and thus they do not appear similar to the corresponding regions in the pBUS image. These regions should be excluded in estimating the registration cost function. In areas where there is a progressive change in the acoustic impedance, ultrasound exhibits significant echogenicity. The homogeneous regions, such as the CSF, and ventricles, on the other hand do not produce echoes, and therefore do not include sharp edges. Thus, the regions with sharp edges are generally identical in pBUS and iBUS images, and the rest of the regions affect the registration wrongly.

In general, pixel intensity gradients are higher for pixels at the edges of a region. Based on this observation, Chalfoun *et al.* [27] proposed a segmentation approach based on thresholding the gradient image. Thresholding has a short execution time and small memory requirements. To alleviate the issue of poor edges in pBUS and iBUS images, a binary map S is formed by including the sharp-edge regions to indicate the pixels to be included for registration. The non-sharp edge regions can be excluded pixelwise. For example, if the absolute intensity gradient(AIG) of a pixel is greater than a threshold τ, S is assigned a 1 at that point. Otherwise, it is discarded by assigning 0 in S. When the aforementioned process is applied to an image, the sharp-edge regions are seen. For various values of τ, the modified pBUS and an iBUS image are displayed in Fig. 3.6. It is observed that many tiny regions appear in the gradient-thresholded image and the boundary information become ambiguous. To overcome this, a patch-based gradient thresholding scheme is proposed below.

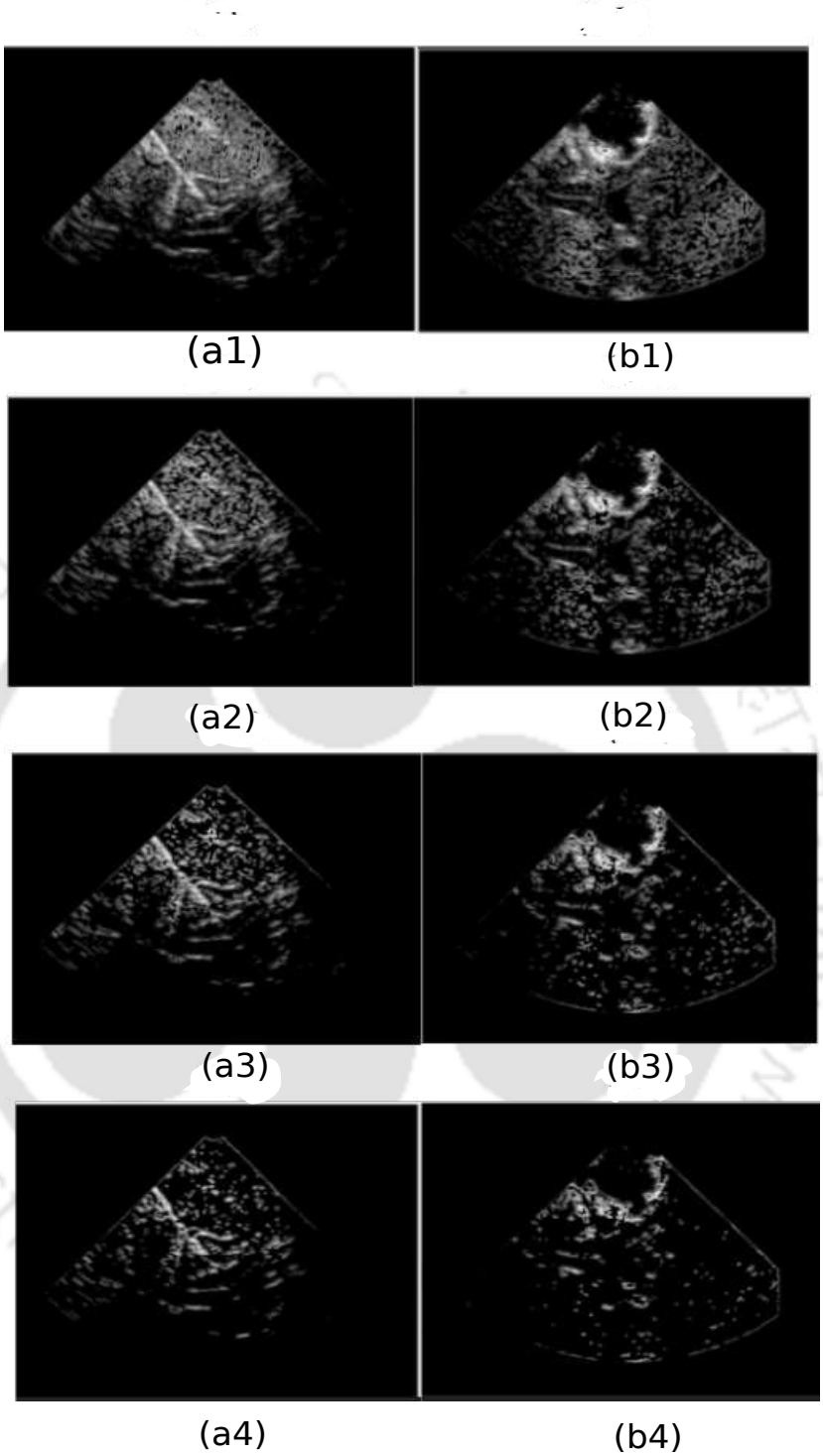


Fig. 3.6: (a1-a4) show the images $\bar{V}$ after gradient thresholding of pBUS image. Similarly, (b1-b4) show $\bar{V}$ of a iBUS image. (a1-b1) images resulted for $\tau = 15$, (a2-b2) images resulted for $\tau = 20$, (a3-b3) images resulted for $\tau = 25$, and (a4-b4) resulted for $\tau = 30$.

3.2.4 Patch-based gradient thresholding

Like edges, noise is also a high frequency component. To exclude this, patch-based thresholding of the AIG is proposed. The averaging process reduces the variance of the noise, and lowers the gradient values.

Suppose, $\mathbf{V}$ is an image of size $M \times N$, and the image grid is uniformly divided into non-overlapping patches $P(m,n)$ of size $w \times w$. The mean intensity of $(m,n)^{th}$ patch is given by $\overline{V}(m,n) = \frac{1}{w^2}\Sigma_{\forall(i,j)\in P(m,n)}V(i,j)$. The patch-based AIG at $(m,n)^{th}$ patch is given by,

$$\left|\nabla\overline{V}(m,n)\right| = \frac{1}{2}\sqrt{\left(\overline{V}(m+1,n) - \overline{V}(m-1,n)\right)^2 + \left(\overline{V}(m,n+1) - \overline{V}(m,n-1)\right)^2} \tag{3.13}$$

The binary edge map $S(i,j), \forall(i,j) \in P(m,n)$ is set for each patch as,

$$S(i,j) = \begin{cases} 1 & \text{if } \left|\nabla\overline{V}(m,n)\right| > \tau \\ 0 & \text{otherwise} \end{cases}$$

where τ is threshold fixed experimentally. To get the edge-rich image ($\widehat{\mathbf{V}}$), S is multiplied element-wise with $\mathbf{V}$. The resulting $\widehat{\mathbf{V}}$ for patch-size of 4×4 and 8×8 is shown in Fig. 3.7 and 3.8 respectively.

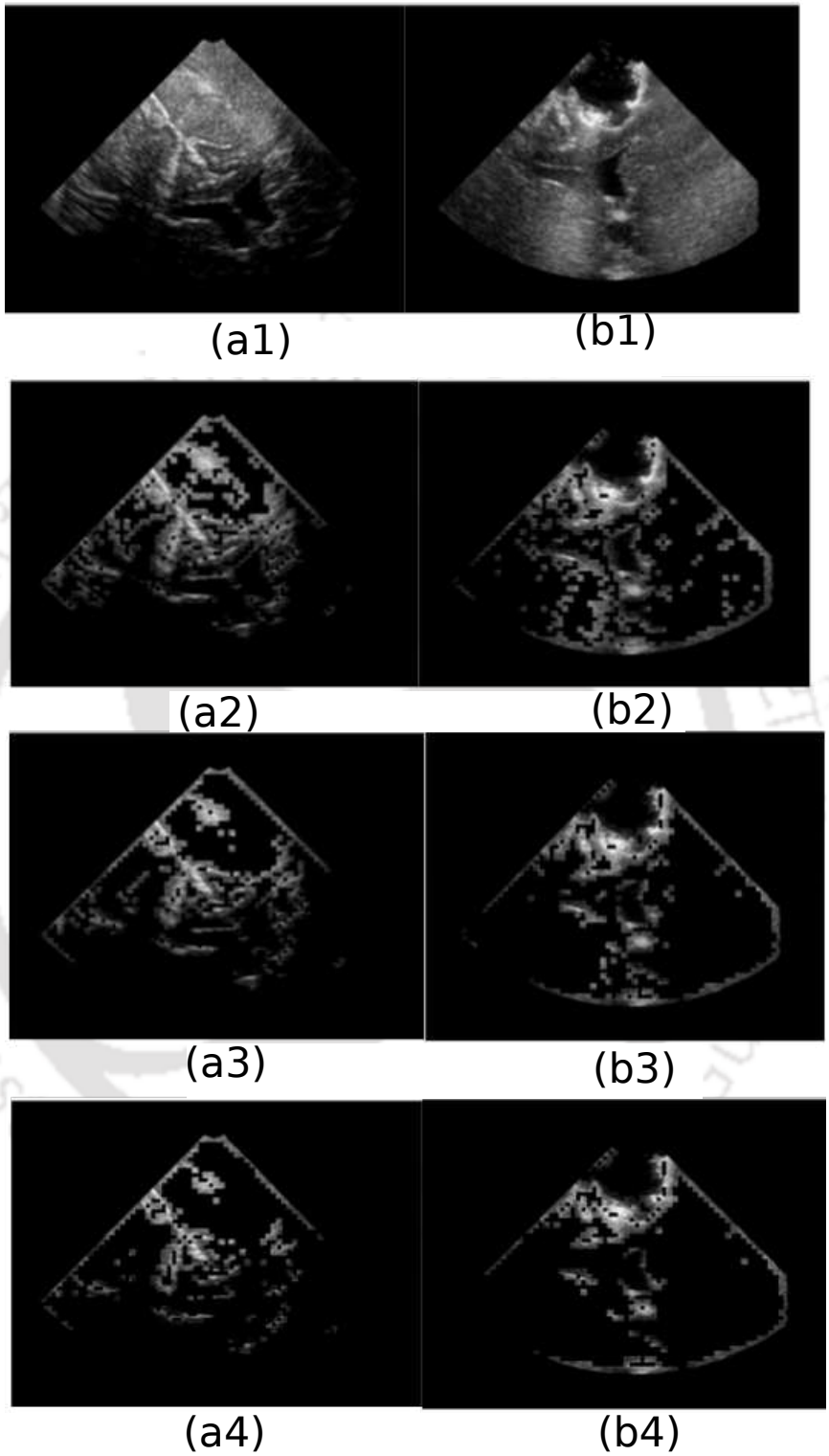


Fig. 3.7: Output of patch-based gradient thresholding is shown by varying $\tau = 15, 20$, and 25 for $w = 4$. Here (a1) and (b1) show the original pBUS iBUS images respectively. The output pBUS and iBUS images for $\tau = 15$ are shown in (a2-b2). Similarly (a3-b3) show the same outputs for $\tau = 20$, and (a4-b4) show the outputs for $\tau = 25$

.

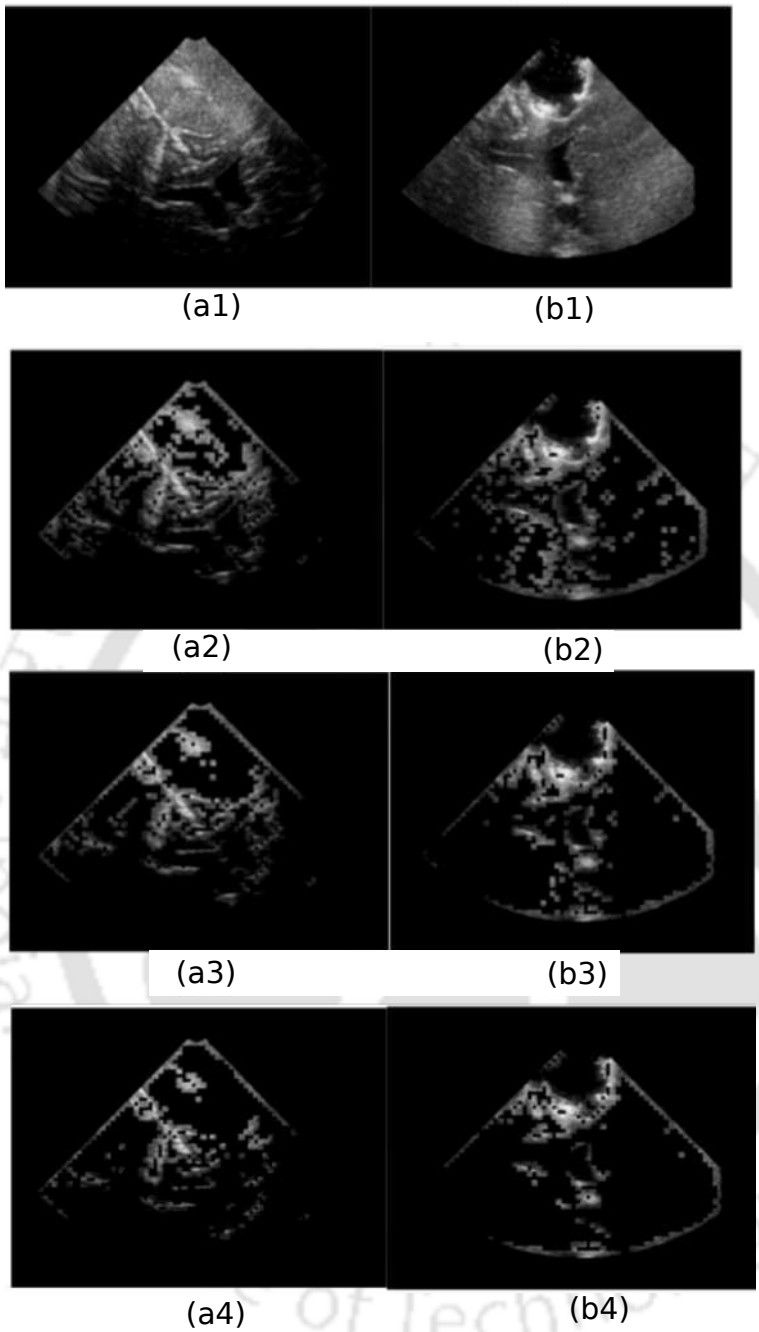


Fig. 3.8: Output of patch-based gradient thresholding is shown by varying $\tau = 10, 15, 20$, and 25 for $w = 8$. Here (a1) and (b1) show the original pBUS iBUS images respectively. The output pBUS and iBUS images for $\tau = 15$ are shown in (a2-b2). Similarly (a3-b3) show the same outputs for $\tau = 20$, and (a4-b4) show the outputs for $\tau = 25$

.

Fig.3.7 shows that there are lower number of tiny regions for a patch size of 4×4 compared to the pixel-based method as shown in Fig.3.6. The selection of the threshold is a crucial point in this experiment. The intensity profiles are different in the RESECT and the BITE datasets. For

obtaining the threshold automatically, we have collected patch-based AIG profile from 10000 images each of the BITE and RESECT datasets. The absolute values of the gradient as shown in Equation (3.13) was estimated patch-wise for each of the images. The gradient-arrays were estimated from 10000 images from each of the datasets. The *normalized histogram of the AIG* (NHAIG) of the BITE and RESECT datasets are shown in Fig. 3.9.(a) and (b) respectively. Both of these images contains two histograms with red and blue plots. It can be carefully seen that the BITE and the RESECT datasets have different NHAIG profile and therefore, the value of τ for them should be different. Otsu's threshold estimation [147] method minimizes the intra-class variance of the average gradient between the background and foreground, and finds τ for separating them.

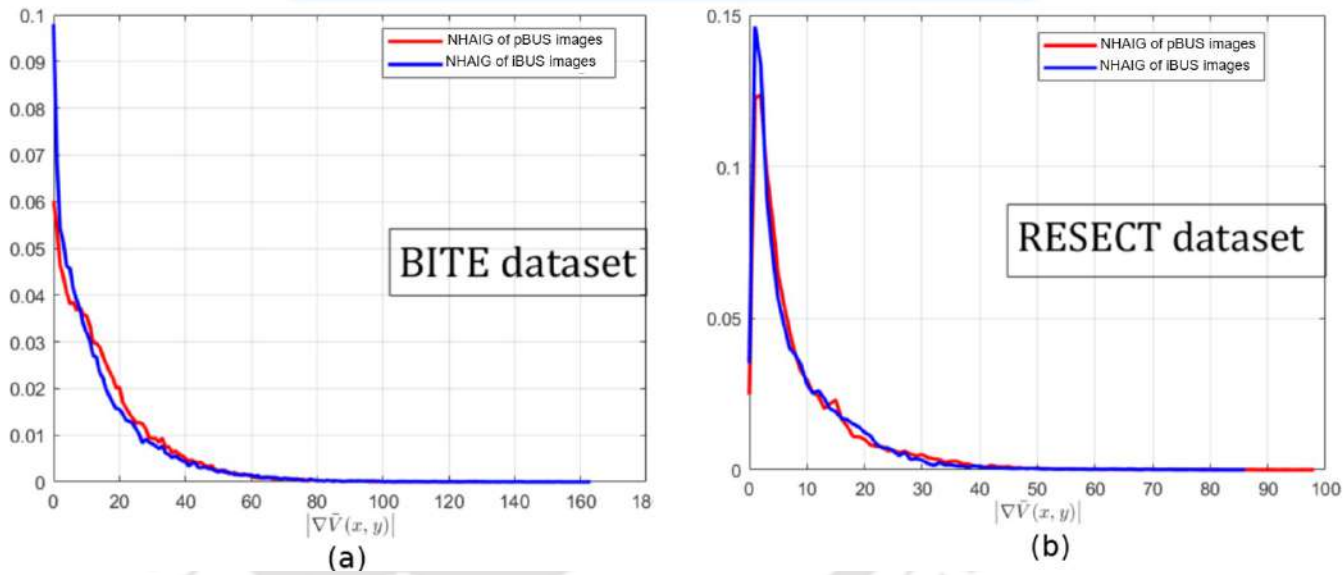


Fig. 3.9: Normalized histogram of AIG for pBUS and iBUS images for a) BITE dataset (b)RESECT dataset.

We have tested effect of the automatically obtained threshold on the pBUS and iBUS images. It was performed in a subjective manner by visually inspecting the pBUS and iBUS images placing them side by side. Various threshold values of NHAIG were applied on both the images. The images for $\tau = 15, 20$ and 25 are shown in Fig. 3.6(a2-b2), (a3-b3) and (a3-b3) respectively. It is observed that the similarity between image pairs improves with the increase in the value of τ from 15 to 25 for the BITE dataset. When τ is increased beyond 25 for the BITE dataset, and 14 for the RESECT dataset, information gets reduced from both the images. Thus $\tau = 25$ and $\tau = 14$ were found experimentally as the best values for the BITE and the RESECT datasets

respectively.

Like τ, selection of w is also important here. We have examined different values of w and recorded their outputs. During this operation, we kept the optimum values of τ for both the datasets. Fig.3.8 shows the resulting images for $w = 8$. Though the number of tiny undesired regions is small, but the boundaries appear like blocks, that distorts the shapes of the important regions. It reduces the similarity between the registering image pairs. Hence, adopting an appropriate patch size is crucial. Experiments were also carried out for $w = 2, 3, 4,$ and 6 with optimum τ for both the datasets. Corresponding results various patch-sizes are shown in Fig.3.10.

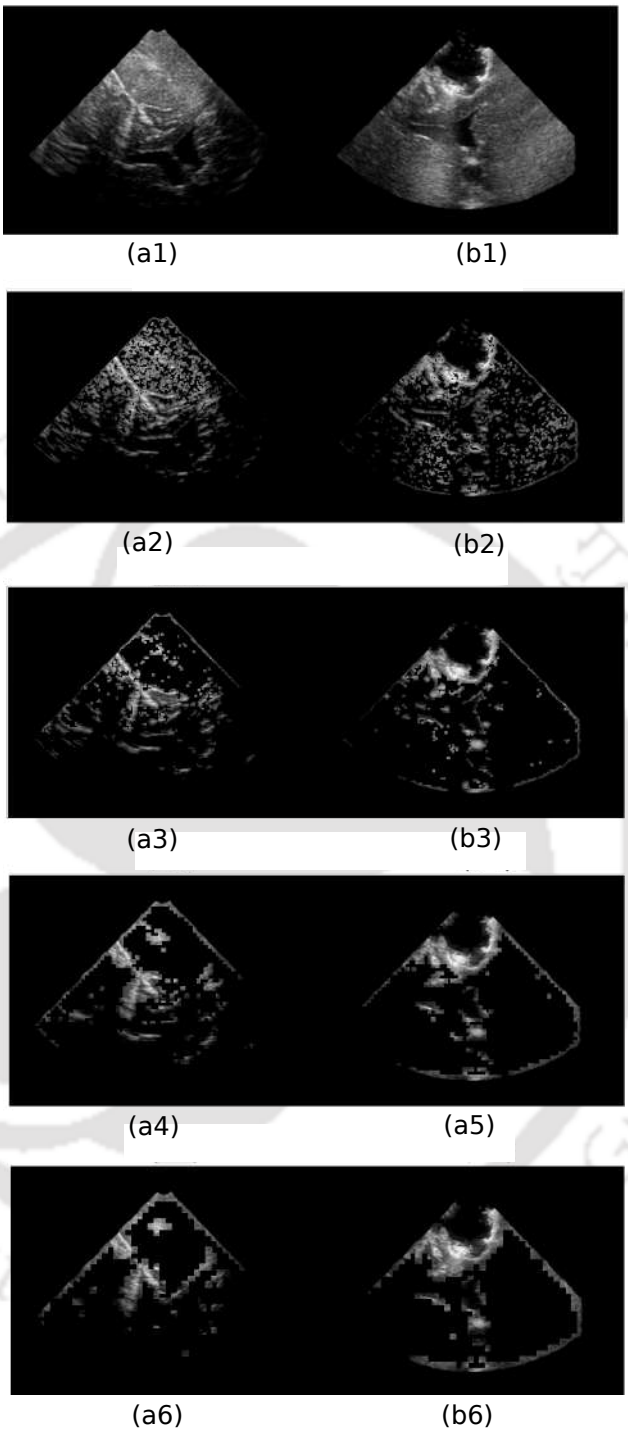


Fig. 3.10: (a1-b1) respectively are the pBUS and iBUS images from the BITE dataset respectively. w is varied from 3,4,5, and 6. $\tau = 25$ was considered for BITE dataset for these images. The resulting images are shown in (a2-b2),(a3-b3), and (a4-b4) respectively.

This figure shows that an increase in the value of w reduces the number of tiny regions until $w = 4$. An increase in w increases the blokieness of and it becomes significant for $w > 4$.

Thus $w = 4$ was was considered optimum for the experiments. The method of obtaining a patch-based gradient thresholded image is summarized in Algorithm 3.

Algorithm 3 Patch-based Gradient Thresholding

Divide the images $\mathbf{V}$ into non-overlapping square grid of size$m \times m$.
for each block $P(m,n)$ in the grid **do**
  calculate the mean $\overline{V}(m,n) = \frac{1}{m^2}\Sigma_{\forall(i,j)\in P(m,n)}V(i,j)$.
  Find the magnitude of the gradient of $\overline{V}(m,n)$ given by,
$\left|\overline{V}(m,n)\right| = \frac{1}{2}\sqrt{\left(\overline{V}(m+1,n) - \overline{V}(m-1,n)\right)^2 + \left(\overline{V}(m,n+1) - \overline{V}(m,n-1)\right)^2}$
  for $(i,j) \in P(m,n)$ **do**
    if $\left|\nabla\overline{V}(m,n)\right| > \tau$ **then** $S(i,j) = 1$
    else$S(i,j) = 0$
    end if
  end for
end for
Obtain the thresholded image $\widehat{\mathbf{V}} = \mathbf{V} \odot \mathbf{S}$, where $\odot$ indicates point-wise multiplication.

3.2.5 Optimization

Optimization is used to minimize the registration error and it plays a critical role in registration. Gradient descent(GD) is a slow method and the initial guess value of the solution should be close to the optimum. Comparatively, the LMA is a fast and efficient method for minimizing a convex cost function. This method was used in the registration of the artery wall of heart US images, [144]. Rivaz and Collins [165] used the stochastic gradient descent method and a patch-based approach in REgistration of ultraSOUND (RESOUND) volumes. In a modified version of RESOUND, Zhou and Rivaz [217] used a gradient-based efficient second-order minimization [117]. Canalini *et al.* [21] applied a quasi-Newton(QN) approach, namely the LBFGS algorithm for registering BUS volumes. Canalini *et al.* [20] performed optimization of SSE for registering pBUS and iBUS images. We choose the LBFGS algorithm for its better convergence characteristics.

Compared to the gradient-based methods, the evolutionary optimization methods perform better search of the global optimum. These methods search for the global minimum in the error hypersurface. Gradient-based methods often fail in multimodal optimization unless a good initial solution is anticipated. Comparatively, heuristic-based PSO may be useful [31, 174]. For exploring the possibility of better results, we have compared the registration results by applying

three different heuristic optimization methods: PSO, biogeography-based optimization(BBO) [183], and simulated annealing (SA) [113]. A brief outline to these algorithms were described in Section 2.5. They all search for the optimum solution over the entire search range and converge to the global optimum. These methods are simple, robust to noise, and also effective in obtaining global optimum for a multimodal objective function. The performance of three different heuristic optimization methods and LBFGS are compared in the next section.

3.3 RESULT AND DISCUSSION

Experiments were carried out to study the performance of the proposed method.43 image pairs from the BITE dataset and 25 image pairs from the RESECT dataset were used for the evaluation of registration performance and the comparison with other BUS image registration methods. Selection of image pairs were done with the help of an expert radiologist. For evaluating the performance of the registration methods, 4-8 common tag-points were chosen from both pBUS and iBUS images. In most cases, the tag-points were selected from the common ridges, ventricle horns, corners, etc. The mTRE was used for the evaluation of the registration performance.

Let the points $\mathbf{x}_i$ and $\mathbf{y}_i$, {i=1,2,3,...,K} be selected from the iBUS image ($\mathbf{V}_S$) and pBUS image ($\mathbf{V}_R$) respectively. The value of mTRE before registration (E_b) can be formulated as,

$$E_b = \frac{\Sigma_{i=1}^{K} ||\mathbf{x}_i - \mathbf{y}_i||}{K} \tag{3.14}$$

For an optimal transformation $\mathbf{T}$, $\mathbf{x}_i$ shifts to $\mathbf{T}\mathbf{x}_i$. Hence, the post-registration mTRE becomes

$$E_a = \frac{\Sigma_{i=1}^{K} ||\mathbf{T}\mathbf{x}_i - \mathbf{y}_i||}{K} \tag{3.15}$$

The method that produces higher value of $(E_b - E_a)$, can be assumed to be a better method.

3.3.1 Results of registration using patch-based thresholding of AIG

As discussed in the previous sections, we performed registration for $w = 1, 4, 8$, and 16. Four different optimization methods, namely LBFGS, BBO,PSO, and SA were applied for minimizing MSE as described in (3.11). All these optimization algorithms are sensitive to

LBFGS	Initial Hessian: Identity matrix of size 6x6	BBO	Keep rate=0.2
	Solution started from zero point.		Population size=100
			$\alpha = 0.9$, $\sigma = 0.02(Upperbound - Lowerbound)$
PSO	Inertia: $0.9 < \eta < 1$	SA	Solution started from zero point
	Population size=100		Cooling Schedule: Linearly decreasing from 300 to 50 over 100 iterations
	$c_1 = c_2 = 0.1$		Neighborhood range: A square region having sides equal to the current temperature.

Table 3.1: Important parameters of four optimization algorithms

their parameters. We varied them within a wide range for obtaining the best results. The values of the various parameters of these optimization algorithms are chosen by trial and error for the best performance and shown in Table 3.1. The nature of the intensity profile and artefacts are different for the BITE and the RESECT dataset. We have conducted the tests separately for the these two datasets. The optimum threshold of AIG for the BITE dataset is considered 25, whereas the same value is 14 for the RESECT dataset as mentioned in Section 3.2.4. It was shown graphically that $w = 4$ is the optimum patch-size, because it increases the similarity between the pBUS and iBUS images without much deterioration of shape on the boundaries. For validating this hypo-thesis, we varied $w = 1, 4, 8$, and 16 for both the BITE and the RESECT datasets. There were 43 pairs of images from the BITE dataset and 25 pairs of images from the RESECT dataset. Depending on the visual similarities of the registering image pairs, we divide them into three subcategories: *easy, moderate* and *difficult*. The image pairs having good visual similarity among them, are considered easy. If the similarity is good and artefacts are more, they are considered moderate. Some image pairs where similarity can be identified by the experts only are referred to as the difficult image pairs. We have found that out of the 68 pairs of images, only 46 pairs are in the category of easy and moderate. We have presented parametric results based on those 46 image pairs only. The image pairs, which could not be registered by the proposed method and the other state-of-the-art methods, are considered as the difficult image pairs and excluded from the performance evaluation. The performance of the competing optimization methods on the images of the RESECT and the BITE datasets, are shown in Fig. 3.11.

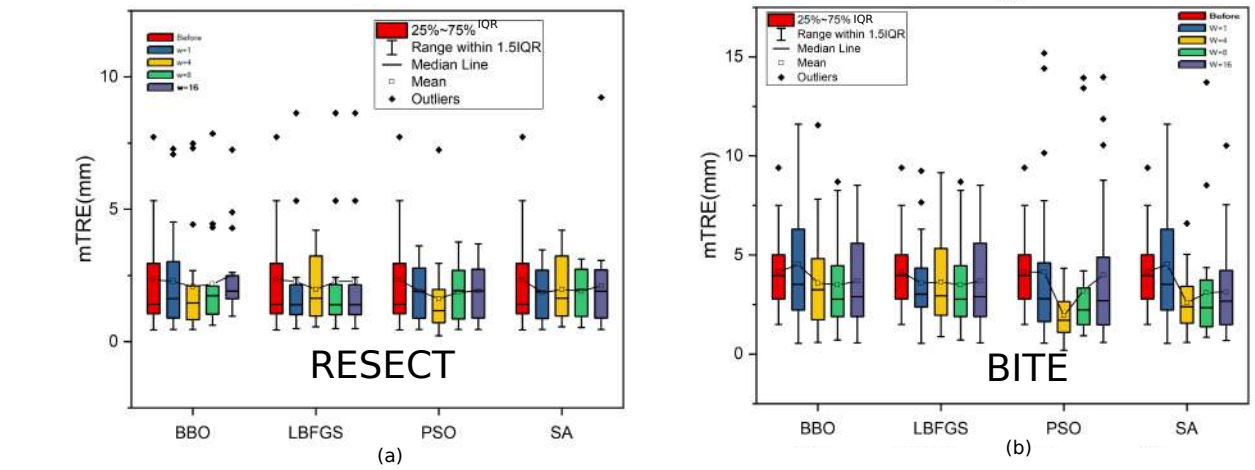


Fig. 3.11: Comparison of the pixel-based and the patch-based thresholding of AIG in the registration of pBUS and iBUS images of the RESECT dataset is shown. (a) shows the comparative mTREs are shown for the images of the RESECT dataset and (b) shows the same for the BITE dataset. Each figure shows the results for LBFGS, PSO, SA, and BBO algorithms. For each of them, five box plots are shown. The leftmost boxplot shows E_b. Subsequent four boxes show E_a for $w = 1, 4, 8, and 16$ respectively.

Fig. 3.11.(a) and (b) show respectively the comparative mTREs for the RESECT and the BITE datasets. Each figure consists of four different subgroups for showing the results by LBFGS, BBO, PSO and SA. Each subgroup contains 5 boxplots. In each subgroup, the left most boxplot shows the mTRE before registration. The subsequent bloxplots show the same results after registration with $w = 1, 4, 8,$ and 16 respectively. These figures show that the mTRE before registration does not improve with pixel-based approach i.e $w = 1$. The average mTRE increases from $w = 4$ to $w = 8$ and it further increases with $w = 16$.

The mean and the standard deviation of the MTREs of the RESECT and BITE datasets are shown in Table 3.2 and Table 3.3 respectively. For the images of the RESECT dataset, the mTRE value of 2.341 mm before registration and reduces to 1.615 mm for $w = 4$. Similarly, for the image pairs of BITE dataset, the mTRE before registration was 4.136 mm. PSO shows the best results and improves mTRE to 1.931 mm. We found that SA is the next best method compared to the BBO and LBFGS. This is to note that the results in Fig. 3.11 and the entries in Table 3.2, Table 3.3 were estimated based on those image pairs which could show improvement in mTRE by at least one optimization method. In some cases, the RC is large and it partially remains in the iBUS image even after segmentation. For those images, the

	(mm)	Before	$w = 1$	$w = 4$	$w = 8$	$w = 16$
LBFGS	Mean	2.341	2.274	1.971	2.274	2.274
	Std. Dev	2.136	2.318	1.261	2.318	2.318
BBO	Mean	2.341	2.274	2.081	2.173	2.523
	Std. Dev	2.136	1.88	1.941	2	1.708
PSO	Mean	2.341	1.916	**1.615**	1.873	1.908
	Std. Dev	2.136	1.011	1.449	0.99	1.026
SA	Mean	2.341	1.854	1.971	1.905	2.114
	Std. Dev	2.136	0.971	1.261	0.905	1.769

Table 3.2: mTRE for different optimization method with $w = 1, 4, 8, 16$ for RESECT dataset.

	mTRE	Before	$w = 1$	$w = 4$	$w = 8$	$w = 16$
LBFGS	Mean	4.136	3.582	3.636	3.487	3.681
	Std. Dev	1.923	2.016	2.202	2.205	2.2
BBO	Mean	4.136	4.539	3.585	3.487	3.681
	Std. Dev	1.923	2.916	2.598	2.205	2.292
PSO	Mean	4.136	4.144	**1.931**	3.206	3.997
	Std. Dev	1.923	3.976	1.041	3.37	3.695
SA	Mean	4.136	4.539	2.585	3.096	3.139
	Std. Dev	1.923	2.916	1.454	2.799	2.355

Table 3.3: mTRE for different optimization method with $w = 1, 4, 8, 16$ for BITE dataset.

RC region does not match with the tumor region in the pBUS image. Those image pairs were found difficult to register by this method. Similarly, for many image pairs, some HE regions appear distinctly in the pBUS image, but connected together in the iBUS image, which leads to failure in registration. Hence, we exclude the registration results for images. It is found that for almost all the images, $w = 4$ results into better mTRE because, thresholding of the AIG with $w = 4$ captures better similarity between pBUS and iBUS images.

3.3.2 Comparison with other methods

As mentioned earlier, there are only a few works reported for the registration of pBUS and iBUS images. The original implementation of Machado *et al.* [108] was performed for 3D pBUS and iBUS volume images. 3D-SIFT features were extracted from the registering image pairs and feature-matching was performed using the maximum likelihood approach. The likelihood function incorporates the position of the feature points, the SIFT descriptor, and the scale. We followed the same methodology, and extracted 2D-SIFT features from the registering image pairs, and applied the maximum likelihood approach for feature matching. Experimental results showed that this feature extraction scheme captures most of the features on the boundary of the FOV regions as shown in Fig. 3.12.

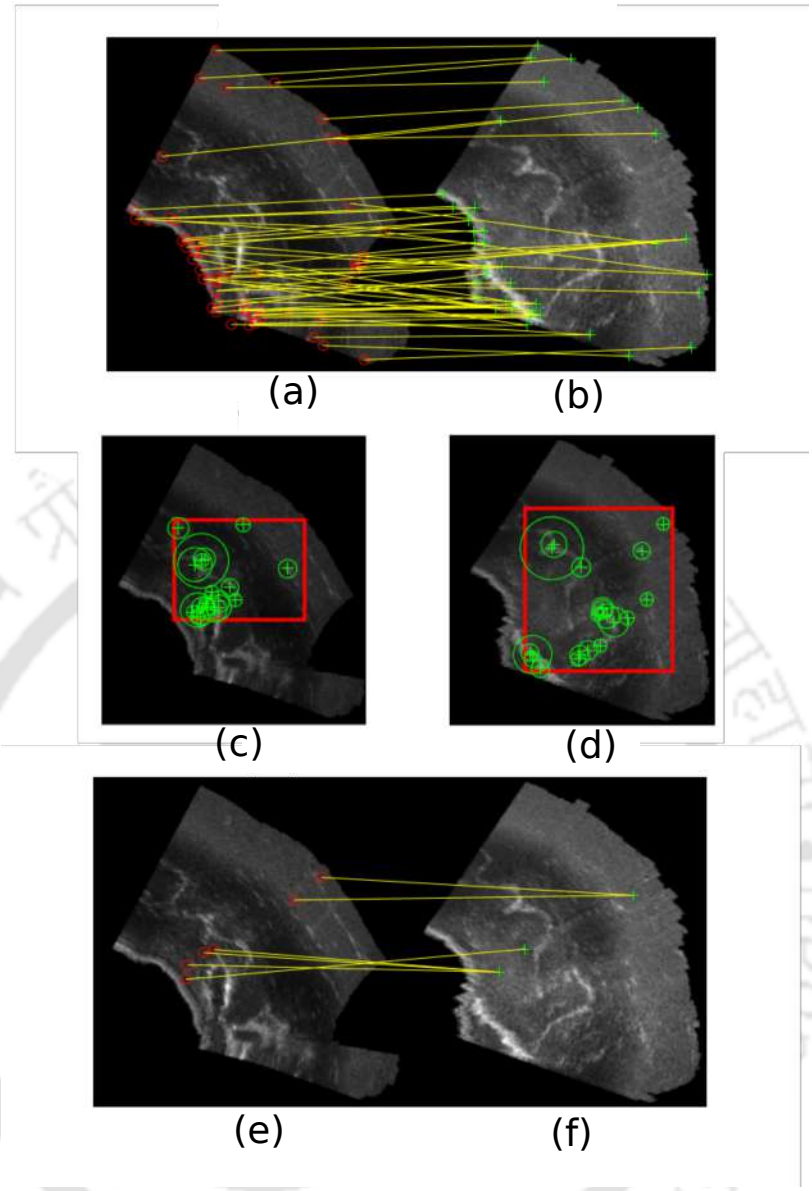


Fig. 3.12: (a) and (b) are the pBUS and its corresponding iBUS images respectively. Two hundred strong matching feature points from each image are shown in (a) and (b). (c) and (d) show 20 prominent features within the rectangular region marked by the red boundary. (e) and (f) show the matching features using the maximum likelihood approach with a radius of search of 20 mm

Fig. 3.12.(a) and (b) show a pBUS-iBUS image pair, and their corresponding matched features. Feature-matching was done with the exhaustive search approach [135]. The maximum distance-constraint between two matching features was set to an empirically obtained distance of 20 mm. Many features in pBUS images do not match properly with the corresponding features in iBUS images. This mismatch leads to degradation of the overall registration performance. The Most of the features are found to be located in the boundary of the FOV, which are not true

features. To avoid this, we have manually selected a rectangular region that covers the large possible area of the foreground, and restricted the feature extraction from that region only. For this approach, 20 strong features are shown in Fig. 3.12.(c) and (d) which are restricted within the red colored rectangular region marked manually. If the shape of the foreground region is complex, then multiple rectangular regions, circular regions or region with manually drawn boundary can also be considered. The mapping of features from Fig. 3.12.(e) to Fig.3.12.(f) show that this unable to match the features.

The work model of Canalini *et al.* [21] was also implemented for 2D pBUS and iBUS images. They separated the RC using a trained U-Net as prework, which requires a large annotated dataset for training. The RC is an anechoic region that is easy to separate. Our proposed method considered the separation of the RC in a simple manner as described in Section 3.2.2. The LBFGS optimization method used by Canalini *et al.* is a gradient-based method that often converges into a local minimum. It performs successful registration when the initial solution point is close to the global minimum. It is illustrated by the pBUS and corresponding iBUS images in Fig. 3.13.(a) and (b). These images show that the same image pairs produce different results for different start points. Fig. 3.13.(c) correctly registers the image pairs when $a = b = d = e = 0.2$ and $c = f = -10$. The same algorithm fails when the initial point is $a = b = c = d = e = f = 0$ which can be seen in Fig. 3.13.(d).

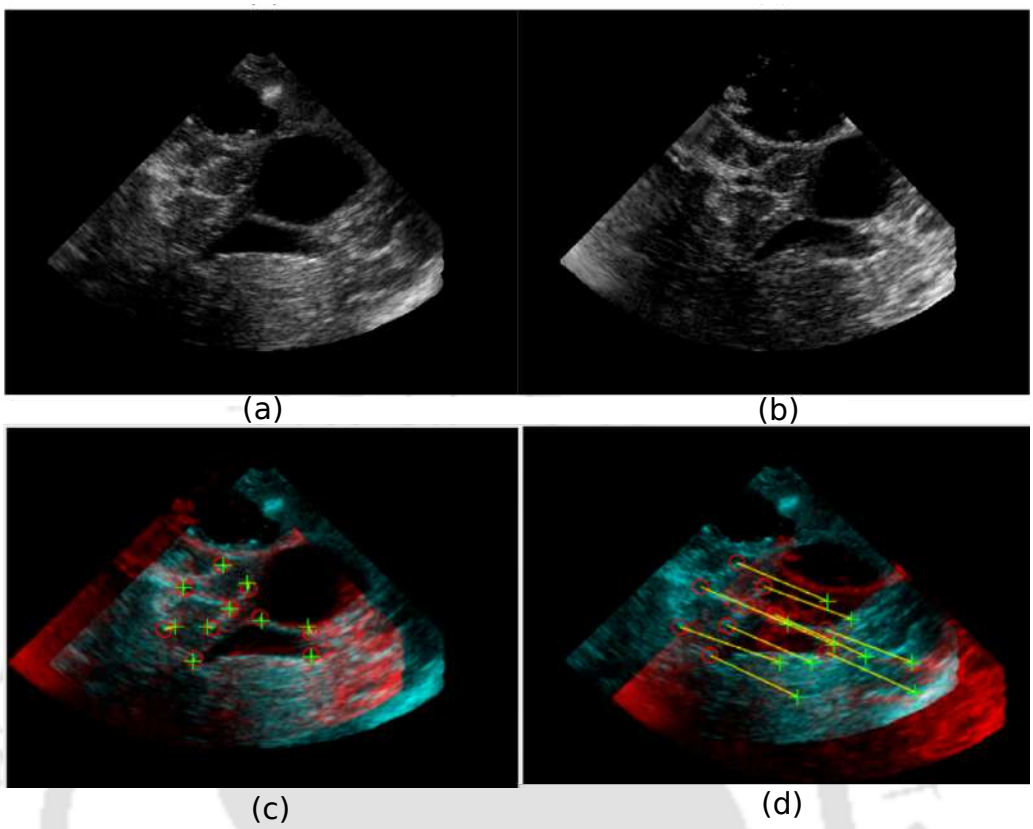


Fig. 3.13: (a) and (b) are the pBUS and the corresponding iBUS images. Registered images are shown in a common frame with the corresponding tag points. (c) shows the results when the initial point is $a = b = d = e = 0.2$ and $c = f = -10$. (d) shows the registering images when the initial point is $a = b = c = d = e = f = 0$

The registration error of the proposed method, the methods of Machado *et al.*, and that of Canalini *et al.* are evaluated. The comparative mTREs for the RESECT and the BITE datasets are shown in Fig. 3.14.(a) and (b) respectively. Each figure contains three groups of box plots, and each group representing one method. Each group of boxes consists of two box plots, where the left box represent E_b and the right box represents E_a. The mean points of E_b and E_a are connected to highlight the nature of the progress of mTRE after registration. These figures show that the improvement of the mTRE is the best for the proposed method for both of the dataset. The parametric results of the mTRE for the BITE and the RESECT datasets are shown in Table 3.4 and Table 3.5 methods respectively.

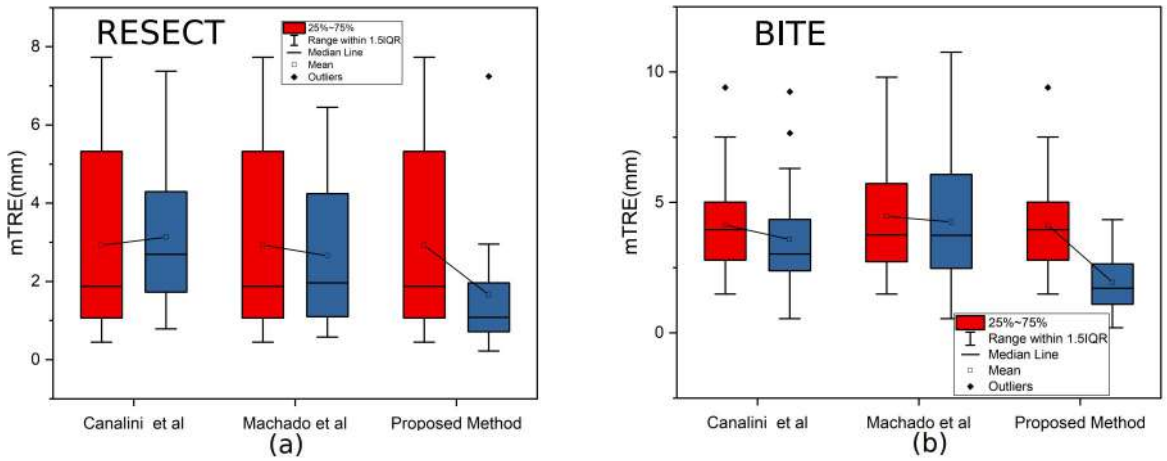


Fig. 3.14: (a) and (b) show respectively the comparative mTRE between the method of Canalini *et al.* [20], Machado *et al.* [108] and the proposed method for 29 pairs of images from the RESECT dataset and 14 pairs of images from the BITE dataset. The mean mTRE points before and after registration are connected for highlighting the changes in the performance of the registration method.

The results showed that the registration improved by the method of Canalini *et al.* is from 4.136 mm to 3.582 mm, which is slightly better than the method of Machado *et al.* The scenario reverses in the case of the RESECT dataset. For the RESECT dataset, the method of Canalini *et al.* increases the mTRE during registration. It registers the original images excluding the resection cavity region. The lack of the similarity between the registering image pairs increased. Capturing the similarity between the pBUS and iBUS images are challenging in this case. Machado *et al.* improved the average mTRE from 2.927 to 2.658 mm. The improvement in the average mTRE after registration is from 2.43 mm to 2.11 mm for the method of Canalini *et al.* It essentially infers that this method fails to register most of the image pairs. In both cases, the proposed PSO-based method significantly improves mTRE. For the BITE dataset, the average mTRE before registration was 4.136 mm that reduced to 1.931 mm. This improvement was noticed for the proposed method, and because of that, it assures better similarity between the registering image pairs.

BITE	Machado *et al.* (E_b)	Machado *et al.* (E_a)	Canalini *et al.* (E_b)	Canalini *et al.* (E_a)	Proposed method (E_b)	Proposed method (E_a)
mTRE (Mean)	4.467	4.247	4.136	3.582	4.136	**1.931**
mTRE (Std.Dev)	2.337	2.364	1.923	2.016	1.923	1.041

Table 3.4: Comparative mTREs for 14 pairs of images from the BITE dataset. The first and the second columns of each method represent respectively the error before and after registration.

RESECT	Machado *et al.* (E_b)	Machado *et al.* (E_a)	Canalini *et al.* (E_b)	Canalini *et al.* (E_a)	Proposed method (E_b)	Proposed method (E_a)
mTRE (Mean)	2.927	2.658	2.927	3.127	2.927	**1.654**
mTRE (Std.Dev)	2.434	2.01	2.434	2.064	2.434	1.708

Table 3.5: Comparative mTREs for 29 pairs of images from the RESECT dataset. The first and the second columns of each method represent respectively the error before and after registration

Fig. 3.15.(a1) and (a2) show an image pair having good visual similarity, Fig. 3.15.(b1) and (b2) show an image pair having moderate similarity, and Fig. 3.15.(c1) and (c2) show an impair with poor similarity and this registration is comparatively difficult compared to the other pairs. The third row of the figure shows overlapped pBUS and iBUS images before registration. Common tag-points are marked on pBUS and iBUS with '+' and 'o' symbols respectively. Common tag-points are connected with a straight line, that shows the initial distances between the tag-points. These distances are higher before registration; thus the value of E_b is high. The fourth to sixth rows of this figure show the same image pairs and the tag points after the registration by the proposed method with PSO as the optimization. It can be clearly seen that E_a reduces significantly by the proposed method compared to the two other methods, which are shown in the fourth, fifth, and sixth row of the Fig. 3.15.

3.4 Limitation and future scope

The thresholding of AIG extracts the edge-rich regions from the registering image pairs thereby increasing the similarity between registering pBUS and iBUS image pairs. For the 46 pairs of images included in the test, the results showed that the proposed method could reduce the registration errors in comparison to the previous studies. Fig. 3.16 shows a summary of the success rates for both datasets. In the work of Canalini *et al.* and the Machado *et al.*, post registration error were within 1.5 mm to 2.75 mm. We define the registration of a pBUS and iBUS image pair successful if its post-registration error is within 2.5mm. We define the success rate as a percentage of successful pairs of images among all the registering image pairs in a group (easy, medium or difficult). In this experiment, only 6 pairs of images (2 pairs from the RESECT dataset and 4 pairs from the BITE dataset) were found easy, and 40 image pairs (25

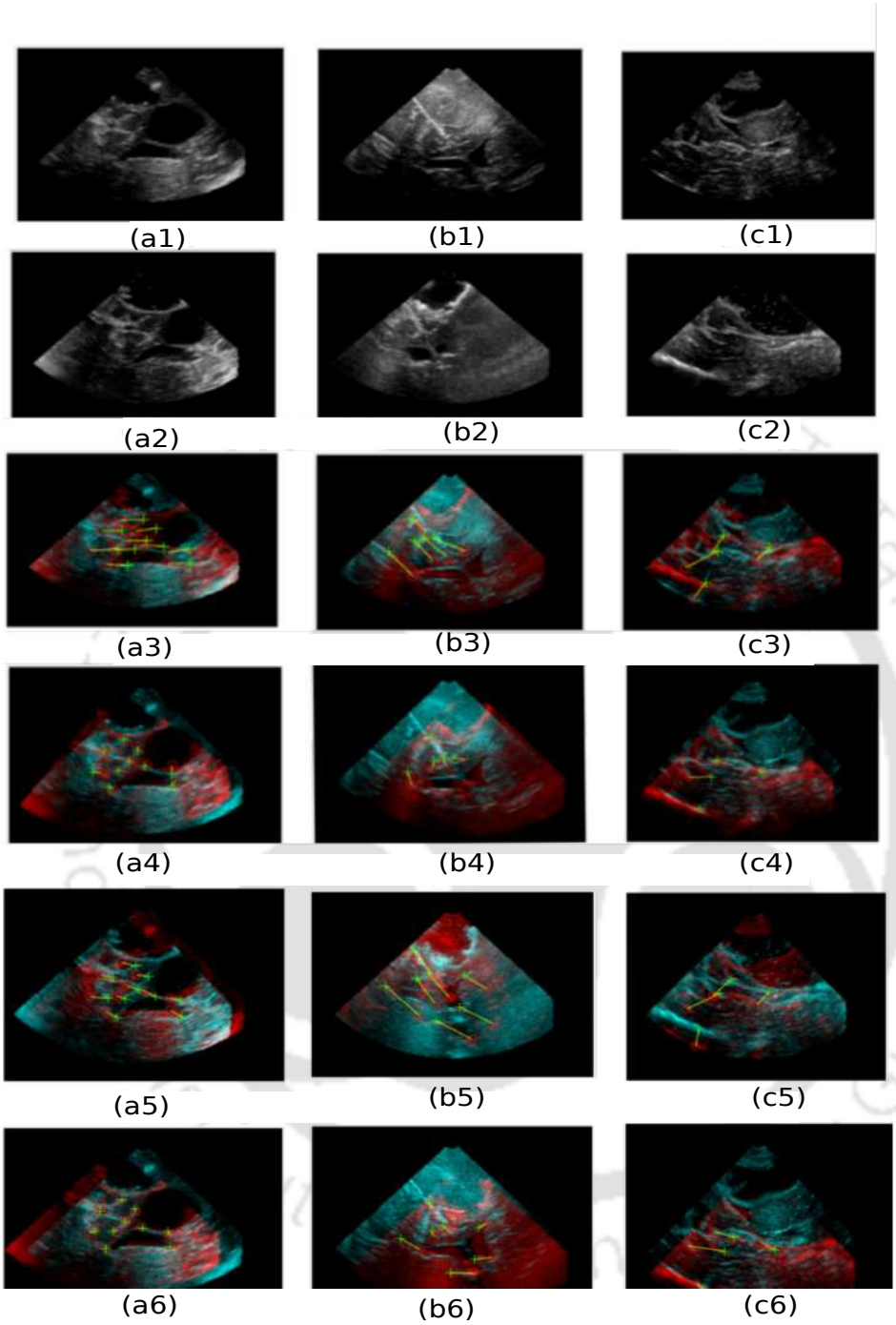


Fig. 3.15: (a1),(b1), and (c1) show three pBUS images, and corresponding iBUS images are shown in (a2),(b2), and (c2) respectively. Respective pBUS and iBUS image pairs before registration are shown in common frame in (a3), (b3) and (c3). Corresponding tag points are denoted in the image pairs with '+'and 'o'. Similarly, the resulting image pairs after registration using the proposed method are shown in (a4),(b4), and (c4). The tag point positions are changed in the iBUS images as per the applied T. (a5),(b5), and (c5) show the resulting images and tag-points after registration using the SIFT feature-based method by Machado *et al.* The images in (a6),(b6), and (c6) are the resulting images with the corresponding tag points resulted by the approach of Canalini *et al.*

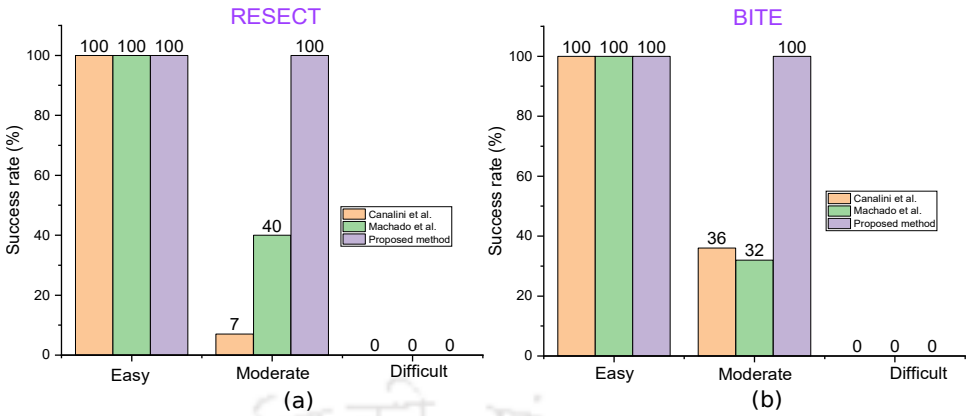


Fig. 3.16: Comparative success rate (in percentage) of registration between three methods for (a) BITE dataset and (b) RESECT dataset.

image pairs from the BITE and 15 image pairs from the RESECT datasets) were found to be moderate. Out of those 40 image pairs, 15 image pairs could be registered by the method of Machado *et al.*, but failed by the method of Canalini *et al.*. It was found that the proposed method outperformed other methods, but the positive results were limited to 46 images out of 68 image pairs, which should be improved further.

3.5 Conclusions

This chapter presented an efficient method for the rigid registration of pBUS and iBUS images. The constraint in this registration is the poor similarity between the registering image pairs due to the presence of various artefacts. This method performs thresholding of absolute intensity gradient of the registering image pairs for improving the similarity by extracting the edge-rich regions from the registering pBUS and iBUS images. Registration is done on the edge-rich regions of pBUS and iBUS images. This chapter compares the performance of three different heuristic optimization algorithms, namely SA, BBO, and PSO, and PSO was found to yield the best registration results. Results were also compared with other state-of-the-art methods. The proposed method showed better registration even in the presence of various artifacts.

4

Registration of pBUS and iBUS images using patch-based segmentation of hyper-echoic regions

Hyper-echoic (HE) regions appear comparatively prominent in both the pBUS and iBUS images. The segmentation of these regions is an effective means to increase the similarity between them. This chapter first proposes a patch-based strategy for segmenting the HE regions. The patchwise average intensity is thresholded using a modified NI-black's approach. The boundary regions are refined with an iterative boundary correction stage with gradually reduced patch sizes. A cubic B-spline contour smoothing technique is applied to obtain smooth boundary between the HE and NHE regions. The efficiency of the proposed method and its convergence speed are compared with state-of-the-art level-set-based methods and a machine learning-based method. The results show that the proposed segmentation method outperforms its competing methods. The HE regions of the pBUS and the iBUS images are registered using gradient-based and heuristic optimization techniques.

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Introduction

The previous chapter presented a method for the patch-based rigid registration of pBUS and iBUS images. It emphasized on the extraction of edge-rich regions from the pBUS-iBUS image pairs. For obtaining these regions, the patch-based AIG was thresholded by a predetermined threshold. This method is successful where registering image pairs contain sufficient sharp regions. Unfortunately, the iBUS images include various artefacts due to the presence of saline water and blood clotting agents in the RC [123, 124]. For many image pairs, the complete edge profiles are lost because of the change in acoustic impedance, and pixel-to-pixel correlation between the pBUS and the iBUS images reduces. For obtaining a better match, we need to find the common regions with high correlation. Fig. 4.1 shows an iBUS image that contains regions with different types of echogenicity. The hyper-echoic (HE) regions like LF and CP, and the anechoic regions like a cyst and ventricles appear prominently in this image. The image also shows comet tail artefacts that adds undesired echogenicity of the comet's tail. Similarly, the hyper-echogenicity, produced due to blood clotting agents in the RC is also seen. Some parts of the ventricle in this iBUS image get hidden due to the derived echogenicity of the commet's tail. Artifacts and shadows reduce the local information. In such a scenario, the regions with high echogenicity can be used for image registration. Thus, the segmentation of the HE regions becomes reasonable.

Based on the echogenicity, the US image regions can be classified into three types viz. a) HE or very bright region, b) hypo-echoic or less bright region (example: caudate nucleus and cerebrum) and c) anechoic or visually black regions (example: all four ventricles, the RC, and cyst.) [123]. We consider both the hypo-echoic and anechoic regions together as non-hyper-echoic (NHE).

For achieving successful registration, segmented HE regions from both the images can be used. Canalini *et al.* [20] followed a similar approach to the registration of pBUS and iBUS images by segmenting two HE structures sulcus and longitudinal fissure. The common HE regions possess higher echogenicity and thus act as relatively better landmarks for successful registration of the iBUS and the pBUS images. Hence, segmentation of the HE regions can be treated as a preprocessing step in the registration process. This chapter first describes an automatic, fast method for the segmentation of HE regions from pBUS and iBUS images. After extraction of the HE regions from the image pairs, registration is performed and its performance

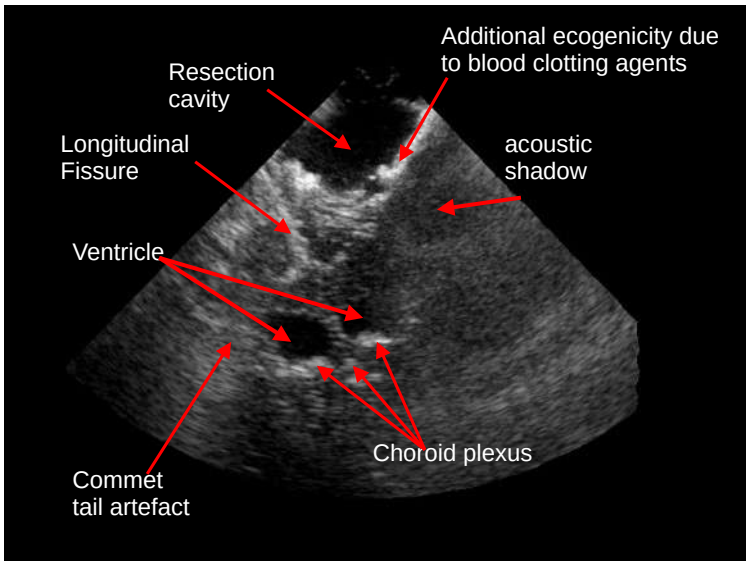


Fig. 4.1: Different echogenic regions and artefacts present in the iBUS image

is compared with the previous approach discussed in Chapter 3, and other registration methods.

The organization of the rest of the chapter is as follows. Section 4.1 provides a detailed literature survey on the segmentation methods of ultrasound images. Section 4.2.1 describes the segmentation method for extraction of the HE regions. Section 4.2 presents the different steps of the proposed segmentation approach. Results of this approach are presented in the Section 4.2.5. Section 4.3 discusses the registration of segmented pBUS and iBUS images. Finally, the conclusion is drawn at Section 4.7.

4.1 Approaches to the segmentation of ultrasound images

The approaches to the segmentation of US images [143] include Markov random field (MRF) [201, 62], region growing [75, 110], support vector machines (SVM) [214, 200], multidimensional space frequency transformation [136], watershed-based segmentation [82], level-set based segmentation [87, 24, 28, 98, 90], and deep learning-based approaches [3, 74, 193, 207, 208]. Local intensity inhomogeneity and large variance are inherent nature of a US image due to the speckle present in the images. Traditional segmentation techniques [69] like pixel-based thresholding, multilevel thresholding, region growing and watershed segmentation fail to segment such noisy

BUS images. Researches are being carried out in this field for the past three decades, and some of the relevant and effective methods are covered in this section.

For noisy image segmentation, the MRF and the AC based approaches [118, 140]were found to be useful. Martin and Alberola [118] proposed a probabilistic Bayesian approach for the segmentation kidney in US abdomen image. It is a semi-automatic approach that needs an initial contour to be drawn by the user in a 2D slice image. The final contour is obtained using restriction of contour smoothness and a Markov field of deformations imposed by the prior. The likelihood function is a conditional probability of the data based on the position of the contour. Nirmala and Sridevi [140] proposed an US image segmentation method for detecting prenatal ventricular septal defect using a Markov random field model. The majority of homogeneous nearby pixels with comparable attributes or features, such as intensity, color, and texture, are visualized in an image using the MRF model-based segmentation approach. In order to complete the segmentation process, the MRF model collects those comparable features from the collection of pixel intensities. Wu and Lu [200] proposed an integrated approach to US image segmentation based on iterative sliced inverse regression (ISIR) from the perspective of statistical classification for image pixels to various spatial structures. For dimensionality reduction, average intensity of the square patches was considered as a feature. For extracting better edges and their orientations, each local square patch is convolved with a bank of Gabor filters with different frequencies and orientations. The initial group labeling was done by K-Means clustering, which is further fine-tuned by a supervised learning approach using SVM.

Recently, deep neural networks(DNNs) are demonstrated to be very popular tool for US image segmentation. Different approaches were presented recently, which are either region-specific or they address the segmentation of the regions affected by some artifacts. Deep learning was applied for the segmentation of many anatomical regions like prostate [197, 86, 214], ovary and its follicles [99], thyroid nodule [105, 218], breast tumor [74, 193, 207], kidney [211, 210], artery [48], heart [134, 80], fetus [121, 156, 204] etc. The popular architectures are U-Net [134, 218], [171], CR-Unet [99], 3D ResNeXt [197, 203], Dual Path U-Net (DPU-Net) [208], DenseNets [48] and the generative adversarial network (GAN) [74]. Most of the models are fully supervised and require radiologists to annotate large-scale training data which is a rather labor-intensive and time-consuming task. To cope with this issue of the radiology experts, many researchers attempted for implementing the segmentation task in a semi-supervised manner.

These semi-supervised approaches need limited annotated images and a large amount of the unannotated data can be utilized. Meng *et al.* [121] applied a semi-supervised approach to find the acoustic shadow in fetal US images. Some other semi-supervised approaches are reported in [114, 61], which used fewer numbers of annotated images.

Applying deep learning for 2D brain US image segmentation is difficult because of non-availability of the annotated 2D dataset of segmented images. Supervised DNN-based segmentation requires thousands of annotated training images. To the best of our knowledge, there are only two datasets BITE [123] and RESECT [202] available on brain US images. There are no ground truth annotated images available in these datasets for the segmentation of HE regions. Canalini *et al.* [20] annotated gyri and sulci regions from these datasets and adopted a U-Net [171] architecture for the segmentation of RC. Because the RC region in iBUS image is dissimilar to the corresponding region in the pBUS region, this region is segmented and excluded from the registering image pairs. This method was able to register pBUS and iBUS images based on those segmented HE sections.

The performance of the pixel-based approaches for US image segmentation deteriorates due to its intensity inhomogeneity, speckle, and various artifacts. Achanta *et al.* [2] introduced the concept of simple linear iterative clustering(SLIC) superpixels. It is a fast method for clustering group of pixels adopting a modified k-means clustering algorithm. It combines color and spatial proximity. Superpixels are tiny clusters of neighboring pixels having similar textures. Superpixel-based segmentation methods show better noise immunity compared to the pixel-based approaches. This was first applied in medical image segmentation by Schwier *et al.* [177] for spine and vertebrae detection in CT images. Nitsch *et al.* [141] proposed a superpixel-based method for BUS image segmentation using a semi-supervised random-forest classifier. A similar superpixel-based semi-automatic approach was proposed by Daoud *et al.* [44], which applied an SVM classifier for breast tumor detection. The perfection of the boundary delineation by the superpixel-based method depends on the ratio of $K = \frac{N}{S}$, where N and S are respectively the number of pixels and superpixels in the image. With the increase in the value of S, performance improves but it also increases the complexity. There is no standard guideline for selecting the optimum value of K and it is chosen within a range of 1 to 40 [177]. The output boundaries produced by these methods are uneven and suffer from the slope discontinuity, because superpixels do not have any regular geometric shape. The slow nature and non-unique

boundary detection of these methods limit their application, where fast image segmentation is desired. Some post-processing steps may be required here to improve the boundaries that further increases the complexity.

An important framework in image segmentation is the level-set-based approaches [24, 28, 64, 12, 182, 103, 90]. Level-set-based methods have wide applications in the segmentation of medical images of various modalities. They are used for boundary curve evolution in segmentation whereas, the changes in the segmented region is guided by an energy minimization framework. The older methods like the active contour(AC) [87], and the geodesic active contour (GAC) [24] do not perform well in brain US image segmentation. However, the advanced approaches like distance regularized level-set evolution (DRLSE) [98], and weighted level-set evolution WLSE [90] reported improved results. Different level-set-based approaches applied to medical image segmentation are described in the next section.

4.1.1 Level-set-based segmentation methods

Active contour or *snake* is an image segmentation approach introduced by Kass *et al.* [87], which adopts energy minimization framework with respect to a boundary contour. The energy is derived from the curve parameters and the intensity profile of the image at the neighborhood of the contour. At an optimum-energy condition, the contour divided the image into inside and outside regions.

Let, $\mathbf{V}$ be an an image defined on $\Omega \subset \mathbb{R}^2$. A parametric boundary curve $C(l) : [0, 1] \rightarrow \mathbb{R}^2$ in $\mathbf{V}$ can be expressed as

$$\inf_{\mathbf{C}} J(\mathbf{C}), \text{ where, } J(\mathbf{C}) \text{ is an energy function.}$$

For the classical AC method,

$$J(\mathbf{C}) = \alpha \int_0^1 \left|C'(l)\right|^2 dl + \beta \int_0^1 |C''(l)|\, dl - \gamma \int_0^1 |\nabla V(C(l))|^2) dl \tag{4.1}$$

where α, β and γ are the constants used to control the weighs of three different energy terms. The first and the second terms in (4.1) refer jointly as the measure of smoothness of the boundary. It is also referred to as the stored energy or the internal energy of the snake. The

third term in (4.1) refers to the external energy. The basic purpose served by the minimization of (4.1) is to locate the contour over the maxima of $|\nabla V|$ preserving the smoothness of the curve.

Level-set is a curve evolution strategy was first formulated by Osher and Sethian [146]. Application of the classical level-set includes evolving interfaces in computation geometry, fluid mechanics, material science, and computer vision and image processing [178]. A level-set curve is a closed contour defined by $\mathbf{C} = \{x, y\}|\phi(x, y) = c\}$, where c is a constant and ϕ is a Lipschitz function. The evolution of a let-set curve initiates from the zero-level-set-curve $\mathbf{C}$, where $\mathbf{C} = \{(x, y)|\phi(x, y) = 0\}$. $\phi(t, x, y)$ is defined as the evolution of the contour at time t. This temporal evolution of the contour that drives the snake towards the normal direction, can be expressed as,

$$\frac{\partial \phi}{\partial t} = \mathbf{F}\,|\nabla \phi| \tag{4.2}$$

where $\mathbf{F} = div\left(\frac{\nabla \phi}{|\nabla \phi|}\right)$ controls the speed of evolution. Hence Equation (4.2) can be expressed as,

$$\frac{\partial \phi}{\partial t} = |\nabla \phi|\, div\left(\frac{\nabla \phi}{|\nabla \phi|}\right) \tag{4.3}$$

$\mathbf{F}$ is known as the force function. In the classical level-set, $\mathbf{F}$ is the curvature of $\mathbf{C}$ that has no relation to the topology of the surface, and its values change freely as per Equation 4.3.
In a level-set-based image segmentation problem, ϕ is defined as a distance function having positive values inside and negative values outside. The segmentation progresses like an evolving region surrounded by the zero-level-set curve. The evolution starts from an initially marked zero-level-set curve $\mathbf{C}$. Unlike the classical level-set, it has to stop at the boundary of the curve, and therefore, it is combined with the active contour to use in a practical image segmentation problem. The term $\mathbf{F}$ in (4.2) becomes dependent on the image intensity in such a manner that it has a high value at a homogeneous region, and zero value at the region boundaries.

The primary purpose of the AC approaches [87, 195, 28, 12, 64, 182, 149, 101, 24] is to locate a topology-independent and regular-shaped boundary between different regions of an image. The AC-based methods involve two steps in each iteration, namely the energy minimization and the curve flow. The segmentation procedure starts from a user-defined boundary which evolves until total energy gets minimized. When the total energy is minimized, the curve reaches the optimal boundary and stops further evolution.

Caselles *et al.* [24] formulated GAC model based on the mean curvature motion equation. It finds an edge map of the image as given by

$$g(|\nabla V(x,y)|) = \frac{1}{(1+|\nabla G_\sigma * V(x,y)|^p)} \tag{4.4}$$

where $*$ denotes convolution and G_σ is a Gaussian filter with standard deviation σ and $p = 1$ or 2. It filters noise present in $V(x,y)$. $g(|\nabla \mathbf{V}|)$ is small at an edge and has higher values in homogeneous regions. $g(|\nabla \mathbf{V}(x,y)|)$ lies in between 0 and 1.

The curve flow in GAC involves the level-set approach for boundary curve evolution. The estimation of $C''(l)$ at each iteration is computationally expensive. For avoiding this, GAC [24] uses a simpler objective function as

$$\inf_C J(C) = 2\int_0^1 |C'(l)|\, g(|\nabla \mathbf{V}(C(l))|)dl \tag{4.5}$$

The level-set curve evolves along the normal direction and its jump depends on a force function. GAC uses

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| \left(div\left(g(|\nabla V|)\frac{\nabla \phi}{|\nabla \phi|}\right) + v g(|\nabla \phi|\right) \tag{4.6}$$

where, $v \geq 0$, is a constant. During curve evolution, when the level-set curve reaches at the desired boundary, $g(|\nabla V_0|$ becomes very small. The use of $g(|\nabla V)|$ in level-set curve evolution shows that the energy at the boundary is minimum, that forces the curve evolution to stop.

Both the classical AC and GAC emphasize on the smoothness of the evolving curve and the length of the curve was included as a regularization term into the objective function. Other level-set-based algorithms such as DRLSE by Li *et al.* [98] and WLSE by Khadidos *et al.* [90] also included a smoothness parameter in the energy function. Chan and Vese [28] introduced ACWE which minimizes the summation of intra-class variances for the segmentation of the background and the foreground from a composite image. This algorithm also incorporated the length of the evolving curve and the area of the foreground region into the objective function as a regular. Li *et al.* [97] addressed the issues related to the intensity heterogeneity in various images and added a concept of local intensity clustering assuming that the intensity variation within a cluster is very less. The intensity clusters and their respective cluster centers are

obtained using k-means clustering. This method formulated a weighted average of intensity variances within the clusters as the energy function. The weights were chosen from a Gaussian kernel as a function of the distance between a pixel coordinate and the center of a neighborhood. A standard level-set approach was used for its curve evolution. Another work, proposed by Li *et al.* [98] used a k-means clustering integrating it with a two-stage level-set curve evolution for minimizing the sum of intra-class variances. This method also considered the length of the boundary curve as a regularizer.

The smoothness regularization is necessary in US image segmentation because it provides a natural shape to the contours and overcomes the effect of local spatial noise. Smoothness can be imposed by other methods. Bernard *et al.* [12] introduced the B-spline-based level set where the evolving curve is represented as a linear combination of B-spline basis functions. Energy minimization is done directly with respect to the B-spline coefficients. B-spline has the advantage that the higher order coefficients can easily be found by the convolution of two lower order B-spline functions and it also has a connate smoothness property which benefits the algorithm. These level-set-based techniques were found to be efficient for the segmentation of the various medical imaging modalities like MR, CT and US. Liu *et al.* [103] proposed a fast level-set method that finds the initial boundary using mean-shift clustering. A level-set-based curve evolution is followed along with the energy function minimization. This method is robust to intensity heterogeneity, which is common in US images. Generally, in level-set curve evolution, a single boundary curve and its closed neighborhood regions are used for energy estimation. The recent WLSE method proposed ka Khadidos *et al.* [90] modified DRLSE and calculated the energy function as a weighted average of the energies estimated from multiple parallel boundaries and their close neighborhoods. The automatic weight estimation involves gradient vector flow (GVF) of the image and the normal vectors of the initially chosen mask. This method was proven to be better for medical image segmentation, especially in the presence of the weak edges.

The level-set process involves the numerical solution of a partial differential equation(PDE), which makes the whole process slow. Shi-Karl [182] proposed a fast level-set-based AC technique which did not need any solution of a PDE. This method evolves the boundary curve within a spatial narrow band of the neighboring pixels. Like GAC and AC, the energy function is curvature dependent, that helps to obtain a regular-shaped boundary contour. This method is

relatively faster compared to GAC, ACWE, DRLSE, and WLSE.

4.1.2 Selection of energy function

The classical AC [87] model includes the spatial curve gradient $\mathbf{C}'(q)$, the curvature $\mathbf{C}''(q)$ and the intensity gradient $|\nabla(V_0(C(q)))|$ in the energy function as shown in Equation (4.1) and (4.5). This method was proven to be effective for the segmentation of photographic images when they have a homogeneous intensity profile and less noise. The large complexities for estimating the curvature and solving the partial differential equation (4.3) at every iteration are the limitations of this method. The GAC [24] model avoided the estimation of curvature $C''(q)$ and used the following energy function

$$J_T = \int_0^1 \alpha |C'(q)|^2 dq - \gamma \int_0^1 |\nabla V(C(q))| dq \tag{4.7}$$

Noise has an impact on the performance of both AC and GAC due to the $|\nabla V(v(q))|$ term. Mumford and Shah [137] presented three energy functionals that can be used to solve the image segmentation difficulties. The original formulation is given by

$$J_T = \int_X (\mathbf{U} - \mathbf{V})^2 dxdy + \alpha \int_0^1 |\nabla q|^2 dxdy + \gamma |\mathbf{C}| \tag{4.8}$$

Except at the border $\mathbf{C}$, $\mathbf{U}$ is a differentiable function assumed to lie within the planner region $\mathbf{X}$. α and γ are the real positive constants. $|\mathbf{C}|$ is the length of the contour $\mathbf{C}$. They proposed another formulation of J_T that incorporates all of the local intensity variances of segments $\mathbb{S}_i$s and can be written as,

$$J_T = \sum_i \int \int_{\mathbb{S}_i} (\mathbf{V} - \mu_i)^2 dxdy + \gamma |\mathbf{C}| \tag{4.9}$$

where γ is a weight of the length constraint of the boundary $\mathbf{C}$. Noise affects the energy functional in (4.5) and (4.8). In contrast to these models, Equation (4.9) estimates the local variances of each disjoint region without using $|\nabla C|$.

Chan and Vese proposed ACWE [28] that minimizes intra-class intensity variance considering the length of the contour, and the area within the contour as constraints. The formulation

of J_T given by Chan and Vese is

$$J_T = \alpha\,|\mathbf{C}| + \beta \underbrace{\frac{1}{2}\oint_{\mathbf{C}} (xdy - ydx)}_{\text{Area inside } u\text{(Green's theorem)}} + \lambda_1 \int_{(x,y)\epsilon\mathbb{S}_i} |V(x,y) - c_1|^2 dxdy + \lambda_2 \int_{(x,y)\epsilon)\mathbb{S}_o} |V(x,y) - c_2|^2 dxdy \tag{4.10}$$

where λ_1 and λ_2 are the constants, $\mathbf{C}$ represents the union of all the boundaries of disjoint regions $\mathbf{S}_1, \mathbf{S}_2$..., etc. $\mathbf{S}_i(q)$ and $\mathbf{S}_o(q)$ are the inside and outside region of $\mathbf{C}$. Therefore,

$$\mathbf{S} = \mathbf{S}_i \bigcup \mathbf{S}_o$$

The strategy of taking mean value of a region reduces the effective noise in the boundary.

A similar region-based approach was proposed by Lankton and Tanenbum [95], that exploited the statistics of close circular neighborhood of the boundary $B(x,y), (x,y) \in \mathbf{C}$ given by,

$$B(x,y) = \begin{cases} 1\, for\ \|x-y\|^2 < r \\ 0\, otherwise \end{cases} \tag{4.11}$$

The energy function considered is

$$J_T = \int_{(x,y)\epsilon\mathbb{S}_o} \delta(\phi(x,y)) \int_{\mathbb{S}_i} B(x,y) F(V(x,y), \phi(x,y)) dxdy + \lambda\delta(\phi(x,y)) \|\nabla\phi(x,y)\| dxdy \tag{4.12}$$

where, $\phi(x,y)$ represents the level-set distance function. Its value if positive inside $\mathbf{C}$ and negative outside. $F(V(x,y), \phi(x,y))$ is the generic internal energy assessing the local adherence to a particular model along the contour, and the boundary region is approximated with a parametric form of the impulse function. Any of the following three energy models can be used as J_T.

- *Uniform energy* (J_{UM})

$$J_{UM} = \int_{(x,y)\epsilon S_i} H(x,y) \left(V(x,y) - \overline{V}_{\mathbb{S}_i}\right)^2 dxdy + \int_{(x,y)\epsilon\mathbb{S}_o} (1 - H(x,y)) \left(V(x,y) - \overline{V}_{\mathbb{S}_o}\right)^2 dxdy \tag{4.13}$$

where, H represents a 2D Heavieside function. Global mean intensities for the inside region ($\overline{\mathbf{V}}_{\mathbb{S}_i}$) and for the outside region ($\overline{\mathbf{V}}_{\bar{\mathbb{S}}_o}$) are respectively given by,

$$\begin{gathered} \overline{V}_{\mathbb{S}_i} = \frac{\int_{\mathbb{S}_i} H(x,y)V(x,y)dxdy}{\int_{\mathbb{S}_i} H(x,y)dxdy} \\ and \\ \overline{V}_{\mathbb{S}_o} = \frac{\int_{\mathbb{S}_o} (1-H(x,y))V(x,y)dxdy}{\int_{\mathbb{S}_o} (1-H(x,y))dxdy} \end{gathered} \tag{4.14}$$

- *Mean separating energy* (J_{MS}):

$$J_{MS} = \int_{x,y \epsilon S} \left(\overline{V}_{\mathbb{S}_i} - \overline{V}_{\mathbb{S}_o}\right)^2 dxdy \tag{4.15}$$

This energy if formulated on the assumption that the foreground and the background regions should have maximally separate mean intensities.

- *Histogram separation energy* (J_{HS}):

$$J_{HS} = \int_z \sqrt{P_{\mathbb{S}_i}(z).P_{\mathbb{S}_o}(z)}dz \tag{4.16}$$

where $P_{\mathbb{S}_i}(z)$ and $P_{\mathbb{S}_o}(z)$ are the smoothed intensity histograms estimated on the inside and outside regions respectively.

Apart from the aforementioned energy functions, Li.*et al.* [97] proposed a region-based ACWE model. For obtaining a smooth boundary, the length of the $\phi(x, y) = 0$ curve was incorporated in the energy function as a regularizer. Bernard *et al.* [12] modified ACWE framework by considering the evolving level-set boundary as a B-spline curve. The B-spline parameters are optimized during energy optimization, resulting an accurate and a smooth boundary.

In all of the said methods like AC [87], GAC [24], ACWE [28], and the method of Lankton and Tanenbum's [95], the contour evolves along the normal direction to it. The level-set curve in WLSE approach [90] propagates towards the direction of the gradient vector flow. The length and the area terms were also included in determining the desired smooth boundaries.

The energy function J_T in WLSE is given by,

$$J_T = \frac{\mu}{2}\int_{\mathbb{S}} \left(|\nabla\phi| - 1\right)^2 dxdy + (1-\omega(\phi,k)) \underbrace{\int_{\mathbb{S}} g(\delta(\phi)\,|\nabla\phi|)dxdy}_{\text{Length of level-set curve } \phi=0} + \omega(\phi,k) \underbrace{\int_{\mathbb{S}} g(H(\phi))dxdy}_{\text{Area } \phi \geq 0} \tag{4.17}$$

where $\omega(\phi,k)$represents the weight for the k^{th} level-set curve such that $\sum_{\forall k}\omega(\phi,k) = 1$ and g is the edge function as described in Equation (4.4).

Another crucial issue is selecting the values of α, β, and γ in (4.1). These are needed to be varied for different images and modalities. ACWE [28] and the method by Bernard *et al.* [12] involved intra-class variance in every iteration, that has a significant computational burden. A large step size or the magnitude of the force function in (4.6) can make the evolution faster, which result into discontinuous boundaries. For avoiding the discontinuities in evolving boundaries, a smaller step size of the force function is needed that results into slow evolution of the level-set curve.

4.2 Proposed segmentation method

In this section, we propose a patch-based approach for the segmentation of the HE regions in a BUS image. As segmentation is a prework to the registration of pBUS and iBUS images, the method should be fast and reliable. The proposed method is automatic, uses simple calculations, and does not need any solution of PDE or any iterative energy minimization, which ensures faster convergence to the actual boundaries between the HE and the NHE regions. This process uses a region-driven patch-based approach, which segments the regions patch-wise based on the average patch intensity.

The proposed work consists of six processing steps on a brain US image V as shown in the block diagram of Fig. 4.2. Generally, a US image occupies a fan shaped FOV region in a rectangular image, and the rest of the part of the image is redundant. The segmentation is needed to be done on that fan shaped FOV region only. So, the first task is to extract the FOV region. This step was discussed in Section 3.2.1. After the separation of the FOV, this region is grossly divided into HE and NHE regions applying patch-based thresholding. The intensity threshold is estimated by combining the global mean and the standard deviation of

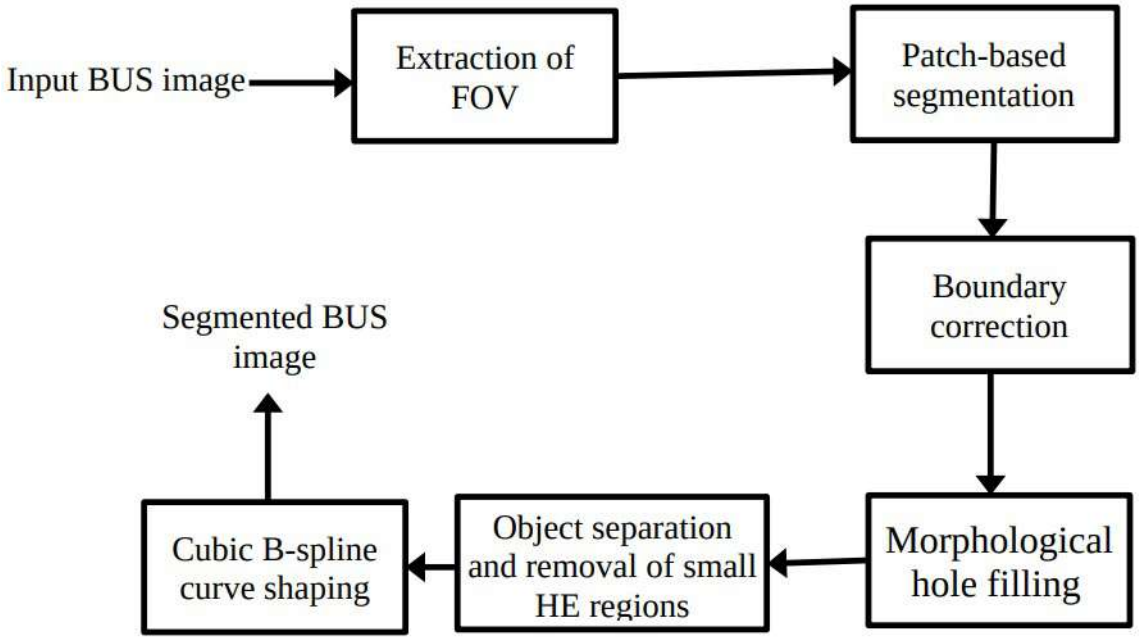


Fig. 4.2: Block diagram of the proposed method for the segmentation of HE regions from BUS images.

intensity in the extracted FOV region. A global threshold is applied on the image patch-wise for separating the HE and the NHE regions, and thus, a preliminary segmentation is obtained. For reducing the blockieness of the obtained boundaries, an iterative boundary correction step is followed. This step removes the inter-region ambiguities due to noise and refines the shapes of the boundary contours. Several small HE and NHE regions appear due to speckle and artefacts, they are removed using morphological hole filling. The boundary is further improved by a cubic B-spline-based contour shaping technique. Detailed descriptions of all the steps employed are given in the following subsections.

4.2.1 Patch-based segmentation

Intensity thresholding is the most basic approach to image segmentation. This technique can be applied pixelwise to segregate them into suitable segments. But for noisy images, pixelwise method mostly fails. The proposed segmentation method employs a patch-based average thresholding scheme which performs better than the pixel-based methods. We brief the related discussion in two parts:

4.2.1.1 Pixel-based thresholding and its limitations

The easiest approach to segment the foreground and the background regions of a photographic image is the intensity thresholding. Among the various thresholding approaches, Otsu's method [69] is the simplest. It is a pixel-based technique for two-class segmentation, that minimizes the intra-class variance. This method is successful for images having a bimodal histogram. Unfortunately, the histogram of an US image is not bimodal and Otsu's method fails to segregate the HE and NHE regions. Fig. 4.3(a)shows a brain US image from the BITE dataset [123]. This image consists of a HE grade-IV glioblastoma tumor and an anechoic cyst. The resulting binary map after the segmentation using Otsu's method is shown in Fig. 4.3.(b). The boundaries of HE regions are not distinct, which is a significant drawback of this method. US images suffer from intensity heterogeneity due to the presence of speckle. A pixel-based segmentation approach fails to segment US images because it compares the intensity pixel-wise with a predetermined threshold. Many pixels within an HE region have intensities lower than the threshold and therefore get excluded from the HE regions. A patch-based threshold comparison may be an effective solution to this problem. Kuo *et al.* [94] followed a multilevel threshold approach for brain ventricle segmentation in a mouse embryo image. For obtaining thresholds, a Gaussian mixture model (GMM) with two mixtures to the histogram of the preprocessed image was employed. The threshold was set at the intersection of the estimated Gaussian distributions. Unfortunately, the histograms of an ultrasound image does not always show a deep valley in the intensity axis. As a result, both the GMM and the Otsu's method fail in most cases.

Niblack [138] proposed a patch-based threshold estimation method for the segmentation of text images by combining the local mean $\mu(x, y)$ and standard deviation $\sigma(x, y)$ of a patch centered at (x, y) as

$$E_T(x, y) = \mu(x, y) + k\sigma(x, y) \tag{4.18}$$

where $E_T(x, y)$ is the threshold and k is a constant, chosen empirically as -0.2 for the best performance.

This method applies different thresholds for different patches. Let a patch be centered at (x, y) having a local patch threshold $E_T(x, y)$. All the pixels within the patch is compared with $E_T(x, y)$ and the pixels having intensity greater than E_T are included in the HE region. The

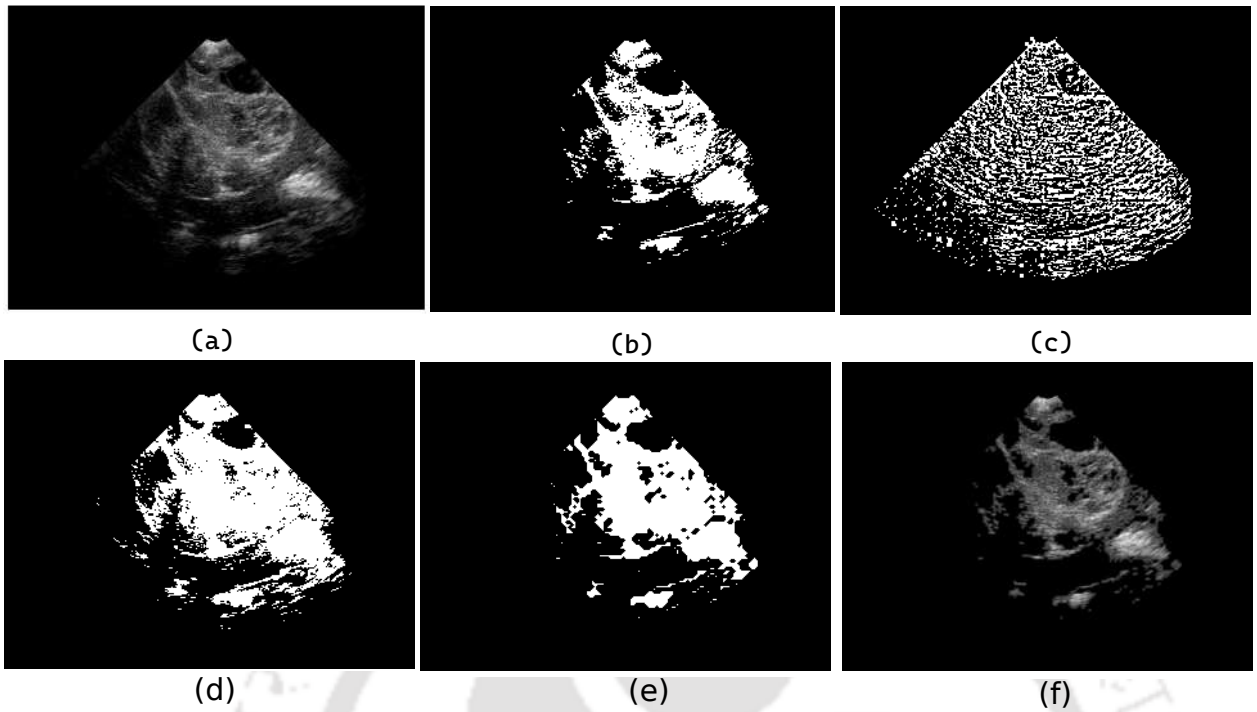


Fig. 4.3: (a) Original image chosen from BITE dataset [123]. (b) Segmentation using Otsu's method. (c) Segmentation using original Niblack's thresholding method. (d) Segmentation by pixel-based global thresholding using Niblack's formula($k_1 = -0.2$) (e) Patch-based segmentation using thresholding and $k_1 = 0.3$. (f) Segmented HE image using proposed patch-based thresholding. The image is generated by multiplying the binary mask in with the original image.

.

binary segmentation map obtained after the Niblack's thresholding is shown in Fig. 4.3.(c). Unfortunately, this image does not provide any regional information because of the local threshold estimation approach. We have first attempted the Niblack's method by replacing its local threshold with a global threshold obtained by using Equation (4.19) and considering $k_1 = -0.2$. The resulting binary image after the pixel-based comparison with the estimated threshold, is shown in Fig. 4.3.(d). This image provides a gross description of the HE regions present in the binary segmentation map. From the images, it can be found that this approach is superior to the local patch-based thresholding shown in Fig. 4.3.(c). However this method suffers from under-segmentation because of the low threshold value and results ambiguous boundaries. For avoiding the ambiguity on the boundary, the proposed method i) fine-tunes the global threshold under the Niblack framework and ii) applies patch-wise thresholding where the mean intensity of each local patch is compared with a global threshold E_T. The detail of the patch-based modification of the method is described below.

4.2.1.2 Patchwise classification of HE and NHE regions

Let the image plane Ω be divided into non-overlapping square patches $(\mathbf{P}_1, \mathbf{P}_2, ..., \mathbf{P}_K)$ of size $(1 + w_{max}) \times (1 + w_{max})$, where w_{max} is an integer power of 2. Suppose μ_k is the average intensity of $\mathbf{P}_k$. $\mathbf{P}_k$ is decided to be HE if $\mu_k > E_T$. E_T is estimated by the combination of the global mean (μ_g) and the global standard deviation (σ_g) of the pixels in the FOV Ω_F and given by

$$E_T = \mu_g + k_1 \sigma_g \tag{4.19}$$

where

$$\mu_g = \frac{1}{|\Omega_F|} \sum_{(x_i, y_i) \in \Omega_F} V(x_i, y_i), \tag{4.20}$$

$$\sigma_g = \sqrt{\frac{1}{|\Omega_F|} \sum_{\forall ((x_i, y_i) \in \Omega_F} (V(x_i, y_i) - \mu_g)^2} \tag{4.21}$$

and $|\Omega_F|$ is the number of pixels within the FOV Ω_F. The value of k_1 is varied gradually from $k_1 = 0$ with a small step of 0.1. For different values of k_1, HE regions in the corresponding output images were observed. Based on the subjective evaluation of the HE regions, the value of $k_1 = 0.3$ was found to be appropriate for both the BITE and the RESECT datasets. For patch-wise classification of the HE and the NHE regions, μ_k is found over local $\mathbf{P}_k$ as,

$$\mu_k = \frac{1}{(1 + w_{max})^2} \Sigma_{(x,y) \epsilon \mathbf{P}_k} V(x, y) \tag{4.22}$$

The patch $\mathbf{P}_k$ is classified as HE if μ_k exceeds E_T. This classification can be formally expressed using a binary segmentation map $\mathbf{S}$ as follows:

Compute μ_k using Equation (4.22) (4.23)

$For\ k = 1\ to\ K$ (4.24)

$If\ \mu_k > E_T\ \text{ then}$ (4.25)

$S(x,y) = 1,\ \forall (x,y) \epsilon \mathbf{P}_k$

$Else,\ S(x,y) = 0,\ \forall (x,y) \epsilon \mathbf{P}_k$

End (4.26)

The proposed thresholding scheme is examined on the image shown in Fig. 4.3.(a). The resulting binary map obtained after this step is shown in Fig. 4.3.(e). The corresponding HE image segments are shown in Fig. 4.3.(f). The segmentation shows less ambiguities on boundaries in Fig. 4.3.(e) compared with the Otsu's method [69] and Niblacks's method [138] as shown in Fig. 4.3.(c) and (d) respectively. Fig. 4.4.(a) shows another pBUS image. Fig. 4.4.(b) shows the corresponding **S** after the proposed thresholding with $w_{max} = 8$. This image provides clear segregation between the HE and NHE regions. Performing thresholding on non-overlapping patches reduces the computational time which is a highly desired goal of this work. Processing with a large patch size results in blocking artefacts on the boundaries. The blocking artefact is manifested as the rectangular step-like shapes near its boundaries. The value of w_{max} is chosen optimally as 8. Increasing w_{max}, lowers the computational complexity, but increases the blocking artefacts on the boundaries. The choice of large w_{max} increases the possibility of merging different regions together. On the other hand, if w_{max} is small, multiple regions may come from a single region resulting in ambiguous boundaries. It was found empirically that, $w_{max} = 8$ is a good choice. However, the blockieness of boundaries resulted from patch-based classification can be made smooth by applying an iterative boundary correction step as described below.

4.2.2 Boundary correction

The patch-based segmentation process selects pixels in groups in HE and NHE regions. Due to the large patch size, multiple regions may get merged together. This phenomenon can be observed in Fig. 4.4.(b). This figure shows the step like boundaries resulting from the large

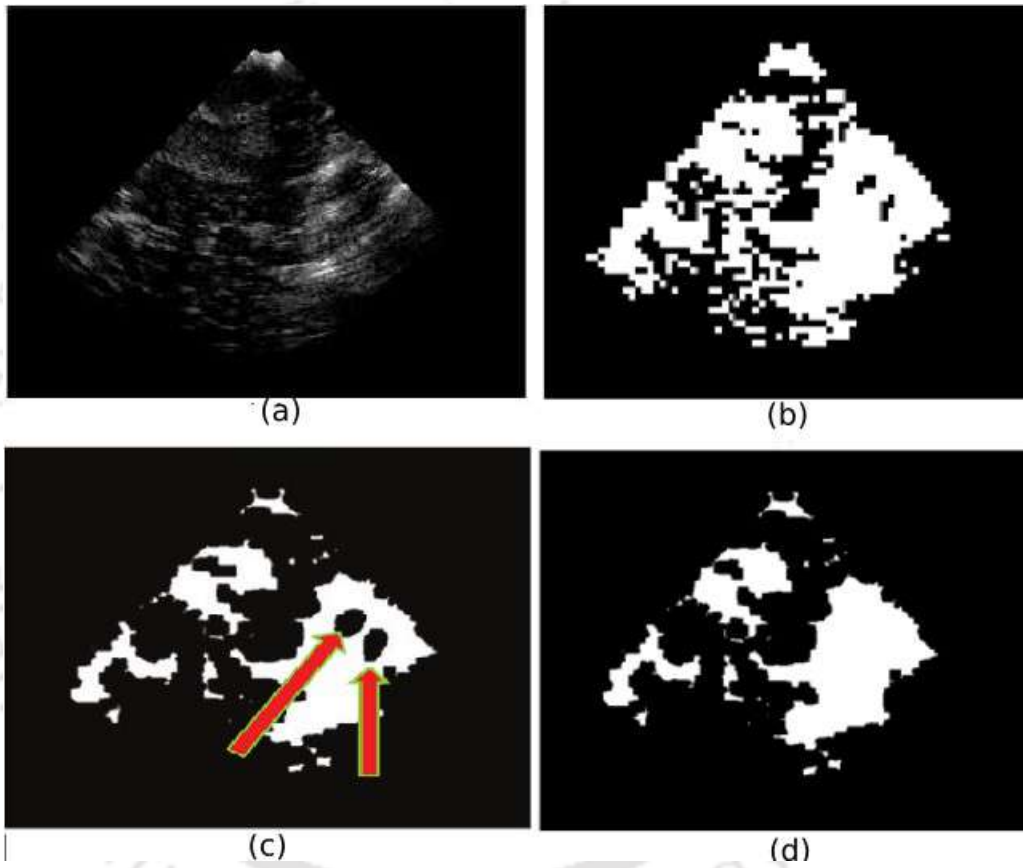


Fig. 4.4: (a) Original brain US Image from the BITE dataset. (b) Binary segmentation map after the patch-based segmentation. (c) Binary segmentation map after the boundary correction. (d)Binary segmentation map after the morphological hole filling.

patch size. The iterative boundary correction step separates distinctly different segments and also assures smooth boundaries S. This operation consists of two steps in each iteration: a) boundary ($\mathbf{C}$) detection from S using Sobel edge operator [69] and b) modifying S by comparing the average patch intensities connected with $\mathbf{C}$. $\mathbf{C}$ is a sparse binary matrix of the same size of $\mathbf{V}$ that contains a binary value of 1 where a boundary pixel exists. After the extraction of $\mathbf{C}$, a local symmetrical patch $\mathbf{P}$ of size w centered at each boundary pixel is chosen. For obtaining the symmetrical shapes of $\mathbf{P}$, the size of $\mathbf{P}$ is chosen as $(1+w)\times(1+w)$ for $w>1$ and 1 for $w=1$. The mean intensity μ_P of the boundary patch $\mathbf{P}$ is computed as $\mu_P = \frac{1}{|\mathbf{P}|}\Sigma_{(i,j)\epsilon P}V(i,j)$. The modification of $\mathbf{S}$ is done by comparing μ_P with E_T as follows:

$$
\begin{aligned}
&For (i,j)\epsilon \mathbf{P}, && (4.27)\\
&If\ \mu_P > E_T\ \text{ then} && (4.28)\\
&S(i,j) = 1\\
&Else\\
&S(i,j) = 0\\
&End && (4.29)
\end{aligned}
$$

The patch size w is reduced to $\frac{w}{2}$ in each iteration, and this step continues until $w=1$. This operation reduces the blocking artefacts with the decrease in w. Progressive reduction of the blocking artifact can be observed with reducing patch sizes $w = 8, 4\ 2$, and 1 respectively in Fig. 4.5.(a)-(d). This step requires $1 + log_2(w_{max})$ iterations and that is the only iterative step in this work. For example, if $w_{max} = 8$, the boundary correction requires only 4 iterations.

4.2.3 Morphological hole filling

The boundary correction step stated in the previous subsection reduces the blockieness on the boundaries between the HE and NHE regions. Some tiny HE and NHE regions are wrongly classified during this step. A close observation over the S shows that many smaller NHE regions appear inside the HE regions due to the presence of some fluid filled regions or acoustic shadows. Similarly, some small HE regions can also be seen within an NHE region, that mostly appear due to speckle. Both these two types of smaller regions are not significant

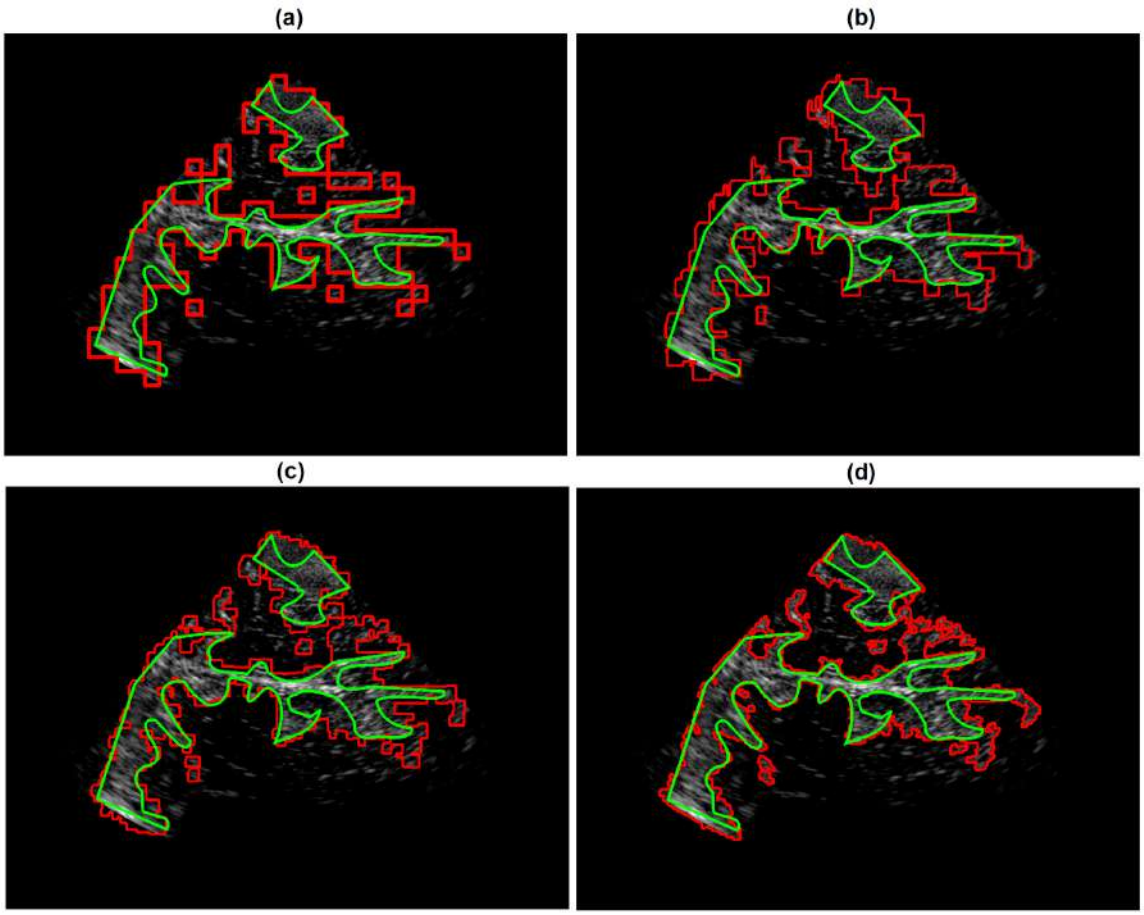


Fig. 4.5: Boundary of HE regions for (a) $w = 8$, (b) $w = 4$ (c) $w = 2$, and (d) $w = 1$. The green boundary shows the ground truth.

because of their less anatomical information. These regions often add error to the similarity measure in a registration problem. For avoiding that error and for reducing the complexity, those tiny regions are needed to be removed from S. A morphological hole-filling algorithm [69] is applied over the HE regions for filling up the small NHE regions and marking them back as HE regions. Fig. 4.4.(c) shows S after the boundary correction. It contains two small NHE regions, highlighted by arrows inside an HE region. The modified S after the hole filling is shown in Fig. 4.4.(d). It does not contain any NHE region inside an HE region.

4.2.3.1 Object separation and removal of small HE regions

After the morphological hole-filling, the segmented HE regions can be represented as $\mathbf{S}^0$: $\mathbf{S}^0 = \{(x, y)|S(x, y) = 1\}$. $\mathbf{S}^o$ can be seen as an union of some disjoint regions:

$$\mathbf{S}^0 = \bigcup_{i=1}^{L_1} \mathbf{S}_i$$

where $\mathbf{S}_i$ is the i^{th} disjoint region in $\mathbf{S}^0$. The disjoint regions $\mathbf{S}_i$s can be separated from $\mathbf{S}^o$ using the connected-component labeling (CCL) algorithm [153]. After completion of the labeling, all the disjoint regions are sorted into a sequence based on their areas. The area A_i of the region $\mathbf{S}_i$ can be described by the number of pixels inside this region. A_i is compared with a predetermined threshold A_T. If $A_i < A_T$, then $\mathbf{S}_i$ is marked as an NHE region and excluded from the HE regions. Using this step, the segmented HE region can now be written as

$$\mathbf{S}^1 = \bigcup_{\forall i: A_i > A_T} \mathbf{S}_i, \tag{4.30}$$

Let $\mathbf{C}_i$ be the boundary of the qualified regions of $\mathbf{S}_i$. The aggregated boundaries of all $\mathbf{S}^1$s can be represented in the form an edge map $\mathbf{C}$ as,

$$\mathbf{C}^1 = \bigcup_{\forall i: A_i > A_T} \mathbf{C}_i \tag{4.31}$$

The boundaries of every disjoint region is a closed contour. Equation (4.31) shows that an increase in A_T decreases the number of disjoint contours, and therefore, selecting a suitable value of A_T is important. A_T depends on how small HE regions the user needs to consider. For the images of size of 480×640, in the BITE dataset, we have chosen a value of $A_T = 350$. Fig.

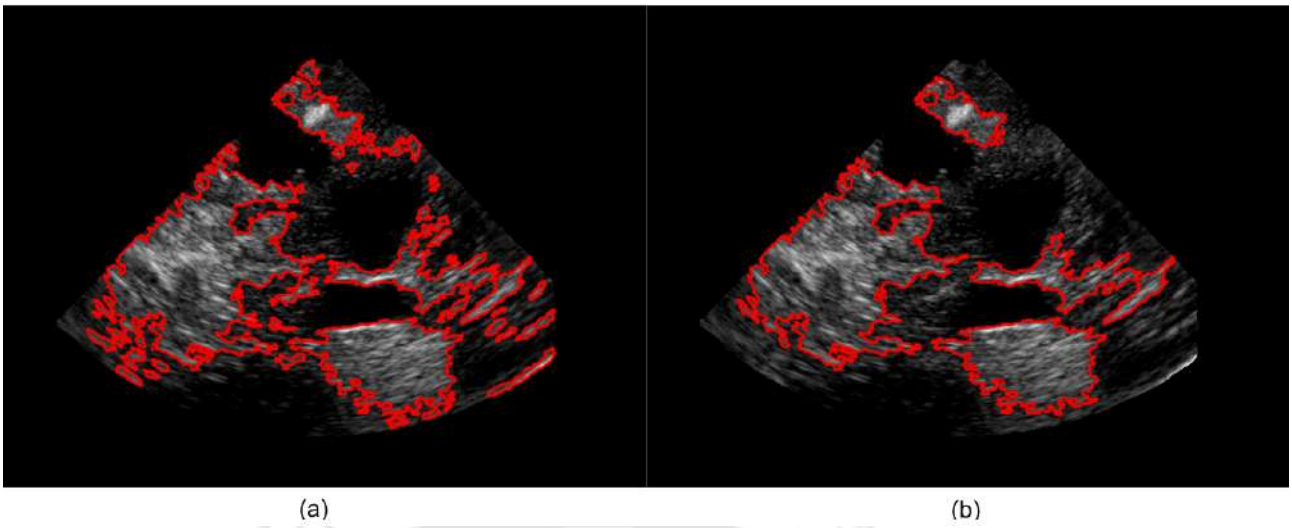


Fig. 4.6: (a) Corrected boundaries after morphological hole filling. (b) Regions after separation of the smaller HE regions as described in the section 4.2.3.1.

4.6.(a) represents the boundaries of all the disjoint regions including the smaller undesired HE regions. The qualified HE regions having an area greater than $A_T = 350$ shown in Fig. 4.6.(b). It can be observed that the tiny HE regions are not present in the output image.

4.2.4 Cubic-B spline curve shaping

Let, $\mathbf{S}^1$ in Equation (4.30) contains L_2 number of disjoint regions. The union of these regions and their respective boundaries can be expressed as, so that $\mathbf{S}^1 = \bigcup_{i=1}^{L2} \mathbf{S}_i^1$ and $\mathbf{C}^1 = \bigcup_{i=1}^{L_2} \mathbf{C}_i^1$. The boundary correction step described in the Section 4.2.2 improves the segmentation at the finest level. Visual inspection of the disjoint boundaries shows the following limitations:

(i) The region boundaries suffer from slope and curvature discontinuities at several points (shown in Fig. 4.7.(b)).

(ii) Self intersection of the boundaries occur in many places. This results in ambiguous boundaries.(shown in Fig. 4.7.(c)).

(iii) Dividing of a distinct region into many connected regions leads to incorrect segmentation.

Because of the above limitations, the contours obtained after boundary correction become irregular and non-distinct. For obtaining a regular boundary without any ambiguity, a B-spline curve shaping method is applied.

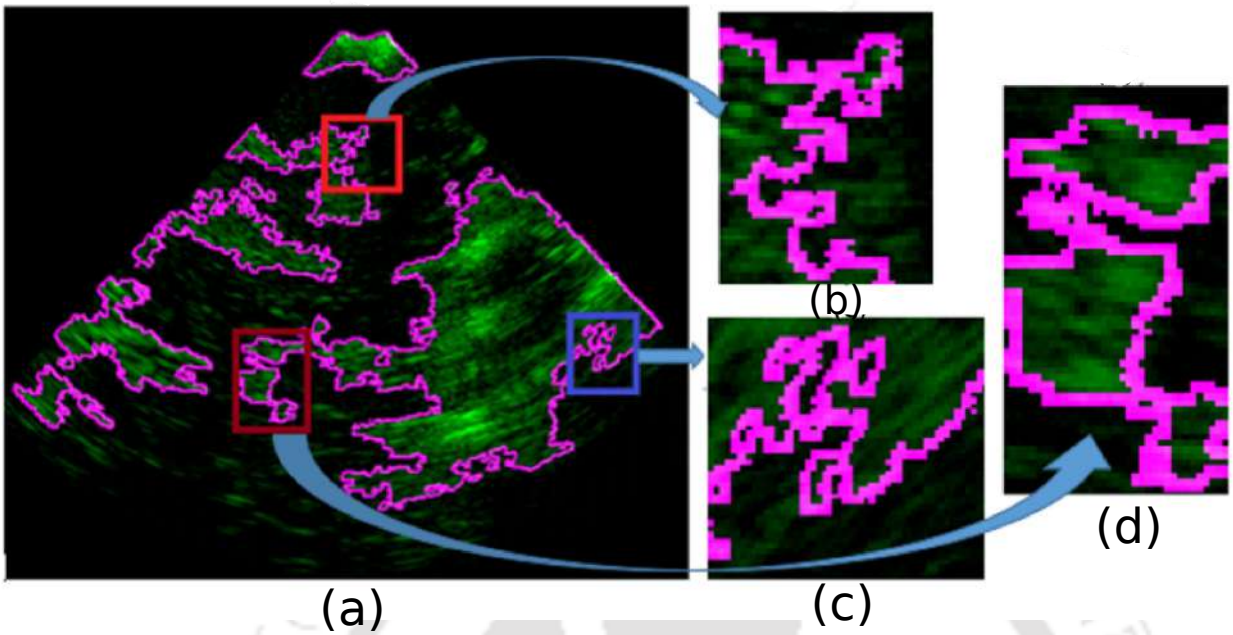


Fig. 4.7: (a) Boundaries of HE regions as described in Section 4.2.3.1. (b) Uneven boundaries without slope and curvature continuity. (c) Connected and repeated boundaries. (d) Incorrect segmentation

A B-spline curve comprises several polynomial curve segments, each of which may be derived by using the De-Boor algorithm [22]. Conventionally a B-spline curve $c(t)$ is defined by i)the order of the spline (p), ii) the control points $d_0, d_1..., d_L$, and iii) the knot vectors $\mathbf{U} = \{\underbrace{z_0, z_1, ..., z_{p-1}}_{\text{p equal knots=a}}, \underbrace{z_p, z_{p+1}, z_{p+2}, ..., z_L}_{\text{L-p+1 other knots}}, \underbrace{z_{L+1}, z_{L+2}, ..., z_{L+p}}_{\text{p equal knots=b}}\}$.

$c(t)$ is given by

$$c(t) = \sum_{i=0}^{L} N_{i,p}(t)\, d_i \tag{4.32}$$

where t lies between a and b and $N_{i,p}$ is the p order or $(p-1)$ degree B-spline basis function. For obtaining a B-spline curve, $(L+p+1)$ elements of $\mathbf{U}$ are used. The first p knots in $\mathbf{U}$ overlap, and so do the last p knots. The basis functions $N_{i,p}$ in (4.32) depend only on the knot locations and can be obtained iteratively by the following equations,

$$N_{i,0}(t) = \begin{cases} 1 & z_i \leq t < z_{i+1} \\ 0 & otherwise \end{cases} \tag{4.33}$$

$$N_{i,j}(t) = \frac{t - z_i}{z_{i+j} - z_i} N_{i,j-1}(t) + \frac{z_{i+j+1} - t}{z_{i+j+1} - z_{i+1}} N_{i+1,j-1}(t) \tag{4.34}$$

Let the index k be used to define the knot intervals $[z_k, z_{k+1}]$ that contains the position t. It is observed in (4.33) that for a given k, the B-spline basis functions are nonzero for $i = k - (p-1), ..., k$, and zero for others. Thus, Equation (4.32) can be reduced to:

$$c(u) = \sum_{i=k-(p-1)}^{k} N_{i,p}(t) d_i \tag{4.35}$$

Above equation states that only p control points are required for the reconstruction of a curve segment. Hence, it models the curve with local information only. The first and the last control points coincide because all the boundaries are closed contours. For a p degree spline curve, first, p knots coincides and so also the last p knots. Other knots are chosen uniformly along the length of the curve.

The above expressions are represented for 1D B-Spline curve construction. However, for 2D and 3D images, the basis functions are calculated separately for different axes. For example, for 2D images, there will be 2 axes X and Y. The knot vectors and control points are 2D vectors where $\mathbf{d}_i = \begin{bmatrix} d_{ix} \\ d_{iy} \end{bmatrix}, 0 \leq i \leq L-1$ and $\mathbf{z}_i = \begin{bmatrix} z_{ix} \\ z_{iy} \end{bmatrix}, 0 \leq i \leq L+P-1$ and

$$\begin{aligned} X(t) &= N_{i,p}(t) d_{ix} \\ Y(t) &= N_{i,p}(t) d_{iy} \end{aligned} \tag{4.36}$$

For obtaining smoothness of the boundary curve, the said cubic B-spline ($p = 4$) curve shaping is applied. The selection of the control points plays a crucial role in this step. The proposed method adopts an innovative way for the selection of the control points that is described below.

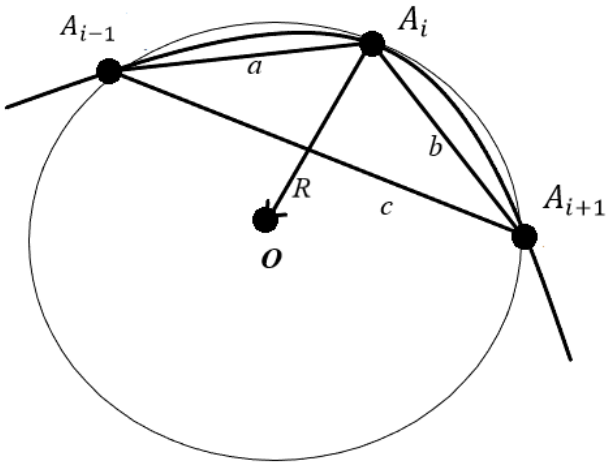


Fig. 4.8: Radius of curvature using three points

4.2.4.1 Selection of control points

The shape of a contour can be described by its curvature values at all points on the contour. Suppose, a contour is a sequence of consecutive M points $A_0, A_1, ..., A_{M-1}$. Let the three consecutive points A_{i-1}, A_i and A_{i+1} form a triangle as shown in Fig. 4.8. Let a, b and c be the lengths of the sides $A_{i-1}A_i$, A_iA_{i+1} and $A_{i-1}A_{i+1}$ respectively. The radius of the curvature (ROC) R can be derived as [129],

$$R = \frac{abc}{4\triangle} \tag{4.37}$$

where $\triangle$ is the area of the triangle. The curvature κ is measured as the inverse of R and given by

$$\kappa = \frac{1}{R} = \frac{4\triangle}{abc} \tag{4.38}$$

We assume that a smoothed B-spline curve should preserve the corners, and hence, these corner locations should be used as the control points. For choosing the sharp corners, the curvature (κ) at each point on the curve is estimated by using (4.38) and thresholded with a predetermined threshold C_T. A closed boundary curve of a segmented region is shown in Fig. 4.9.(a). Values of κ at all the points on the closed contour were estimated. The curvatures C_T of all points on C_i are compared with an empirically determined threshold $C_T = 0.6$. If the value of the curvature at any point $\kappa > C_T$, considered as the candidate control points.

The points on the boundary contour where $\kappa > 0.6$ are shown in Fig. 4.9.(b). It is observed from the figure that there are many closely spaced points at the corners. The inclusion of all those control points in cubic B-spline interpolation may lead to over-fitting. For avoiding this, some closely situated points are needed to be removed. Suppose, d_i is a control point from the candidate list. Then the nearest point in the list with the city-block distance greater than the threshold d_T is chosen as the next control point d_{i+1}. This process is continued until all candidate control points in the list are processed.

Fig. 4.9.(a), shows a total of 181 control points, obtained by thresholding of κ with empirically obtained $C_T = 0.6$. The distance between subsequent control points is thresholded with $d_T = 5$ and only 67 points are retained, which are shown in Fig. 4.9.(b). It is also seen that the regions with small curvature do not contain any node. Curve reconstruction in these regions may not follow its shape if no node is chosen in these large gaps. Hence, we put some nodes in between 2 nodes if its gap is more than 50 pixels. In such a manner, additional nodes are put in the curve as shown in Fig.4.9.(c). The final boundary curve $\mathbf{C}_i^o$ is obtained after applying Equation (4.32)-(4.35) over the selected set of control points and shown in Fig. 4.9.(d). It is evident from the figure that the final boundary $\mathbf{C}_i^{final}$ is smooth and more regularly shaped. Thus the B-spline interpolated boundaries distinctly separate the HE and NHE regions. $\mathbf{C}_i^{final}$ is a closed boundary and the region within it represents $\mathbf{S}_i^{final}$.

Because of the B-spline smoothing of the boundaries, each HE segmented region S_i^1 is now modified. Suppose there are L_2 disjoint regions. Then the final segmented HE region is given by

$$\mathbf{S}^{final} = \bigcup_{i=1}^{L_2} \mathbf{S}_i^{final} \tag{4.39}$$

4.2.5 Results of the proposed segmentation method

The performance of the proposed segmentation method can be evaluated on different aspects. Experiments were performed on 10000 pBUS and iBUS image slices of BITE and RESECT [202] datasets. The quality of the images in the BITE dataset lacks due to higher speckle and various artefacts. The images in the RESECT dataset contain comparatively less speckle and artefacts [141]. Visual inspection reveals that the contrast and the intensity variations of the images in the RESECT dataset are less compared to those of the BITE dataset.

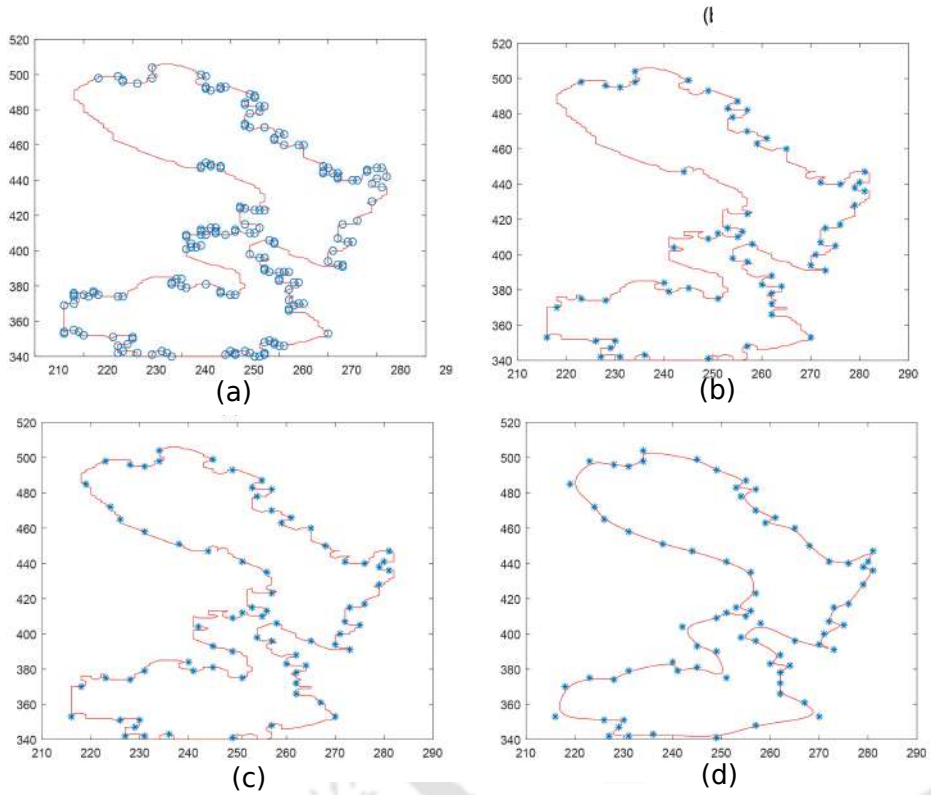


Fig. 4.9: a) A closed contour obtained after the removal of the small HE regions. The initial control points selected by thresholding $\kappa > .06$. b) The control points after the distance thresholding with $d_T = 5$. c) Additional nodes are inserted where there is large gap (more than 50 pixels) between two consecutive nodes. d) Smooth curve obtained after cubic B-spline curve shaping using the selected control points.

Metric	Formula [173]
Precision	$\frac{\lvert\mathbf{S}_R \cap \mathbf{S}_{GT}\rvert}{\lvert\mathbf{S}_R \cup \mathbf{S}_{GT}\rvert}$
Accuracy	$TN = \lvert(\mathbf{S}_R \cup \mathbf{S}_{GT})^c\rvert, TP = \lvert\mathbf{S}_R \cap \mathbf{S}_{GT}\rvert$ $FN = \lvert\mathbf{S}_R \cap \mathbf{S}_{GT}^c\rvert, FP = \lvert\mathbf{S}_R^c \cap \mathbf{S}_{GT}\rvert$ $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
Dice similarity	$D = \frac{2\lvert\mathbf{S}_R \cap \mathbf{S}_{GT}\rvert}{\lvert\mathbf{S}_R\rvert + \lvert\mathbf{S}_{GT}\rvert}$

Table 4.1: Region-based segmentation quality metrics.

The contribution of the proposed method is significant in two aspects. First, it adopts a patchwise average intensity thresholding approach to BUS image segmentation, that is more efficient than the pixel-based approach. We present the improvement in segmentation performance in Section 4.2.6 comparing the results by the patch-based method for different patch-sizes and contour smoothing. Second, we have compared the performance of the proposed method with the other state-of-the art methods in Section 4.2.7.

The performance of segmentation can be shown in both subjective manner and with parametric evaluation. For parametric comparison of image segmentation, ground truth labeled images are essential. Unfortunately, both the BITE and RESECT datasets do not provide labeled 2D images for the parametric evaluation of segmentation. For preparing ground truths, we have selectively chosen 60 images from BITE dataset [123] and 30 images from RESECT [202] dataset and annotated the data in a MATLAB application *ImageAnnotationBot* [39]. We have chosen images having good anatomical information. All the bigger HE regions are segmented manually in coordination with an experienced radiologist.

Performance evaluation of the proposed method was done in terms of both region-based metrics: precision, accuracy, and Dice, and edge-based metrics like the asymmetric squared contour distance (ASCD), and the root mean squared contour distance (RMSCD) [173].

Let $\mathbf{S}_R$ and $\mathbf{S}_{GT}$ be the resulting segmented area and corresponding ground truth respectively. The formulation of the above mentioned performance metrics are shown in Table 4.1, where $|\ .\ |$ represents the number of pixels within a region. ASCD and RMSCD are used to measure the dissimilarity between two closed contours. Let $\mathbf{C}_R$ and $\mathbf{C}_{GT}$ be the boundaries of $\mathbf{S}_R$ and

$\mathbf{S}_{GT}$ respectively. The ASCD between $\mathbf{S}_R$ and $\mathbf{S}_{GT}$ can be expressed as

$$ASCD(\mathbf{S}_R, \mathbf{S}_{GT}) = \frac{1}{|\mathbf{C}_R| + |\mathbf{C}_{GT}|} \times \left(\sum_{c_R \epsilon \mathbf{C}_R} d_E(c_R, \mathbf{C}_{GT}) + \sum_{c_{GT} \epsilon \mathbf{C}_{GT}} d_E(c_{GT}, \mathbf{C}_R) \right) \tag{4.40}$$

where $|\mathbf{C}_{GT}|$ defines the length of the contour $\mathbf{C}_{GT}$. $d_E(\mathbf{c}_{GT}, \mathbf{C}_R)$ is the shortest Euclidean distance between a point $\mathbf{c}_{GT}$ and $\mathbf{C}_R$ is expressed as,

$$d_E(\mathbf{c}_{GT}, \mathbf{C}_R) = \ min_{\mathbf{c}_R \epsilon \mathbf{C}_R} \ ||\mathbf{c}_{GT} - \mathbf{c}_R| \, |_2$$

. $d_E(\mathbf{c}_R, \mathbf{C}_{GT})$ is defined similarly.

The RMSCD between $\mathbf{S}_R$ and $\mathbf{S}_{GT}$ can be expressed as,

$$RMSCD(\mathbf{S}_R, \mathbf{S}_{GT}) = \sqrt{\frac{1}{|\mathbf{C}_R| + |\mathbf{C}_{GT}|}} \times \sqrt{\left(\sum_{\mathbf{c}_R \epsilon \mathbf{C}_R} d_E^2(\mathbf{c}_R, \mathbf{C}_{GT}) + \sum_{\mathbf{c}_{GT} \epsilon \mathbf{C}_{GT}} d_E^2(\mathbf{c}_{GT}, \mathbf{C}_R) \right)} \tag{4.41}$$

4.2.6 Results of iterative boundary correction

Patch-based segmentation compares μ_k with E_T as described in Section 4.2.1.2 and decides a patch either as HE or NHE. The segmentation output exhibits blockieness over the boundaries of S due to initial large patch size for $w = 8$. The iterative boundary correction step with reducing patch sizes ($w = 8, 4, 2,$ and $, 1$) respectively is applied later. The similarity of the resulting boundaries compared to the ground truth is analyzed using both the region-based and boundary-based metrics discussed above. The improvement was noticed in the smoothness of the boundaries as shown in Fig. 4.5.(a)-(c). The segmented regions obtained using $w = 8, 4, 2,$ and 1 are compared with the corresponding region in the ground truth. The parametric evaluation of this experiment is performed on 90 images (60 images from the BITE dataset and 30 images from the RESECT dataset) and the respective results are shown in Table 4.2. During the progress from iteration 1 ($w = 8$) to 4 ($w = 1$), it was observed that the mean precision improved from 0.817 to 0.846. It improved furhter to 0.876 after cubic B-spline

Metric	Error(mm)	Boundary correction				Cubic B-spline curve shaping
		$w = 8$	$w = 4$	$w = 2$	$w = 1$	
Accuracy	**Mean**	0.928	0.946	0.947	0.946	0.982
	Standard deviation	0.057	0.030	0.100	0.115	0.075
Precision	**Mean**	0.817	0.826	0.838	0.846	0.876
	Standard deviation	0.114	0.104	0.173	0.179	0.139
Dice	**Mean**	0.826	0.897	0.835	0.880	0.895
	Standard deviation	0.150	0.092	0.108	0.124	0.095
ASCD	**Mean**	2.287	2.039	2.044	1.934	1.394
	Standard deviation	0.773	0.752	0.919	0.997	0.522
RMSCD	**Mean**	3.183	2.843	2.518	2.507	1.839
	Standard deviation	0.625	0.982	1.212	0.829	0.928

Table 4.2: Improvement in segmentation after iterative boundary correction and B-spline curve shaping

curve smoothing. The mean Dice similarity changed from 0.826 to 0.895. The ASCD has reduced 2.287 mm to 1.394 mm and the RMSCD reduced from 3.183 mm to 1.839 mm. Fig. 4.10.(a1)-(a3) show the rectangular image sections of three different brain US images taken from the BITE dataset. Other images in Fig. 4.10 show the corresponding sections with the resulting segmented boundaries during the iterative boundary correction steps. Over iterations, the progress of the boundary smoothness after $w = 8, 4, 2$ and 1 are clearly observed in Fig. 4.10.(b1)-(b3), (c1)-(c3), (d1)-(d3) and (e1)-e(3) respectively. It is found that the blocking artefact gradually reduces and smoothness becomes optimum at $w = 1$.

The boundaries obtained after cubic B-spline curve shaping are shown in Fig. 4.10.(f1)-(f3). The boundaries appear distinct and have no slope and curvature discontinuities. During these experiments, the values of C_T and A_T were chosen 0.6 and 350 respectively. The boundaries obtained after this step preserve the actual shape of the region along with a smooth periphery.

4.2.7 Overall segmentation performance evaluation and comparison with other methods

Among the discussed approaches in subsection 4.1.1, the active contour and the level-set-based methods are found capable in segmenting brain US images. We used open access

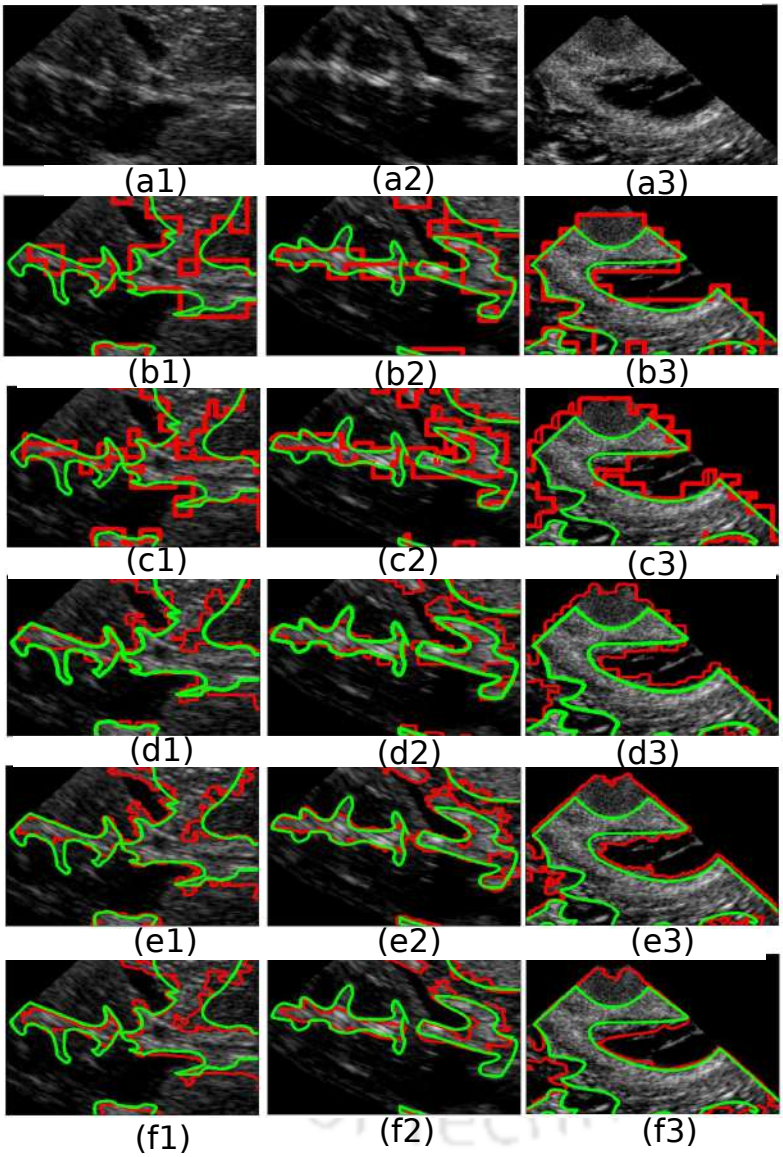


Fig. 4.10: (a1)-(a3) Original images. Outputs after boundary correction using (b1)-(b3) $w = 8$, (c1)-(c3) $w = 4$, (d1)-(d3) $w = 2$ and (e1)-(e3) $w = 1$. (f1)-(f3) show the boundaries after cubic B-spline curve shaping. Green boundary in each figure represents the ground truth and the red boundary represents the resulting segmentation after a step of the proposed method.

segmentation software CREASEG [50] that is freely available for academic and research purposes. CREASEG [202] employs six different level-set-based algorithms [12, 24, 28, 182, 97, 95] for the segmentation of all kinds of images. This software was used on the selected set of brain US images; but comparable results were obtained from ACWE [28] and Shi-Karl [182] methods only. Other algorithms were not capable of the segmentation of brain US images. These algorithms either converge too early to a sub-optimal boundary or enclose the entire fan shaped region into one segment. The performance of each algorithm was evaluated by comparing their efficiencies in isolating the disjoint HE regions and their closeness to the respective ground truths. Both the ACWE and the Shi-Karl methods require an initial boundary to be chosen manually and they should be close to the actual boundary of an HE region. Failing it, the final growth of the boundary become slow or it converges to a sub-optimal boundary. In addition to GAC and ACWE, comparisons have been done with two other level-set-based methods, namely DRLSE [98] and WLSE [90]. The MATLAB program for DRLSE is openly available in https://www.mathworks.com/matlabcentral/fileexchange/68185-drlse-image-segmentation", whereas, the MATLAB code of WLSE was implemented by author. The detailed subjective and parametric studies are presented below.

For evaluation, each of the algorithms was run 10 times by changing their controlling parameters. The initial rectangular boundaries were chosen as close as possible to the actual boundaries. The best results out of 10 trials, were considered for final comparison with the other methods. Many algorithms extend boundaries beyond the fan shaped region. In such cases, a hard boundary was set to the fan shaped region only. It has been observed that the algorithm of Chan-Vese required 800-1500 iterations with a mean of 1280 iterations for reaching to the optimal solution. Similarly, the method of Shi-Karl continues up to 150-200 iterations with a mean of 182 iterations to converge to the final boundary. Both the methods DRLSE and WLSE spend 450-800 iterations with a mean of 523 and 620 iterations respectively. The proposed patch-based method involves simple calculations and requires only four iterations (for $w_{max} = 8$) during boundary correction, hence, its complexity is less.

We have recorded results of a total of 90 images from the BITE and the RESECT datasets. For giving an idea of the performance of the segmentation on individual images, the detailed results in terms of RMSCD and precision values of 24 randomly selected images (14 images from

Sl. No.	File name	RMSCD values					
		ACWE (mm)	*Shi-Karl et al. (mm)*	*DRLSE (mm)*	*Khadidos et al. (mm)*	*Nitsch et al. (mm)*	*Proposed method (mm)*
1	1a.2dus.00088sm.mnc	7.198	4.976	13.209	11.411	8.634	3.784
2	1a.2dus.00187sm.mnc	6.653	3.719	7.517	9.517	6.225	4.684
3	1a.2dus.00247sm.mnc	2.317	4.804	7.018	9.018	2.197	2.172
4	1a.2dus.00332sm.mnc	1.365	9.538	6.99	8.99	0.892	3.027
5	1a.2dus.00270sm.mnc	2.28	1.107	10.005	12.005	2.144	0.839
6	2a.2dus.00036sm.mnc	6.807	5.296	9.53	11.53	4.991	3.901
7	2a.2dus.00057sm.mnc	4.593	3.322	10.632	12.632	3.829	2.053
8	2a.2dus.00130sm.mnc	8.957	4.848	10.071	12.071	8.93	5.655
9	3c.2dus.00182sm.mnc	1.769	1.465	13.793	15.793	1.26	2.16
10	3c.2dus.00244sm.mnc	3.066	4.076	9.356	11.356	2.725	3.998
11	3c.2dus.00331sm.mnc	0.765	2.673	10.495	12.495	2.242	3.331
12	3c.2dus.00308sm.mnc	4.403	4.013	17.148	19.148	3.682	2.079
13	3c.2dus.00415sm.mnc	1.679	2.81	18.146	20.146	1.438	2.216
14	4a.2dus.00001sm.mnc	4.163	2.723	12.926	14.926	3.916	4.821
15	aftercase1z287	3.838	3.986	13.189	15.515	3.128	3.514
16	beforecase1y204	8.75	11.041	21.268	23.432	8.541	3.421
17	case4beforez184	1.032	10.498	31.335	18.77	0.852	1.173
18	beforecase1x223	1.775	7.894	27.594	30.022	1.442	6.638
19	case4beforex116	2.515	3.268	11.466	14.715	2.092	0.999
20	aftercase1x231	0.934	0.992	8.707	11.573	2.978	0.757
21	case4beforey92	0.526	0.225	11.499	14.375	1.645	1.293
22	beforecase1z125	2.083	1.209	11.817	13.303	1.481	1.2
23	case2duringz235	6.414	7.884	8.033	10.18	5.723	3.783
24	case2duringx128	11.71	9.249	16.174	17.637	3.653	2.674
	Mean	3.983	4.651	13.247	14.607	3.527	**2.924**
	Standard deviation	2.984	3.068	6.043	4.859	2.408	1.541

Table 4.3: Comparison of **RMSCD** between segmented regions and the ground truths for 24 images from BITE and RESECT datasets.

the BITE dataset and ten images from RESECT dataset) are shown in Table 4.3 and Table 4.4 respectively. Summarized results for all 90 images in terms of accuracy, Dice, and ASCD are shown in Table 4.5. The value of k_1 in (4.19) was chosen between 0.5 - 1.2 for obtaining the best results for the images of the BITE dataset and 2 - 2.5 for the images of the RESECT dataset. It has already been discussed that the images of the RESECT dataset contain less noise compared to those of the BITE dataset, and as a result, their local intensity variance is less. For compensating the low variance, the value of k_1 in (4.19) was chosen higher for the RESECT dataset. From Table 4.3, Table 4.4, and Table 4.5, it is seen that the proposed method resulted in mean values of precision, accuracy, Dice, ASCD, and RMSCD as 0.828, 0.982, 0.895, and 1.394 mm, and 2.924 mm respectively. The second-best method ACWE resulted in the mean values of the metrics as 0.798, 0.921, 0.918, 2.184 mm, and 3.683 mm respectively. From the above results, it was found that the proposed method outperformed all other level-set-based methods.

Sl. No.	File name	Precision values					
		ACWE	*Shi-Karl et al.*	*DRLSE*	*Khadidos et al.*	*Nitsch et al.*	*Proposed method*
1	1a.2dus.00088sm.mnc	0.724	0.769	0.456	0.601	0.818	0.797
2	1a.2dus.00187sm.mnc	0.756	0.761	0.759	0.759	0.757	0.756
3	1a.2dus.00247sm.mnc	0.817	0.797	0.705	0.705	0.872	0.754
4	1a.2dus.00332sm.mnc	0.945	0.795	0.800	0.800	0.835	0.953
5	1a.2dus.00270sm.mnc	0.815	0.856	0.572	0.572	0.909	0.860
6	2a.2dus.00036sm.mnc	0.855	0.839	0.685	0.685	0.897	0.891
7	2a.2dus.00057sm.mnc	0.889	0.886	0.698	0.698	0.893	0.885
8	2a.2dus.00130sm.mnc	0.823	0.808	0.671	0.671	0.839	0.790
9	3c.2dus.00182sm.mnc	0.849	0.816	0.585	0.585	0.934	0.838
10	3c.2dus.00244sm.mnc	0.709	0.680	0.584	0.584	0.729	0.801
11	3c.2dus.00331sm.mnc	0.801	0.742	0.535	0.535	0.760	0.751
12	3c.2dus.00308sm.mnc	0.840	0.786	0.439	0.439	0.723	0.791
13	3c.2dus.00415sm.mnc	0.725	0.695	0.406	0.406	0.749	0.739
14	4a.2dus.00001sm.mnc	0.626	0.624	0.433	0.433	0.718	0.602
15	aftercase1z287	0.636	0.637	0.196	0.290	0.780	0.621
16	beforecase1y204	0.837	0.834	0.363	0.288	0.798	0.987
17	case4beforez184	0.814	0.815	0.326	0.678	0.850	0.894
18	beforecase1x223	0.800	0.770	0.232	0.286	0.833	0.791
19	case4beforex116	0.639	0.722	0.488	0.440	0.715	0.854
20	aftercase1x231	0.867	0.757	0.378	0.441	0.804	0.991
21	case4beforey92	0.858	0.909	0.525	0.616	0.701	0.853
22	beforecase1z125	0.705	0.749	0.475	0.482	0.818	0.693
23	case2duringz235	0.936	0.829	0.699	0.726	0.248	0.990
24	case2duringx128	0.783	0.677	0.555	0.490	0.895	0.823
	Mean	0.794	0.773	0.523	0.550	0.786	**0.828**
	Standard deviation	0.086	0.072	0.158	0.148	0.131	0.114

Table 4.4: Comparison of the **precision** between the output boundary and the ground truth by different methods for 24 images from BITE and RESECT datasets

Method	Precision		Accuracy		Dice		ASCD (mm)		RMSCD (mm)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
ACWE	0.798	0.082	0.921	0.096	**0.918**	0.099	2.184	0.724	3.683	1.882
Shi-Karl	0.782	0.084	0.863	0.106	0.878	0.088	2.118	1.019	4.242	2.024
DRLSE	0.562	0.168	0.92	0.07	0.703	0.185	6.124	3.02	12.245	4.264
WLSE	0.548	0.11	0.895	0.126	0.718	0.192	5.831	2.548	13.068	4.232
Nitsch et al.	0.786	0.131	**0.993**	0.144	0.836	0.161	2.567	0.65	3.527	2.408
Proposed method	**0.828**	0.116	0.982	0.05	0.895	0.095	**1.394**	0.272	**2.924**	1.442

Table 4.5: Performance comparison of six different methods in terms of the precision, accuracy, Dice, ASCD, and RMSCD are shown. Results have been evaluated on 90 images for the level-set-based methods and the proposed method. The results for the method of Nitsch *et al.* are estimated from 24 test images.

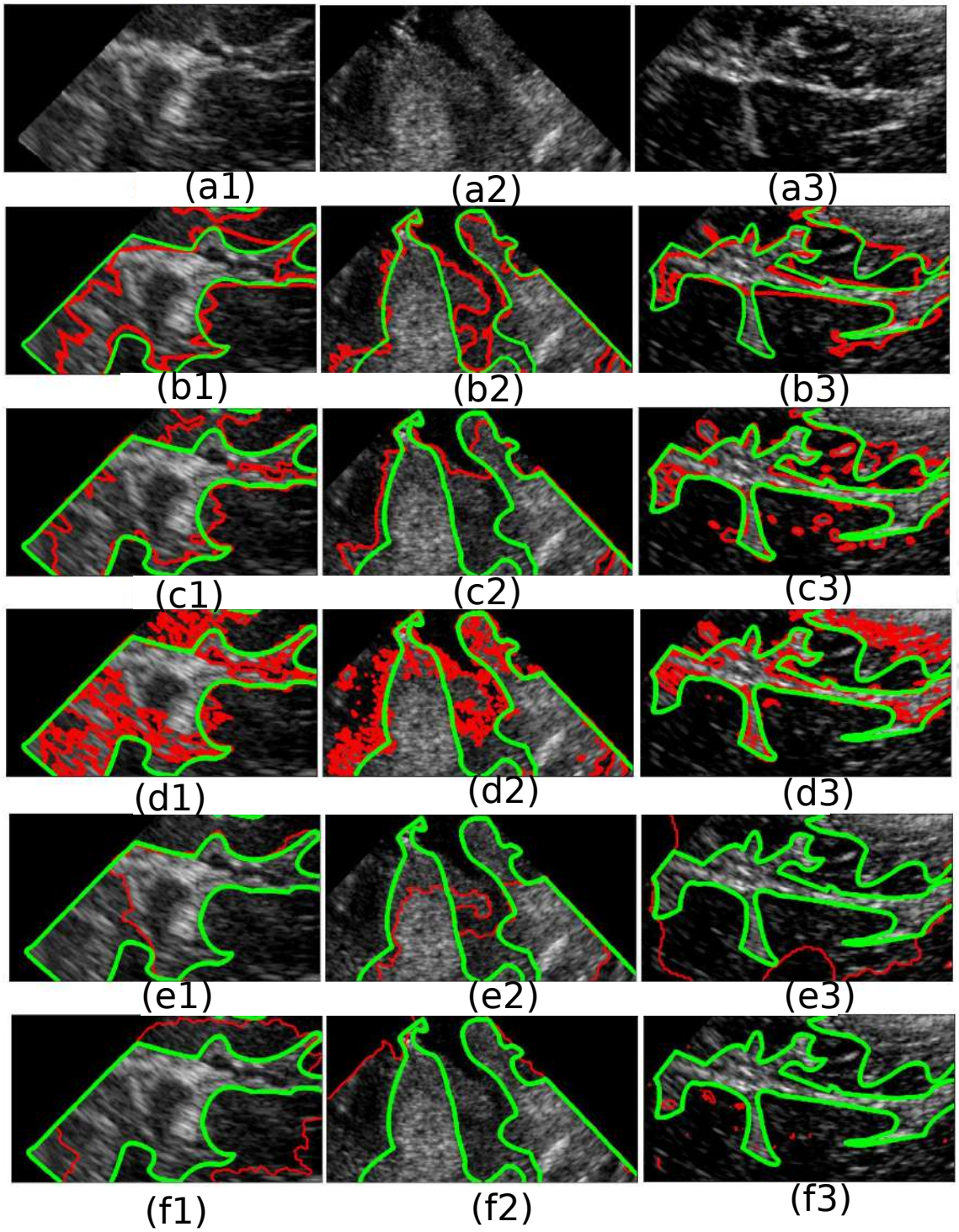


Fig. 4.11: (a1-a3) Original brain US image sections. Comparison of the ground truth by (b1-b3) proposed method, (c1-c3) ACWE, (d1-d3) method of Shi-Karl, (e1-e3) DRLSE, and (f1-f3) WLSE method. The green boundary in each figure represents the ground truth and the red boundary represents the output of a method.

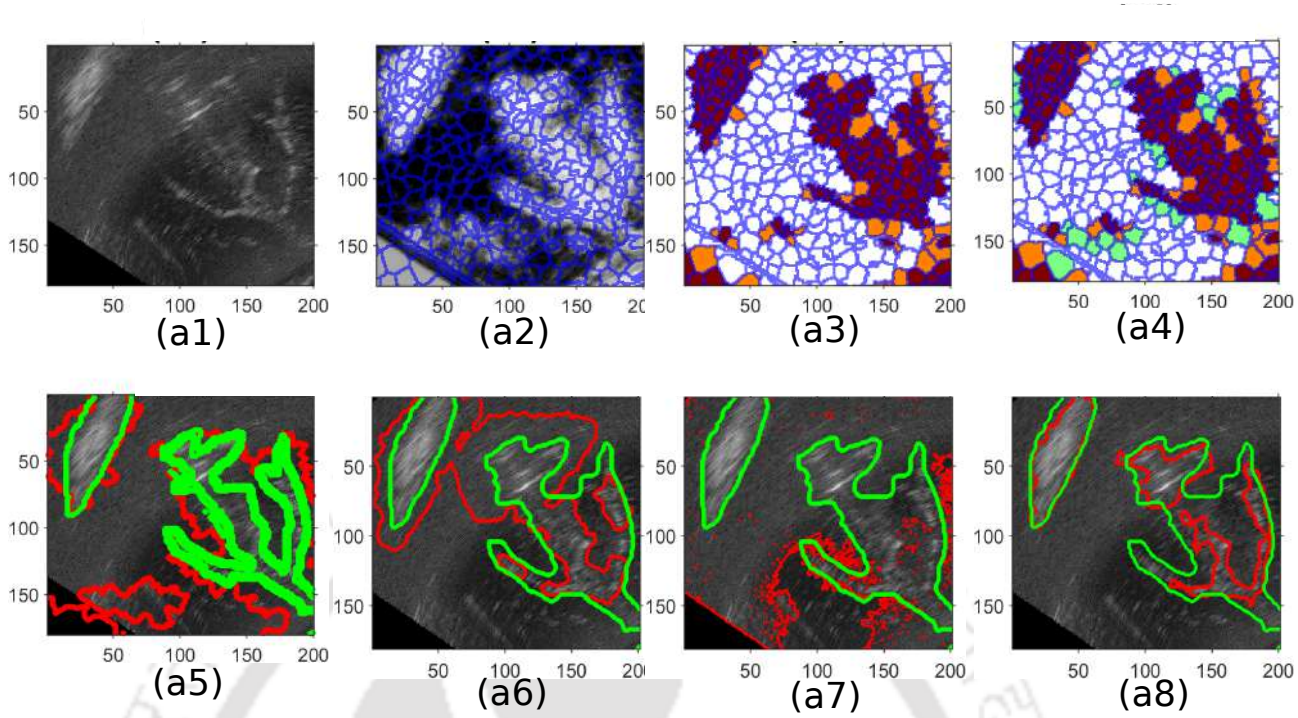


Fig. 4.12: (a1)Original rectangular brain US image section from RESECT dataset. (a2) Orientation of SLIC superpixels with N=1000 and the intensity of a superpixel indicates its weight λ_i. (a3) The superpixels with *magenta* colour have class weight ($\lambda_i > 0.59$) and the orange coloured superpixels have ($0.59 \geq \lambda_i > 0.45$) (a4) The green patches have ($0.45 \geq \lambda_i \geq 0.3$). After combining these three types of superpixels, the resulting segmentation with corresponding ground truth is shown in (a5). (a6) shows the output of the ACWE method, (a7) shows the output of the WLSE method, and (a8) shows the output of the proposed method. The green boundaries in (a5)- a(8) represent the ground truths and the red boundaries represent the resulting segmentations.

The comparative performances of the algorithms on three images is shown in Fig. 4.11. Each figure in (a1-a3) consists of rectangular sections of brain US images from BITE dataset. Respective ground truths of those image sections are shown with green color. The resulting outputs of different algorithms are shown in red color in the rest of these figures. The outputs of the proposed method, ACWE, Shi-Karl, DRLSE and WLSE methods are shown in (b1)-(b3), (c1)-(c3), (d1-d3), (e1)-(e3) and (f1)-(f3) respectively. It can be observed that the proposed method results in closer boundaries to the ground truth compared to other methods.

Apart from the level-set-based methods, both subjective and parametric comparisons have been performed with the random forest-based superpixel classifier approach by Nitsch *et al.* [141]. Random forest is popular for those classification problems where the size of the training dataset is small [141]. Out of 90 ground truth images prepared by us, 56 images were used for

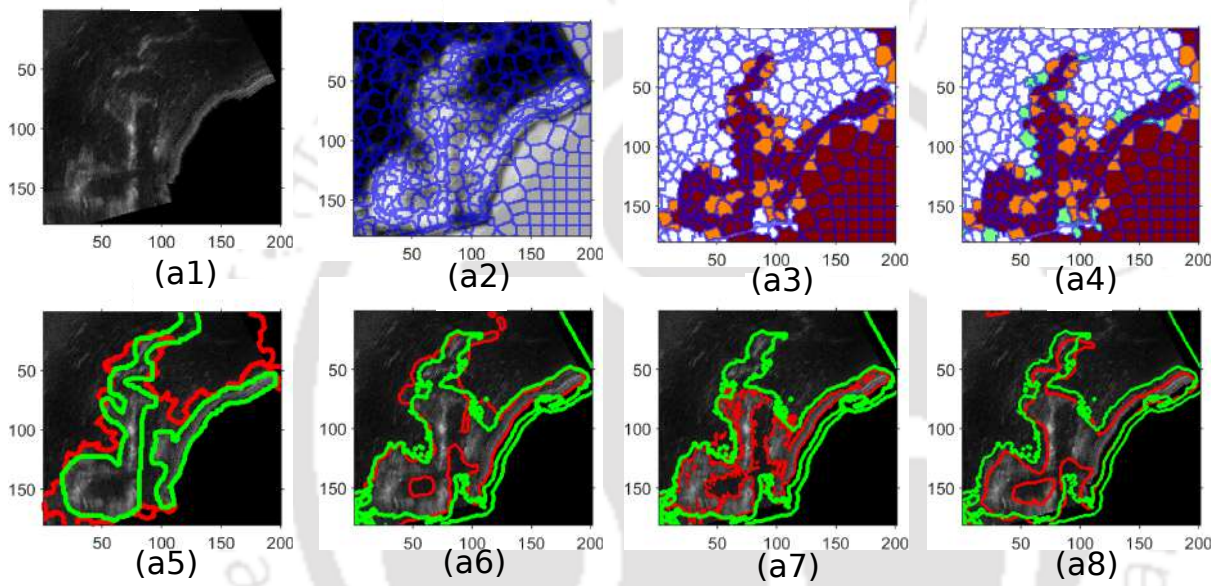


Fig. 4.13: (a1)Original rectangular brain US image section from RESECT dataset. (a2) Orientation of SLIC superpixels with N=1000 and the intensity of a superpixel indicates its weight λ_i. (a3) The superpixels with *magenta* colour have class weight ($\lambda_i > 0.59$) and the orange coloured superpixels have ($0.59 \geq \lambda_i > 0.45$) (a4) The green patches have ($0.45 \geq \lambda_i \geq 0.3$). After combining these three types of superpixels, the resulting segmentation and corresponding ground truth are shown in (a5). (a6) shows the output of the ACWE method, (a7) shows the output of the WLSE method and (a8) shows the output of the proposed method. The green boundaries in (a5)- a(8) represent the ground truths and the red boundaries represent the resulting segmentations.

training of the classifier and remaining 24 images as shown in Table 4.3 and Table 4.4 were used for testing. This classifier provides a weight vector that indicates the probability of a superpixel lying in a particular class. Some rule-based thresholding of that weight is applied later to concatenate the superpixels and obtain the final segmented cluster. We have examined this method for different implementation parameters. It can be observed that the kernel size of the stick filter plays an important role in the performance of the overall method. It should be greater than 5 and less than 15 to obtain its optimal performance without losing the edge details. Another important parameter is the number of SLIC superpixels to be used to divide the image. It was found that the segmentation performance increases with the increase in the number of superpixel clusters used. But increasing the number of superpixels also increases the computational complexity. Experiments were conducted with the different numbers of superpixels (N) and found that for a 480×640 image, $N = 2500$ provides the best performance. But in most cases, this classifier under-segments a region. Due to the rule-based concatenation of the superpixels, many regions get joined together and therefore, the overall performance degrades.

The performance of the proposed algorithm on two images from the RESECT dataset is shown in Fig. 4.12 and Fig.4.13. Fig. 4.12.(a1) and Fig. 4.13.(a1) show the cropped rectangular sections. Images in Fig. 4.12.(a2) and Fig. 4.13.(a2) show the orientation of superpixels within the image sections. During classification, the random forest classifier assigns a weight (λ_i) to the i^{th} superpixel in the HE region. Fig. 4.12.(a3) and Fig. 4.13.(a3) show the superpixels with two coloured zones. The superpixels in the central zone with the *maroon* colour have a weight greater than 0.59. Similarly, the weights of the *orange* superpixels are greater than 0.45 and less than 0.59. For the *green* coloured superpixels as shown in Fig. 4.12.(a4) and Fig. 4.13.(a4), the threshold weights are greater than 0.3 and less than 0.45. The rule-based concatenation is performed as described by Nitsch *et al* [141]. During concatenation of the superpixels, the first *maroon* superpixels are chosen initially as the segmented HE region. In the next iteration, the orange superpixels, contiguous to the *maroon* region are added to the segmented region. In a similar way, those green superpixels contiguous to the composite segmented region made of *maroon* and *orange* superpixels are also added to the segmented region and thus the final segmented region is obtained. The final boundaries obtained by this method are shown in Fig. 4.12.(a5) and Fig. 4.13.(a5) respectively. In Fig. 4.12.(a5)-(a8) and Fig. 4.13.(a5)-

(a8), the red coloured boundaries represent the output of a segmentation method and green coloured boundaries are the respective ground truths. Fig. 4.12.(a6) and Fig. 4.13.(a6) are the outputs of ACWE and Fig. 4.12.(a7) and Fig. 4.13.(a7) show the outputs of WLSE methods. The outputs of the proposed method are shown in Fig. 4.12.(a8) and Fig. 4.13.(a8). Visual inspection of the segmented results and respective ground truths show that the performance of the proposed method is more accurate than all other methods. Quantitative evaluation of the results on 24 test images are shown in Table 4.2, Table 4.3, and Table 4.4. The mean accuracy, precision, Dice, ASCD and RMSCD for the method of Nitsch *et al* were found to be 0.993, 0.794,0.895, 2.567 mm and 2.204 mm respectively. The results were obtained for 2500 superpixels per image.

In addition to the adequate segmentation performance, this patch-based method is computationally faster. All the discussed methods were implemented in MATLAB installed computer, having an eighth generation Intel core i5 processor (2.30 GHz), and 24 GB DDR4 RAM. We note its run times for different steps and show them in a pi-chart in Fig. 4.14.(b). This method does not require any iterative optimization; therefore it is much faster compared to other level-set-based methods. The run time for the steps gross segmentation, b) boundary correction, region filling and smaller HE region removal, boundary smoothing using cubic B-spline are found to be 0.057 secs., 0.3249 secs., 0.057 secs. and 0.0987 secs. respectively. We noted the average run time of other four level-set-based algorithms ACWE, DRLSE, method of Shi *et al.*, and WLSE on the same images, and estimated their average run times, which are plotted in Fig. 4.14.(a). All the level-set-based methods were initiated with the same rectangular boundary close to the desired regions. The average run time to reach to the final boundary are shown in the bar chart of Fig. 4.14.(a). All the times mentioned here are for running the main steps of each algorithms only. Time to achieve final segmentation for the methods ACWE, Shi-Karl, DRLSE and WLSE were noted as 34.23 secs., 22.01 secs., 28.86 secs., and 32.28 secs. respectively. The test part of the method of Nitsch *et al.* includes time for feature extraction and the rule-based concatenation of superpixels. The mean time required for feature extraction was 1.21 seconds and for analyzing the data and rule-based concatenation of superpixels, it took 0.342 secs. Thus, the proposed segmentation method is more time-efficient than other four level-set-based methods, and the random forest-based method.

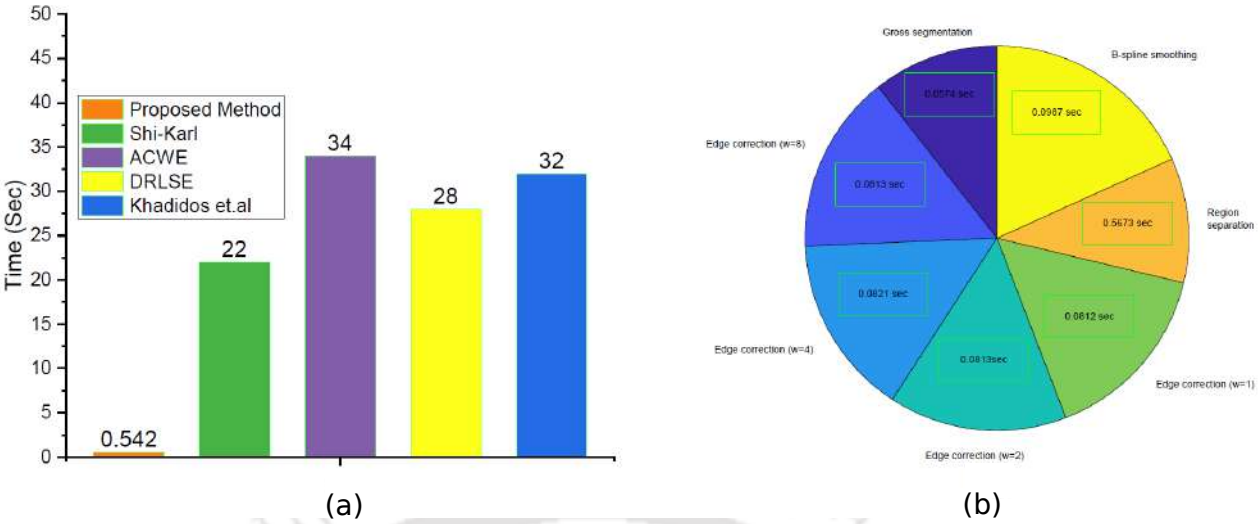


Fig. 4.14: (a) Comparison of runtime of different algorithms and (b) time taken for different steps of the proposed algorithm.

4.3 Registration of pBUS and iBUS images using the segmented HE regions

The previous section described a simple and efficient patch-based segmentation method for extracting the HE regions from pBUS and iBUS images. HE regions are relatively prominent structures that do not disappear completely in the iBUS images. Fluid filled regions like ventricles and blood vessels are also prominent regions that appear as anechoic and often remain connected to the acoustic shadows. It is difficult to determine their boundaries. Segmentation of HE regions from the registering image pairs increases the similarity between them and helps in easy registration. In the previous section, we discussed about a patch-based segmentation method of HE regions from pBUS and iBUS images. In the next section we perform image registration of segmented HE regions of pBUS and iBUS images.

4.4 Proposed registration method

The overall registration framework is shown in Fig. 4.15. After obtaining the HE regions from the pBUS and iBUS image pair, the proposed method registers them by minimizing the

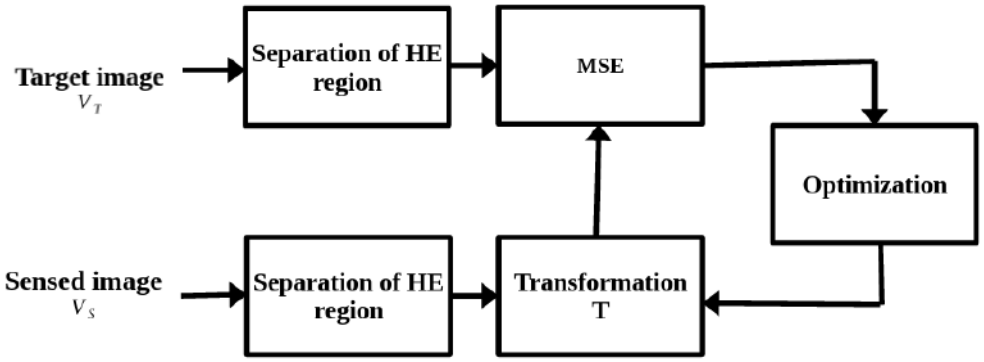


Fig. 4.15: Block diagram of the proposed registration method

MSE using PSO. We discussed in the previous chapter that PSO works better for noisy image registration than the gradient-based and other heuristic optimization methods [31, 33]. In this registration scheme, MSE is used as the objective function. The similarity measure, the deformation model, and the optimization for this registration are discussed in the following subsections.

4.4.1 Selection of the similarity measure

The selection of a good similarity/dissimilarity measure is crucial for successful image registration. Among the various dissimilarity measures, MSE [33], normalized cross-correlation (NCC) [122], and RapTOR [165] are popular. These measures were found to work well in US-US and US-MR image registration. The selection of the similarity measure also depends on the various modalities of the registering image pairs. Efstathiou *et al.* [56] reported a comparative study of six different similarity measures such as MSE, NCC, CC, histogram-based mutual information, histogram-based normalized mutual information, and Mattes mutual information.

4.4.2 Selection of the optimization algorithm

In most cases, the error hypersurface is multimodal and gradient-based optimization methods such as Newton's method, LMA, and LBFGS do not perform well, because they converges into local optima. Efstathiou *et al.* [56] compared Nelder-Mead downhill simplex, Powell direction set LBFGS (quasi-Newton), Polak-Ribiere (conjugate gradient), Regular Step Gradient Descent (RSGD) methods in 3D ultrasound image registration, and reported that LBFGS performs best among them. Canalini *et al.* [20] used the LBFGS algorithm for optimizing the

SSE. Gradient-based methods solve convex-optimization problems faster than heuristic optimizations. In a practical scenario, the error hypersurface for pBUS-iBUS image registration is non-convex over a wide search region. Heuristic optimizations perform better in non-convex optimization because do not use gradient information of the error hypersurface, They selects best solutions based on the ranks of their function values; hence, they do not get affected by the minor variations and unevenness in the error hypersurface. In this registration problem, we compare a gradient-based optimization LBFGS with three other heuristic optimization methods: BBP, SA, and PSO.

4.4.3 Deformation model

The pBUS and iBUS images of BITE and the RESECT datasets are not out of scale because the scan depth was set during the capturing of these images. Hence, the deformation is considered rigid and can be approximated by successive combinations of translation and rotation. Let (x', y') be the coordinate of a point (x, y) after deformation. It can be modeled by a unified representation as

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} cos\theta_0 & -sin\theta_0 & 0 \\ sin\theta_0 & cos\theta_0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{4.42}$$

where x_0 and y_0 indicate the translation, and θ_0 is the measure of anticlockwise rotation with respect to the center of the image. The cost function is given by

$$J = \frac{\Sigma_{x=0}^{M-1}\Sigma_{y=0}^{N-1}(V_S(x,y) - V_T(x,y))^2}{MN} \tag{4.43}$$

, where $\mathbf{V}_S$ is the sensed image and $\mathbf{V}_T$ is its corresponding target image. Each of these images has a size of $(M \times N)$.

A detailed description of the working of PSO was described in Section 2.6.0.3. At each iteration, this algorithm estimates the cost function, swarm velocity, and the current position of the swarm using Equations (2.22) and (2.23). p_1 and p_2 are the random numbers between 0 and 1. $\alpha(n)$ and $\beta(n)$ are two user-defined numbers between 0 and 1, and ξ is known as the inertia parameter that has a random value nearly equal to 1. The dynamic of PSO assures gradual improvement

of the solution process and convergence of the error hyper-surface [18]. The search ranges for the parameters were kept at $[-100\ 100]$ and θ_0 $[-30^\circ\ 30^\circ]$ by trial and error.

4.4.4 Results and discussion

We have selected a total of 68 pairs of pBUS and corresponding iBUS image pairs from the BITE and the RESECT datasets for the evaluation of the performance of the proposed algorithm. A registering image pair represents nearly the same regions of the brain but their appearances are different because of different levels of speckle and artifacts. We divide all images into three groups: *easy*, *medium*, and *difficult* depending on the visual similarities present between registering image pairs. Detailed discussion on the selection of the images can be found in Section 3.3.1. The following table shows the number of images of the BITE and the RESECT datasets present in each category.

Number of image pairs	Easy	Medium	Difficult
BITE	4	25	14
RESECT	2	14	9

Table 4.6: Number of image pairs in *easy*, *medium*, and *difficult* category.

The proposed method minimizes MSE for obtaining optimum x_0, y_0, and θ_0. Compared to the deformation model adopted in Chapter 3, it is simple, and contains only three registration parameters. We applied LBFGS, PSO, BBO, and SA methods for evaluating their performances in this scenario. The results were validated by comparing the common tag-points in pBUS and iBUS images. The tag-points were chosen from the registering image pairs with the assistance of the radiologist using graphical user interface tool developed in MATLAB. Generally, the junction points of the sulci, ventricle horns, and bends of the choroid plexus are chosen as the common tag-points in the two images. The mTRE before registration (E_b) and the mTRE after registration (E_a) are estimated using Equation (3.14) and Equation (3.15) respectively.

The proposed registration method segments the corresponding HE regions from the registering image pair as described in Section 4.2.1.2. k_1 is the only user-defined parameter needed to be varied for obtaining the desired gross boundaries. This step produces some undesired tiny

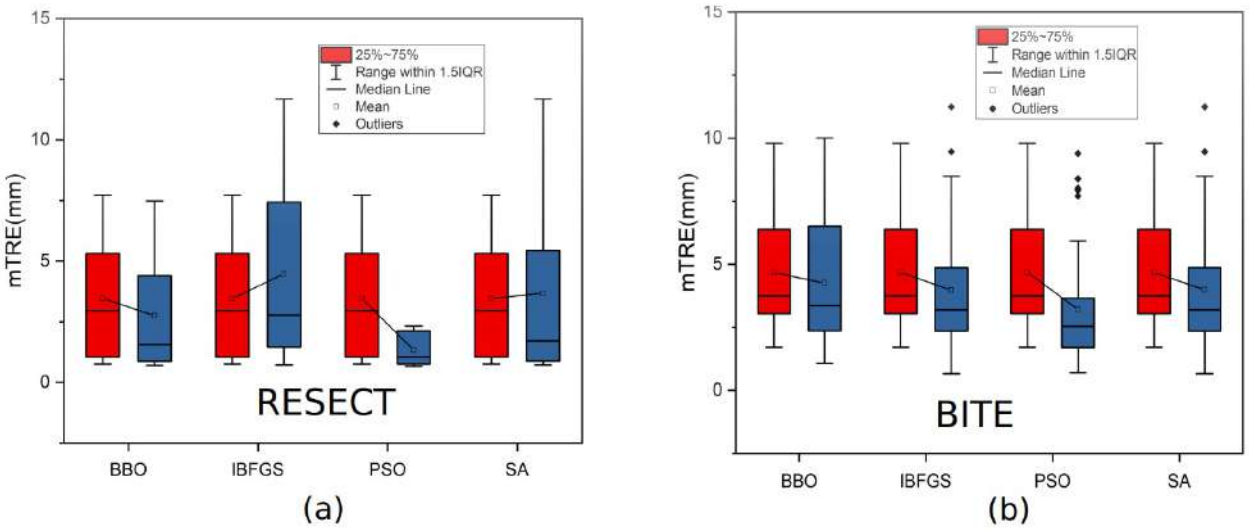


Fig. 4.16: Comparison of the mTRE performance of BBO, LBFGS, PSO, and SA algorithms on the images of (a) the RESECT dataset and (b) the BITE dataset. For a particular method, the red boxplots and blue boxplots represent the mTREs before registration and after registration respectively. The mean values of mTREs before and after registration are connected to highlight the improvement of a method.

regions that are removed because of no anatomical information. These disconnected regions were identified by connected component labeling [68]. If the area of a region is found to be greater than 300 pixels, it was labeled as an inside region. The experimental results for the RESECT and the BITE datasets are presented in Fig. 4.16.(a) and Fig. 4.16.(b) respectively.

Fig. 4.16.(a) shows the mTRE values of 23 pairs of images from the RESECT dataset and Fig. 4.16.(b) shows the results for 39 pairs of images of the BITE dataset. All those images were taken from the stack of moderate and difficult category images. This figure shows that PSO achieved better mTREs than the other three optimization methods. Fig. 4.16.(a) shows that the mTRE before registration is 3.45 mm, that is improved to 2.76 mm, 4.47 mm, 1.33 mm, and 3.68 mm by BBO, LBFGS, PSO, and SA respectively. Similarly, the results in Fig. 4.16.(b) shows the results for the images of BITE dataset. The mTRE before registration is 5.04 mm, that is improved after registration to 4.41 mm, 4.27 mm, 3.67 mm, and 5.12 mm by BBO, LBFGS, PSO, and SA respectively. These results show that PSO performs better compared to other three optimization methods.

The inertia parameter (ξ) of the particle velocity equation of PSO is set to a random number

between 0.9 and 1.1. Both p_1 and p_2 were set to 0.3. The swarm size was varied gradually from 20 to 80 with a step of 10. All these values of the parameters were found by trial and error for producing the best results. Observations revealed that when swarm size is less, PSO continues for more iterations to converge to the optimum. For example, it takes an average of 20 iterations to converge for a swarm size of 80, whereas for a swarm size of 20, it takes approximately 60 iterations. The choice of the search ranges of the parameters is a difficult task. It depends on the size of the images, and the nature of the deformation. An efficient optimization algorithm can find the global optimum within a wide search range. In this experiment, we have chosen the range of x_0 and y_0 from -100 pixels (equiv. to -20 mm) to 100 pixels (equiv. to 20 mm). The upper bound and the lower bound of θ were kept at $-30°$ and $30°$. It was found that for most image pairs, the relative translation and rotation were within these limits.

The performance of the proposed method was compared with the SIFT-feature-based technique developed by Machado *et al.* [108] and the PSO-based method in the previous chapter. The original method of the Machado *et al.* is for the registration of 3DUS images. We have implemented the same steps for 2D pBUS and iBUS images. 2D-SIFT features were extracted from the registering image pairs and an exhaustive search technique was used for matching the obtained features. Parameters of the SIFT features such as the edge threshold, contrast threshold, and the spread parameter of the Gaussian filter were chosen empirically. The threshold for normalized contrast was chosen between 0.01 to 0.1. The edge threshold was chosen between 8 to 15. All these parameters were varied to find the optimally matched features between the registering image pairs.

The comparative results of the proposed method, the method in the previous chapter, and the feature-based method by Machado *et al.* [108] are shown in Fig.4.17. The SIFT features were extracted from the segmented HE regions only. Fig. 4.17.(a) shows the comparative results for 41 pairs registering images in 41 columns. Each column represents mTRE values after registration by three different methods. Every columns consists of one colored marker representing the mTRE value produced by each method. The black squares represent the mTREs before registration. The green triangles represent the mTREs by the proposed method. The red circles show the results for the method discussed in the previous chapter. The blue triangles represent the results for the SIFT-feature-based method. It is found that most of the green triangles have lower values compared to the other two methods. The results are

summarized in Fig. 4.17.(b) in the form of box plots. It is found that the initial mTRE before registration was 5.87 mm, which was improved by the method of the previous chapter to 2.91 with a standard deviation of 2.42 mm. The mean and the standard deviation of mTRE produced by the method in Chapter 3 are 4.79 mm and 2.13 mm respectively. The feature-based method fails to register in most cases because of its poor feature-point localization. The results are shown in the box plot of Fig. 4.17.(b).

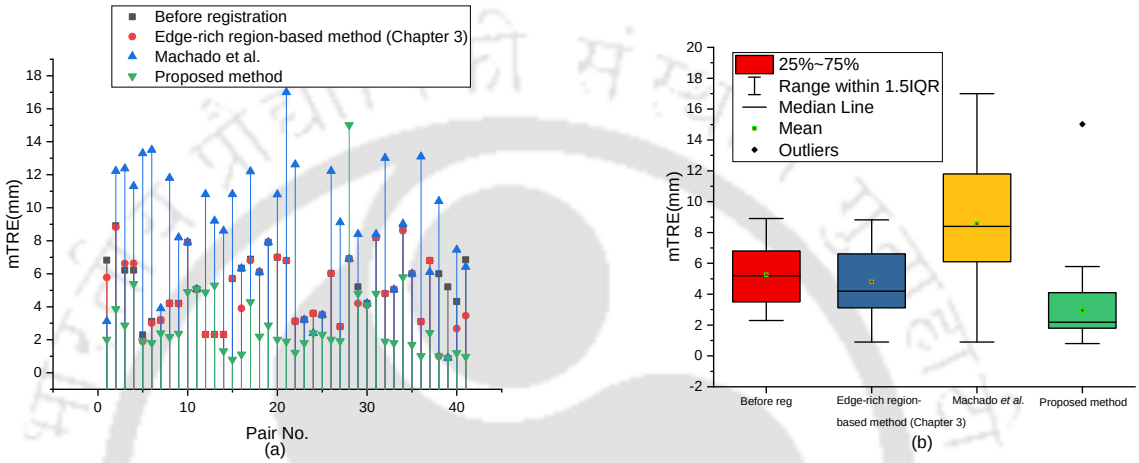


Fig. 4.17: (a)Figure shows mTRE of 41 pairs of images from the BITE and RESECT datasets in 41 columns. Each column represent the result for registration of one image pairs by three different methods. Every column consists of four colored markers representing the mTREs by different method. The red circles represent the mTRE produced by the method in the previous chapter. The blue triangles represent the mTREs produced by the method of Machado *et al.*, the green colored traiangles represent the mTREs produced by the proposed method. [108]. (b) shows 4 box plots. The first box represents the mTRE before registration. Rest three boxes represent the mTRE after registration by the three competing methods

Fig. 4.18 also compares visually the performance of the proposed method with the other methods. Fig. 4.18.(a1-d1) show the pBUS images and Fig. 4.18.(a2-d2) show the corresponding iBUS images. Both the pBUS and iBUS images before registration overlap in Fig. 4.18.(a3-d3). Corresponding image pairs after registration are overlaid in Fig. 4.18.(a4-d4). Registration should reduce mTRE between the corresponding tag points in the registering image pairs. The tag points in the pBUS and iBUS images were shown in red and blue colors respectively. In all the cases, the mTRE value decreases after registration by the proposed method, which reflects better efficiency compared to the other methods.

SIFT-feature-based registration method can hardly find matching features from the pBUS and iBUS image pairs because of the deterioration of the intensity by speckle and artifacts. Fig.

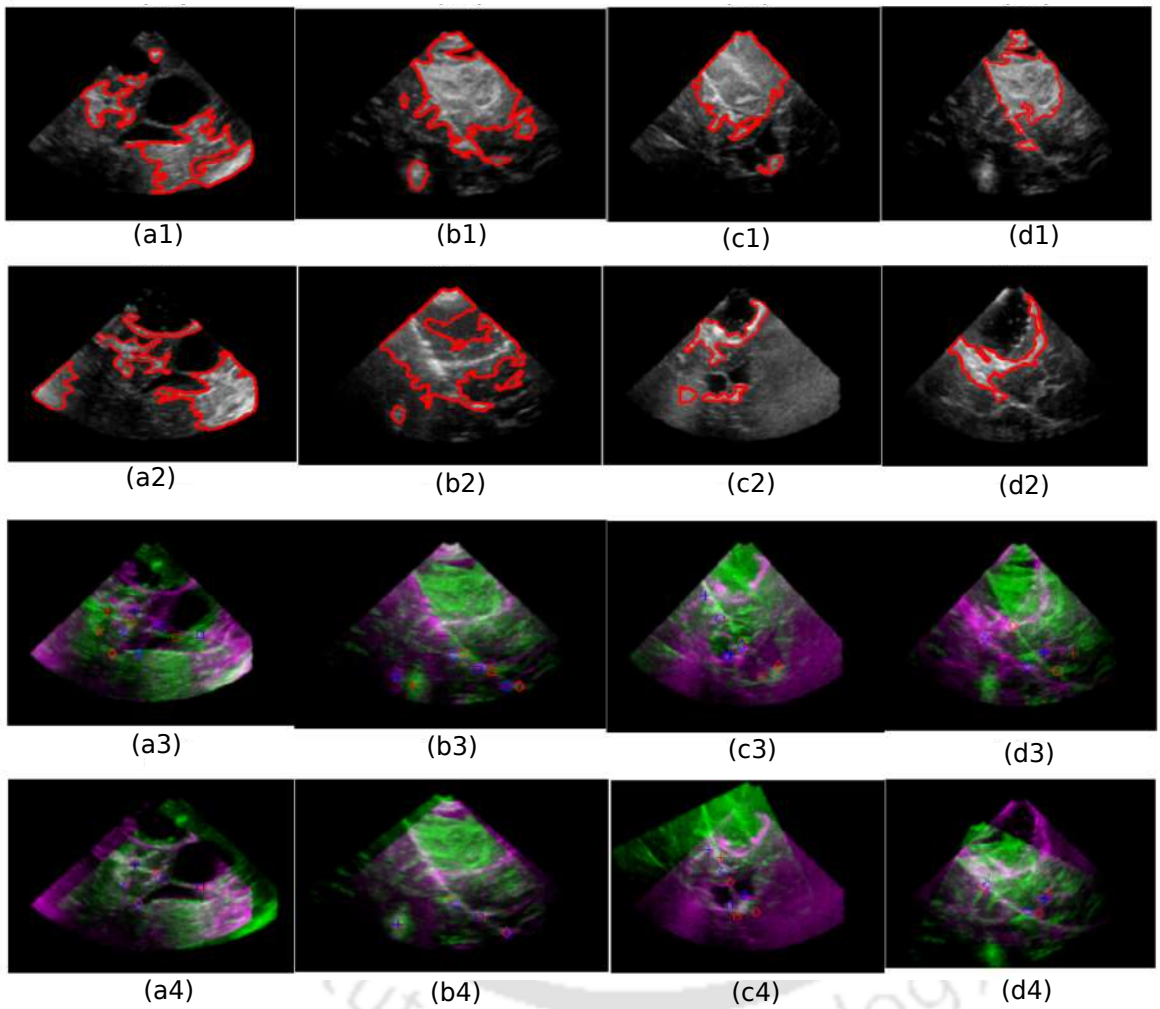


Fig. 4.18: (a1)-(d1) shows the segmented HE regions of pBUS images. (a2)-(d2) shows the segmented HE regions in corresponding iBUS images. The pBUS and corresponding iBUS images and their matching tag-points are overlaid in (a3)-(d3) respectively. Matching tag-points can be identified by the shapes of the markers. (a4)-(d4) show the image pairs after registration with the new position of tag-points in the moving image.

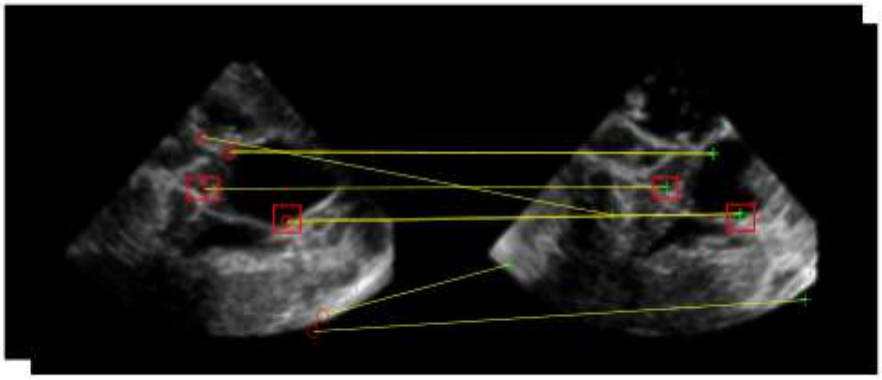


Fig. 4.19: A pair of pBUS and iBUS images is shown. Yellow lines show the mapping of the 2D-SIFT features after feature matching using an exhaustive search technique. Out of five tag points, two points have a match, as shown by red rectangles, and three other points map to the incorrect locations.

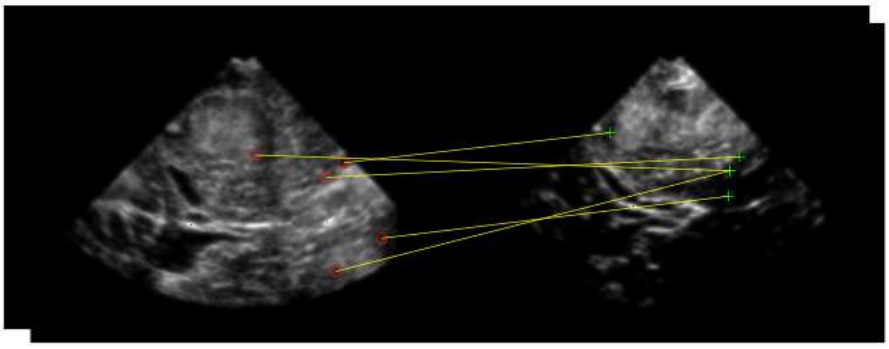


Fig. 4.20: A pair of pBUS and iBUS images is shown. Yellow lines show the mapping of 2D-SIFT features after feature matching using the exhaustive search algorithm. There is no matching of the feature-points, because the iBUS image is suffering from acoustic shadows.

4.19 shows a registering image pair. Two nearly matching tag-points are found from the pair of images, that are not sufficient to perform a successful registration. Another pair of images is shown in Fig. 4.20, where no matching tag-points are found. The mismatch was understood as the effect of speckle and artifacts.

There are some image pairs that consist of similar regions, but those regions remain connected to other HE regions. For example, Fig. 4.21.(a) and (b) show a pair of pBUS and iBUS images, where significant areas have similar intensity profiles. Observing the boundaries of the HE regions, it is found that multiple HE regions are connected together, and therefore it is difficult to achieve similarities between the discrete regions. For those images, manual cropping of the similar regions were required. Those manually cropped one or more rectangular regions from pBUS and iBUS images are copied into blank frames which are registered for obtaining the deformation model. After obtaining the deformation model, it is applied on the full sensed image for registration. It is not essential that the rectangular regions in pBUS and iBUS images are exactly matched; but both the cropped part should include same anatomical regions. Fig. 4.21.(a) and (b) show a cropped rectangular region from pBUS and iBUS images shown in blue coloured boxes. Fig. 4.21.(c) shows the cropped sections of those images in a common frame. Registration is performed based on these cropped images. The deformation model obtained through this registration, is applied on the iBUS image. The image pairs before registration along with the corresponding tag-points are shown in Fig. 4.21.(d) and the positions of the tag-points after registration are shown in Fig. 4.21.(e). Registration reduces the mTRE, and the tag-points are found closer in Fig. 4.21.(e).

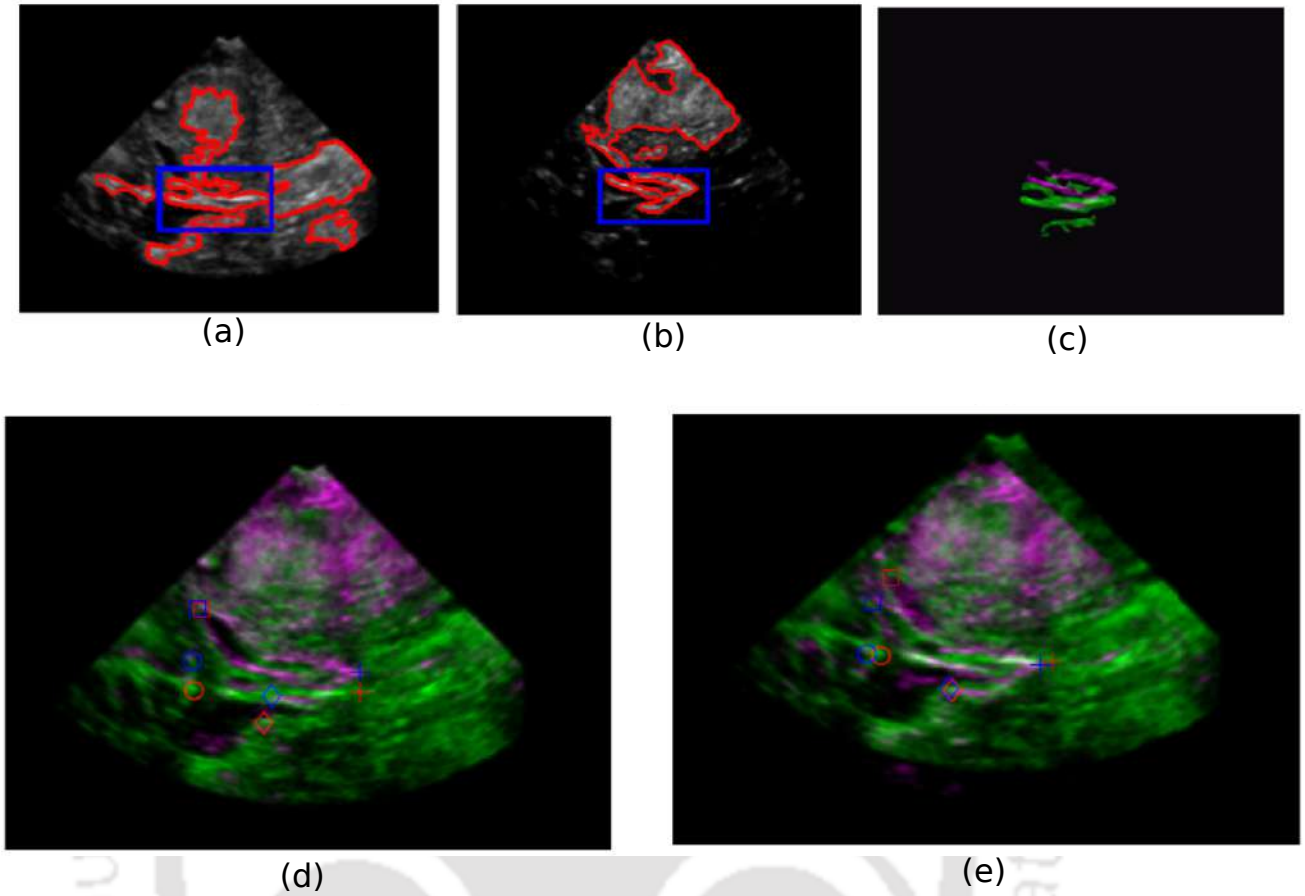


Fig. 4.21: (a) and (b) show the pBUS and corresponding iBUS image. A region shown by a blue rectangle is cropped from both images and placed in blank frames at their unchanged positions that are shown in (c). Corresponding tag points before registration are shown in the overlaid image frame as shown in (d). The images after registration are overlaid in a common frame in (e).

4.5 Success rate comparison of the proposed method with the other methods

The gradient-based registration method described in Chapter 3 applied thresholding of AIG, and $w = 4$ was found optimal. This method has a couple of limitations, such as an increase in the patch size, increased blockiness on the boundaries, and the actual shape of the object gets distorted. Secondly, the threshold values of the AIG chosen were dataset-dependent. They were derived using a statistical analysis of thousands of images from each dataset. But, a constant threshold always does not appear appropriate for many image pairs. Therefore, the

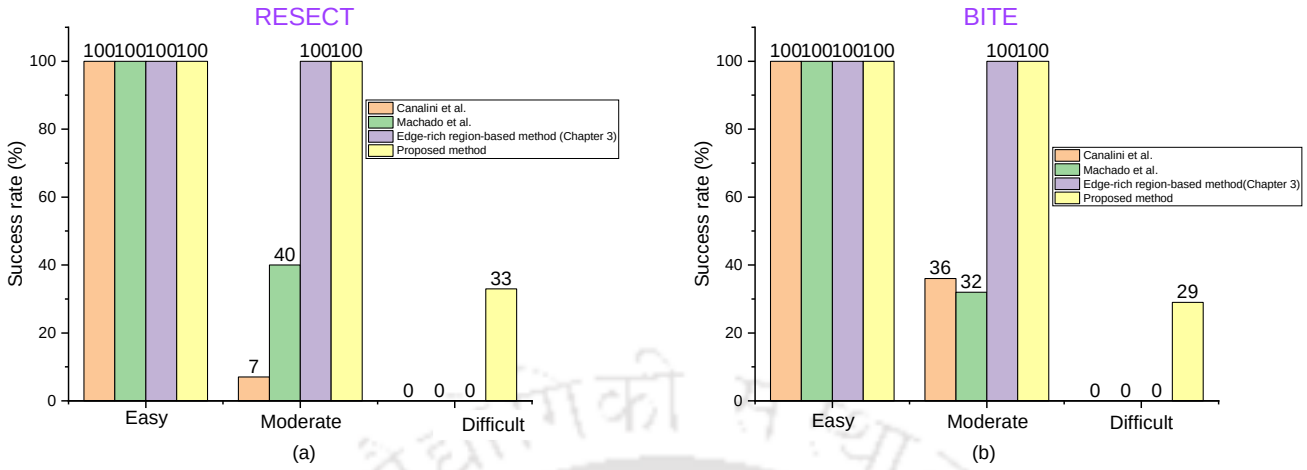


Fig. 4.22: Success rate percentage for the proposed method with the method of Canalini *et al*, Machado *et al*, gradient-based method in Chapter 3, and the proposed method for three different categories of images. The comparative results are shown in (a) for the RESECT dataset, and in (b) for the BITE dataset.

difficult image pairs could not be registered. The proposed method adopts an automatic image-dependent thresholding scheme, which is relevant for different types of regions within a BUS image. On the other hand, it adopts an edge correction technique, that recovers the shape of a region distorted during the gross segmentation. It thus improves the similarity between the common HE regions of the pBUS image and its corresponding iBUS image. Fig. 4.22 shows the improvement in the success of the images in the difficult category. Out of 23 pairs of difficult image pairs, the proposed method could register 7 pairs(4 pairs from the BITE and 3 pairs from the RESECT dataset) of images with a nominal mTRE. Apart from that, the mTRE values for the image pairs of the easy and medium categories were also reduced using this method. The percentage of success rate for each category of images for the BITE and the RESECT datasets are provided in Fig. 4.22. Compared to the methods of Machado *et al.*, Canalini *et al.*, and the gradient-based method in Chapter 3, it performs better by registering difficult image pairs. In Chapter 3, we saw that no methods could register any difficult image pairs. Using the proposed method, the success rate for the registration of the difficult image pairs obtained was 33% for the RESECT dataset and 29% for the BITE dataset. It is a significant improvement over the gradient-based approach presented in Chapter 3.

4.6 Limitations and scope of future work

There are some limitations and challenges in this methods and the performance can be improved by addressing those challenges. The principle challenge is the avoidance of undesired non-matching HE regions in the registering images due to the automatic approach. Some HE regions appear in both images, but they do not have any spatial correlation among them. Hence, the method should be modified for achieving an improved success rate. Both the methods in Chapter 3 and 4 cannot handle artifacts that produce undesired echogenicity and acoustic shadows. The tumor region in pBUS image appears HE but the same region in iBUS image is anechoic because of the RC. It reduces the similarity considerably. The methods discussed in Chapter 3 cannot address this issue. Among the difficult image pairs from the BITE and the RESECT datasets, only seven image pairs could be registered by the proposed method in this chapter, which requires significant improvement.

4.7 Conclusions

This chapter highlighted various challenges in the registration of pBUS and iBUS images. The effect of different artifacts in the iBUS image makes the registration challenging, and therefore the conventional registration methods fail. There are two primary contributions of this chapter. First, a region-driven patch-based segmentation method extracts the HE regions. This segmentation increases the similarity between the registering image pairs. The patch-based approach makes the registration robust to noise, reduces the complexity, and the runtime. It uses simple mean intensity thresholding for obtaining gross boundaries between HE and NHE regions. Blocking artefacts produced by the this step are reduced by an iterative boundary correction process. A curvature-controlled cubic-B-spline interpolation was applied to avoid the gradient and curvature continuity. The segmented boundaries correlate well with the ground truth boundaries. Unlike level-set-based methods, the proposed method does not require any complex energy optimization. The performance of the proposed method was compared with other four level-set-based methods and a random forest classifier-based method. The proposed method outperforms the competing methods in terms of objective metrics and visual results. The second part of the chapter describes a simple image registration method based of the segmented HE regions of pBUS and corresponding iBUS images. It adopts the MSE as the

objective function and applies various heuristic optimization algorithms to obtain the optimum registration parameters. A comparative study among the various optimization methods and competing registration methods showed that the proposed method outperforms the other competing methods and PSO was found as the best-performing optimization algorithm.

5

Semi-automatic registration of pBUS and iBUS images based on fast level-set segmentation of HE and NHE regions.

The methods for the segmentation of BUS images typically target one specific type of echogenic region. The previous chapter proposed an fully automatic segmentation method for extracting the HE regions only. There are some hypo-echoic and anechoic regions in the BUS images, which can work as important landmarks in the registration of pBUS and iBUS images. This chapter presents a fast and semi-automatic level-set method for segmenting the hyper-echoic, the hypo-echoic, and the anechoic regions. The proposed method is a combination of four procedures: a) identification of region type and estimation of region-specific thresholds, b) unidirectional level-set curve flow for obtaining a gross segmentation, c) bidirectional level-set curve flow using a logarithmic patch size control to refine boundaries, and d) cubic B-spline based curve shaping for smoothing the boundaries. A semiautomatic registration method is proposed, where the user marks one or more rectangular regions of interest in both pBUS and iBUS

images. Those regions are segmented by the proposed method and used for registration.

5.1 Introduction

The segmentation of ultrasound(US) images is important in the diagnosis of various tissue abnormalities. Applications include segmenting the fetus [158, 212, 157], the prostate [65], kidneys [189], blood vessels [106, 128], etc. Segmenting a US image is also necessary to estimate the shapes and sizes of the diseased tissues automatically. The previous chapter introduced a fully automatic patch-based method for extracting the HE regions from the pBUS and iBUS images. Generally, the HE regions possess higher similarity between the pBUS and iBUS images. But many image pairs have smaller common HE regions. The extraction of those regions alone cannot ensure much similarity between the registering image pairs. Hence, the registration of an image pair with small common areas often fail. On the other hand, anechoic and hypo-echoic regions present in a pBUS and a iBUS image also contain significant information that helps to obtain better registration results. The inclusion of the anechoic and the hypo-echoic regions, improves the similarity between the registering pBUS and iBUS images. This chapter focuses on the segmentation of specific regions with different level of echogenicity. The patch-based method described in the previous chapter does not target any specific regions and grossly find the HE regions, which may include various HE artifacts. This chapter presents a method which is semi-automatic, requires minimal user intervention to mark the desired region for segmentation.

HE regions produce large echoes and appear bright. There are a number of hyper-echoic regions in a BUS image, such as the CP, tumors, LF, and corpus callosum (CC) [123]. In a B mode US image, nerves also reflect large echos [88]. All four ventricles, sub-arachnoid space, blood vessels, cyst, etc. appear anechoic. Some other regions that show moderate echogenicity like thalamus, caudate nucleus, cerebrum, etc. are found to be hypo-echoic regions. The method described in Chapter 4 extracts all the HE regions together, that may include various artefacts such as, border-line artefacts and comet tail artefacts. It does not consider any specific HE region for segmentation. This chapter highlights the importance of extracting selective hyper-echoic, hypo-echoic, or anechoic regions from both the pBUS and the iBUS images. Extracting multiple types of regions increases the common area between the registering images and thus improves in registration performance.

5. Semi-automatic registration of pBUS and iBUS images based on fast level-set segmentation of HE and NHE regions.

The segmentation of a US image is challenging because of poor resolution, the presence of speckle, acoustic shadows, and unnatural scattering due to the presence of blood clotting agents in the RC [124, 123]. Traditional segmentation methods [69] like region growing, region splitting, morphological segmentation etc. have limited applicability due to the intensity inhomogeneity of the US images. The state-of-the-art approaches to the segmentation of US images were discussed in Chapter 2. In this chapter, we briefly mention a few of them, that have connections to our proposed methodology.

Osher and Sethian [146] first proposed the level-set method for curve evolution for various applications. This approach was employed largely in US image segmentation and related discussion was presented in Section 4.1.1. It consists two major steps: *energy function formulation* and *level-set curve evolution*. There are various approaches to level-sets, such as GAC [24], ACWE [28], and DRLSE [98]. For the segmentation of speckle-free images, the performance of these methods is satisfactory; however, when speckle increases, the performance of these algorithms degrades. Fig. 5.1 shows a synthetic image and speckle of difference variances is added to it. The speckle was simulated in MATLAB using the methodology outlined in [212]. Each image contains five shapes having different intensity values and intensity gradients on their boundaries. The hexagonal region has highest and the pentagonal region has the lowest intensity gradients at the boundaries. Higher intensity gradient makes the segmentation easy. The segmented outputs of the GAC, ACWE, and DRLSE on speckle-free images are shown in Fig. 5.1. (a1), (b1), and (c1) respectively and are marked by the red-colored boundaries. These figures show that all the three algorithms successfully segment all five zones in the absence of speckle. However, as the speckle variance increases, their performance degrades. The outputs for a speckle variance of 0.01 are shown in Fig. 5.1. (a2), (b2), and (c2), which show that ACWE and GAC segment better than the DRLSE algorithm. (a3), (b3), and (c3) show the outputs for speckle variance of 0.1. It is seen that the DRLSE fails to segment all the regions. ACWE and GAC roughly segment the regions other than the pentagonal region, but the accuracy of the boundaries is not satisfactory.

Based on these findings, we conclude that the level-set-based approaches are significantly affected by speckle. Level-set curve evolution is a state-of-art approach for region-based image segmentation. Our proposed solution uses a patch-based level-set-based strategy that is fast,

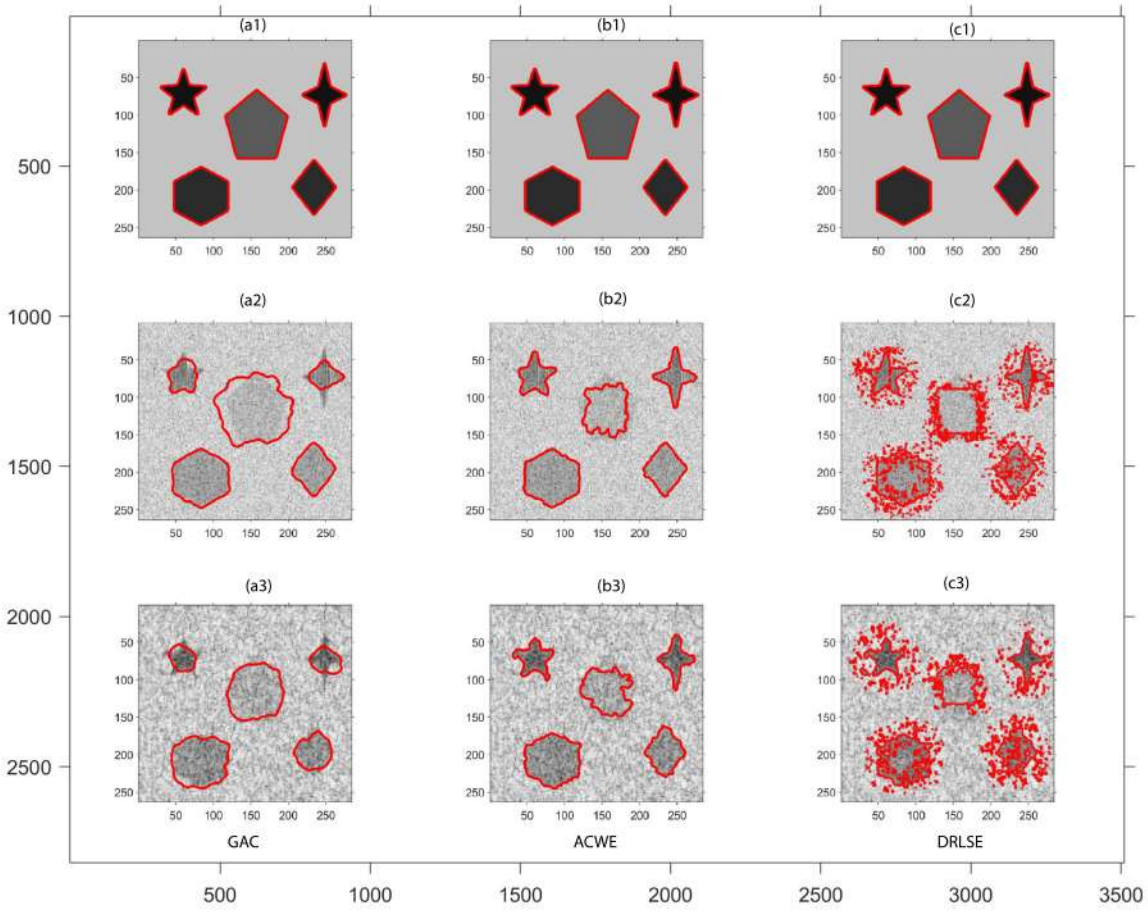


Fig. 5.1: (a1), (b1), and (c1), are specklefree phantom images. The results of the various segmentation algorithms are represented by red boundaries across these images. The GAC method's outputs are (a1) to (a3), the ACWE method's outputs are (b1) to (b3), and the DRLSE method's outputs are (c1) to (c3). The outputs for the speckle of variance 0.01 are shown in rows (a2)-(c2), while the outputs for the speckle of variance 0.1 are shown in rows (a3)-(c3).

speckle-resistant, and yields better accuracy. It is topology-independent and produces regular-shaped smooth boundaries of the desired regions. The detailed description of the proposed method is provided in the next section.

5.2 Proposed method

The statistical nature of the intensity profile of a segmenting region is not taken into account by the existing level-set-based methods. Anechoic, hypo-, and hyper-echoic regions in BUS

images have distinct statistical properties. Hence, a technique that effectively segments an anechoic region may not work for a hyper-echoic region. The proposed approach reads the local statistics of a image region, and offers a strategy that works equally well in all three kinds of regions. The level-set-based image segmentation framework serves as the foundation for the proposed work. Its complexity is avoided by excluding the energy minimization step. This step is substituted by a simple patch-based intensity thresholding. Three main issues are addressed here:

- Stop curve evolution at the boundary.
- Avoid boundary ambiguity caused by speckle.
- Obtain a smooth boundary curve.

As discussed earlier, the spatial curve gradient, whose magnitude depends on the image intensity profile, determines the direction of the iterative growth of the initial level-set curve. The boundaries of speckle-filled images are uneven; hence, a move along the normal direction makes the boundary more distorted. To prevent this, the length and the area covered by the boundary, are added to the energy function as regularizers. The proposed method does not apply any energy optimization, but it adds a contour smoothing step to ensure a smooth boundary of the segmenting regions.

We described the advantages of the patch-based thresholding over pixel-based thresholding in the previous chapter. The method proposed in this chapter also adopts patch-based thresholding in various steps. The steps of the proposed method are shown in Fig. 5.2. At the start of the segmentation process, the user marks a rectangular region of interest(ROI) denoted by Ω_R. The selection is made in such a way that Ω_R contains only one type of echogenic region. The region type of Ω_R is identified by using a thresholding method as described in Section 5.2.1. The next step applies a patch-based thresholding of Ω_R. The statistical natures of three types of echogenic regions are different, and therefore they need different formulations for the estimation of region specific threshold ranges. For the segmentation of hyper-echoic, hypo-echoic, and anechic regions, a lower threshold T_l and an upper threshold T_h are needed to determined. The formulation of T_l and T_h for the three different types of echogenic regions are described in Section 5.2.1.1. The first step of the level-set curve evolution is the estimation of the ini-

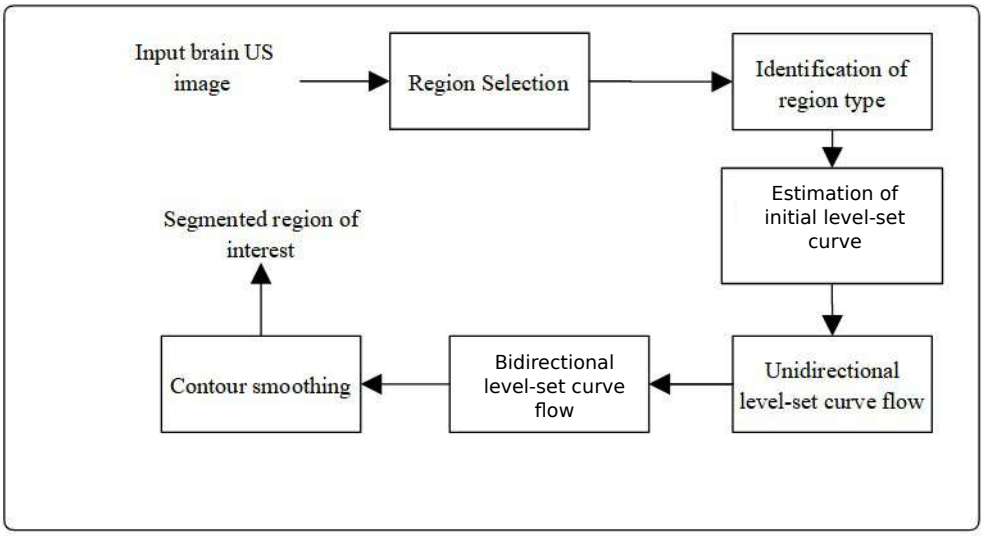


Fig. 5.2: Proposed level-set framework for the segmentation of BUS images

tial level-set curve $\phi(0, x, y)$, which is described in Section 5.2.2. Section 5.2.3 describes how $\phi(0, x, y)$ expands over time by a patch-based unidirectional level-set curve flow (ULSCF) step. This step results in a gross segmentation of the region; however, the resulted boundaries suffer from blockiness due to the patch-based processing. The larger is the size of the patches, the higher is the blockieness. The bidirectional level-set curve-flow (BLSCF) step adopts a patch-based bi-directional level-set evolution, that increases the proximity of the resulting boundaries to the actual boundaries as described in Section 5.2.4. The resulting boundary after the boundary correction is not smooth and suffers from slope and curvature discontinuities. It is rectified by a contour shaping step, which is described in Section 5.2.5.

5.2.1 Identification of region-type

For minimal user intervention, automatic identification of the region type is necessary. Three different types of regions can be identified by the statistics of the intensity profile of the region. First, we find a range of the mean intensities of the regions. For ascertaining the range of mean intensities of hyper-echoic, hypo-echoic and anechoic regions, rectangular sections are cropped from the hyper-echoic, the hypo-echoic, and the anechoic regions of 160 images from the BITE [123] and the RESECT [202] datasets. Assume a rectangular region Ω_R is marked on an image

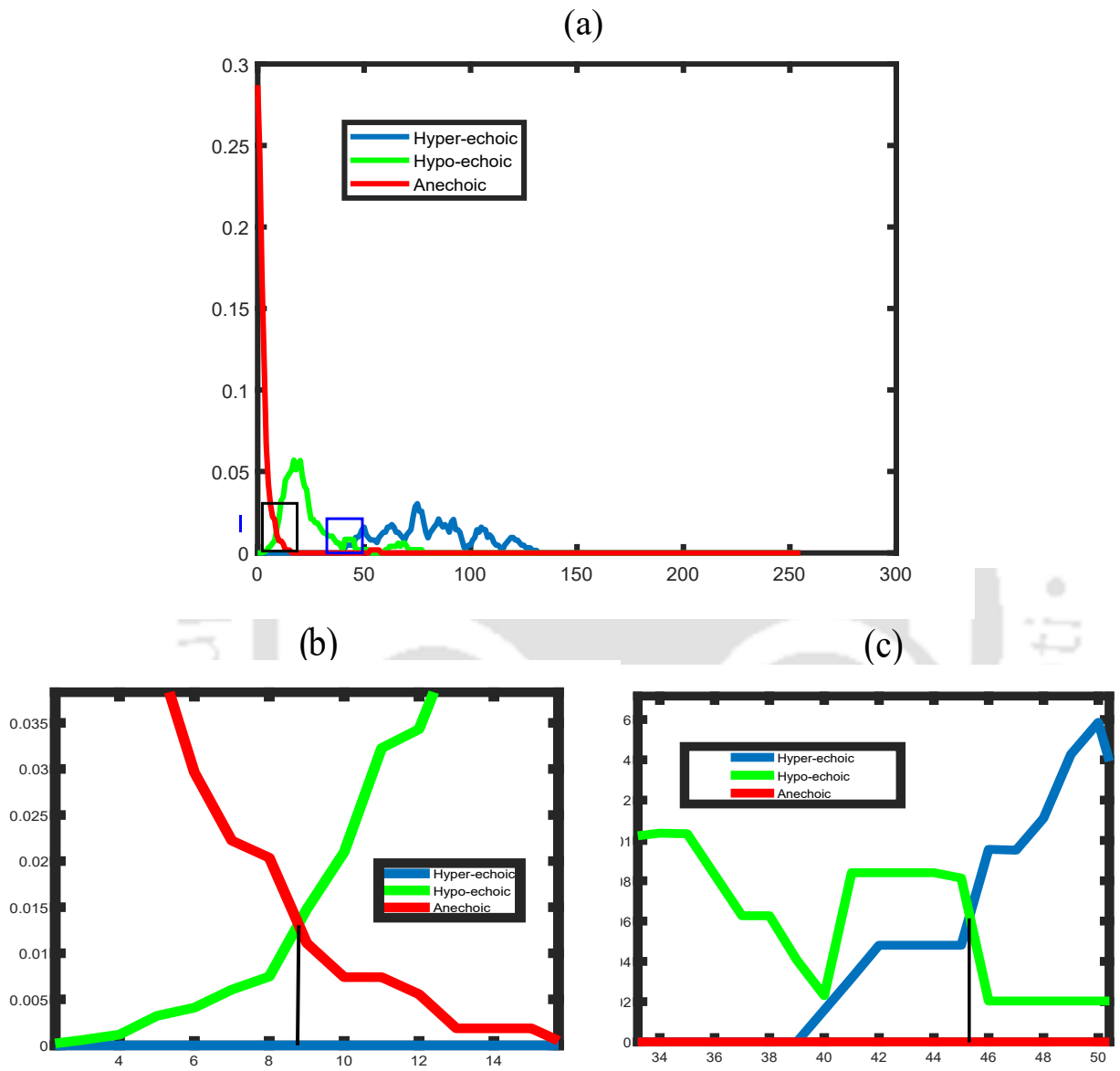


Fig. 5.3: (a) Intensity histogram of the three different echogenic regions. Colored boxes highlight the intersection between different histograms. (b) is the enlarged box that highlights the intersection between the anechoic and the hypo-echoic histograms. (c) is the enlarged box that highlights the intersection of the hyper-echoic and the hypo-echoic histograms.

V. The mean intensity (μ_R) of a chosen rectangular region Ω_R can be obtained by

$$\mu_R = \frac{1}{|\Omega_R|} \sum_{(x,y)\epsilon\Omega_R} V(x,y) \tag{5.1}$$

where $|\Omega_R|$ is the number of pixels in Ω_R. Ω_R is chosen in such a way that it covers the maximum possible area of a region. The μ_R values for each types are plotted in a normalized histogram as shown in Fig. 5.3.

We observed that for an 8-bit image, the mean intensity of the ROI varies between 0-15 for an anechoic region, 4-75 for a hypo-echoic region, and 45-150 for an HE region. The types of regions are determined by comparing the average intensity (μ_R) of the user-defined region to the said ranges of intensities. The thresholds are decided using the maximum likelihood principle. Fig.5.3.(b) highlights the intersection between the anechoic and the hypo-echoic regions, and it was found that both the histograms intersect at 8.2. Thus, we set a threshold value of 8.2 between the anechoic and the hypo-echoic regions. From Fig. 5.3.(c), we find the intersection between the normalized histograms of the hyper-echoic and hypo-echoic regions, and 45.6 was obtained as the demarcation level between these two types of regions. Hence, if $\mu_R \leq 8.2$, the region can be labeled as anechoic. Similarly, if $8.2 \leq \mu_R \leq 45.6$, it is hypo-echoic, else it is HE.

The proposed method adopts a patch-based level-set strategy for the segmentation of the desired region. The user-selected region Ω_R partially includes the desired region. The part of Ω_R that consists of the desired region is defined as the inside region and the rest is termed as the outside region. The growth of the segmentation of the desired region starts from the initial level-set curve. This step divides Ω_R into multiple square patches of size $w \times w$, where w is chosen to be either 8 or 16. Depending on the region type, two thresholds T_h and T_l are estimated based on the statistical nature of Ω_R. The estimation of T_l and T_h is described in the following subsection.

5.2.1.1 Estimation of T_h and T_l

Once the region-type is identified, the pixels in Ω_R are to be labeled as members of the inside or the outside region. As in the previous chapter, this association is done patch wise by comparing the average patch intensity with two thresholds T_l and T_h. These thresholds are used for the patchwise association of the pixels to the desired region. Noting that the standard

deviation controls the spread of the data, we formulate the T_l and T_h as

$$T_l = \mu_R - k_1\sigma_R \tag{5.2}$$

$$T_h = \mu_R + k_2\sigma_R \tag{5.3}$$

where σ_R is the standard deviation of the image in Ω_R, k_1 and k_2 are constants. The values of k_1 and k_2 are determined from the data for different region types and thus, T_l and T_h are formulated as follows:

1) *Anechoic region*

$$\begin{aligned} T_l &= 0 \\ T_h &= \mu_R + 0.5\sigma_R \end{aligned} \tag{5.4}$$

2) *Hypo-echoic region*

$$\begin{aligned} T_l &= max(0, \mu_R - \sigma_R) \\ T_h &= min(255, \mu_R + 1.5\sigma_R) \end{aligned} \tag{5.5}$$

3) *Hyper-echoic region*

$$\begin{aligned} T_l &= max(0, \mu_R - 0.5\sigma_R) \\ T_h &= min(255, \mu_R + 2.0\sigma_R) \end{aligned} \tag{5.6}$$

The *max* and *min* terms in (5.5) and (5.6) are used to limit the values of T_h and T_l with the dynamic range of intensity 0-255 for 8 bit images. Both T_l and T_h have an important role in determining the initial level-set curve $\mathbf{C}_0$ and estimating the level-set function $\phi(n, x, y)$ as described in Sections 5.2.2 and 5.2.3 respectively. HE artifacts have very high echogenicity. If the mean intensity of a HE patch is higher than T_h, it is treated as an artifact, and is not included in the HE region. Thus this approach suppresses HE artifacts also.

5.2.2 Estimation of initial level-set curve

Let an input image $\mathbf{V}$ be defined on the domain Ω and $\Omega_R \subseteq \Omega$ be the initial ROI of size $M \times N$ marked by the user. Let Ω_R be divided into non-overlapping square patches $\mathbf{P}(i,j) : 1 \leq i \leq K$ and $1 \leq j \leq L$ of size $w \times w$. The value of w is an integer power of 2. The region Ω_R is extended minimally so that M and N are integer multiples of the patch size $w \times w$. The mean intensity of $\mathbf{P}(i,j)$ is given by

$$\mu(i,j) = \frac{\Sigma\Sigma_{(x,y)\in\mathbf{P}(i,j)} V(x,y)}{w^2} \tag{5.7}$$

$\mu(i,j)$ is compared with T_l and T_h. If $T_l \leq \mu(i,j) \leq T_h$, $\mathbf{P}(i,j)$ is considered as inside patch, and included in Ω^{inside}. Hence, the inside region, extracted from Ω_R can be expressed as,

$$\Omega^{inside} = \bigcup_{\forall(i,j):T_l\leq\mu(i,j)\leq T_h} \mathbf{P}(i,j) \tag{5.8}$$

The boundary of the region Ω^{inside} is the initial level-set curve $\mathbf{C}_0$. We define a level-set function $\phi(n,x,y)$ as follows.

$$\phi(n,x,y) = \begin{cases} = 1 & (x,y) \in \Omega^{inside} \\ = -1 & (x,y) \in \Omega^{outside} \end{cases} \tag{5.9}$$

We define the zero level-set function as the interface between $\phi(n,x,y) = 1$ and $\phi(n,x,y) = -1$ regions. In Fig. 5.4, the gray patches satisfy the threshold condition. The boundary of this region represents $\mathbf{C}_0$ that encompasses a part of the desired segmenting region. For a region with complex shape, multiple rectangles can also be marked. Overlapping between these rectangles is also allowed during this step. The number of regions and their sizes play important roles in the overall speed of the segmentation. A large rectangular region may include many undesired patches, and corresponding μ and σ also change a lot due to these patches compared to its actual value. It may results into an ineffective segmentation. On the other hand, multiple rectangles increase user intervention. For example, Fig. 5.4.(b) shows two rectangular regions

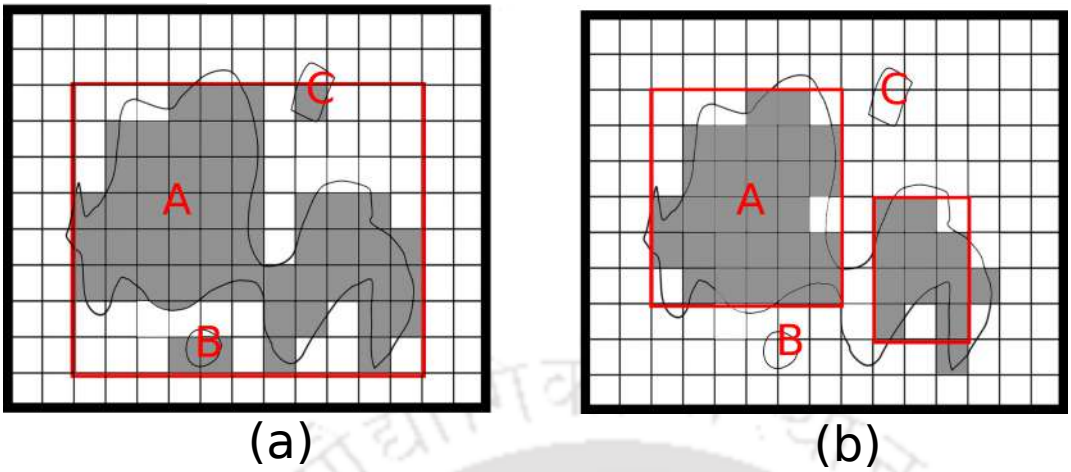


Fig. 5.4: Ω_R is marked in red color to select the ROI. In (a), the ROI encompasses multiple objects. In (b), the ROI is divided into two rectangles.

and the ϕ maps of the regions. The same segmentation can be performed with any one rectangle also. But it takes more iterations during the level-set curve flow; thus the overall complexity increases.

5.2.3 Unidirectional level-set curve flow(ULSCF)

As discussed in Section 5.2.2, all the pixels in a $w \times w$ patch have the same value of ϕ, and thus are expressed patch-wise. Let $\phi(n, i, j)$ be the level-set function value of $(i, j)^{th}$ patch at iteration n. The patch $\mathbf{P}(i, j)$ is given by $\mathbf{P}(i, j) = \{(x, y)|(i-1)w \leq x \leq (i+1)w - 1$ and $(j-1)w \leq y \leq (j+1)w - 1\}$. In the proposed method, $\phi(0, i, j)$ is obtained by threshold estimation and comparison as described in Section 5.2.2. The initial level-set curve flow starts from $\phi(0, i, j)$, and it moves toward the outward direction from $\mathbf{C}_0$. This step is iterative and in every iteration, the boundary moves to the outward direction. Suppose $\mathbf{C}_k$ is the boundary at k^{th} iteration. A one patch wide neighborhood N_k is defined in the outside region of $\mathbf{C}_k$. If a patch in N_k satisfies the condition $T_l \leq \mu(i, j) \leq T_h$, its ϕ value changes from -1 to +1, and the patch is included in the inside region. It is similar to the pixel-based flow in [182], that drives the boundary in the outward direction. Unlike the method in [182], the flow in this step is allowed to one patch outward per iteration from $\mathbf{C}_k$. This movement is allowed to all outward neighboring patches of $\mathbf{C}_k$. This approach is faster because of the bigger movement of the boundaries. Secondly, patchwise averaging has a smoothing effect that makes it robust to

noise. The flow of the level-set function values over time can be expressed as

$$\phi(n+1,i,j) = \phi(n,i,j) + 2\underbrace{H(\mu(i,j) - T_l)H(T_h - \mu(i,j))}_{force function} \quad \forall (i,j) \in N_n \tag{5.10}$$

where $H(.)$ is the Heavieside function. The change in the value of ϕ depends on a force function. The force function is a Heavieside function which results into unity when a patch satisfies the threshold condition. The level-set curve flow is comparatively faster at a homogeneous region and becomes slow when approaching boundary of the level-set curve [182]. The flow in the proposed method is uniform, and grows with larger step size. Fig. 5.5.(a) shows the growth of ϕ during this step by the cyan color. It takes mostly seven iterations to reach the boundary of the desired region. If 2 rectangles marked by red colored boundaries are considered in the beginning as shown in Fig. 5.4.(b), the same region can be covered within one iteration only. It describes about the flexibility in ROI selection, that can reduce the execution time. Like other conventional level-set methods, this method can join multiple growing regions during this step. It can also split from the initial ROI to multiple regions during initialization of ϕ_0. These properties make this method topology-independent.

The evolution of the $\phi = 0$ stops when there is no change in ϕ. In other words, iteration continues until $\Sigma_{\forall(i,j)} |\phi(n,i,j) - \phi(n-1,i,j)| = 0$

5.2.4 Bidirectional level-set curve flow(BLSCF)

For obtaining smooth boundaries, other level-set methods add the curvature to the energy function as a regularization term. These methods require extra calculations in every iteration for the estimation of the curvature of $\mathbf{C}_k$. We ensure curve smoothness in a simple and fast manner. The ULSCF described in the previous subsection performs patch-wise segmentation in the selected region and results in block-like boundaries. Let the stopping criterion be met after J iterations, and it produces a blocky boundary $\mathbf{C}_J$. To reduce the blockiness of the obtained boundary, both the inside and the outside patches connected to $\mathbf{C}_J$ are needed to be observed closely on a finer scale. Fig. 5.5.(a) shows the graphical representation of $\phi(J,i,j)$ after the ULSCF for a patch size of $w \times w$. Here the gray patches represent the initial inside region. Before starting the level-set curve flow, both white and cyan patches remain in the outside

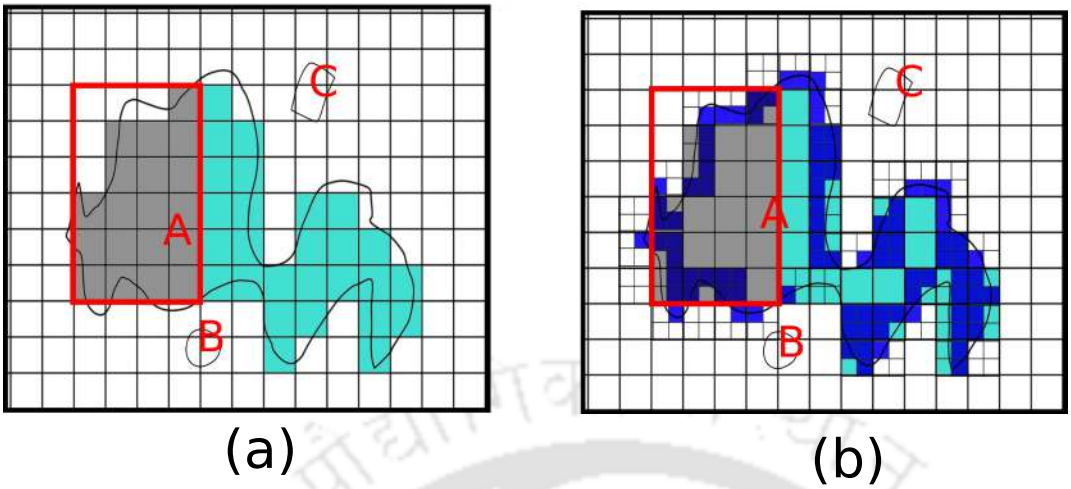


Fig. 5.5: (a) ULSCF starts from the boundary of the gray region. The *cyan* region of ϕ is grown during uni-directional level-set curve flow. (b) During boundary correction, every patch connected to C_0 is divided into 4 equal-sized square subpatches. If a subpatch maintains the prescribed threshold condition, is remained/added to $\phi = 1$ region, and otherwise its value become $\phi = -1$. This modification is shown in blue color.

region. Out of those patches, the cyan patches are within the desired region which is needed to be labeled as inside, and the white patches are needed to be labeled as outside patches. For all the inside patches, the value of $\phi = 1$; $\phi = -1$ at the outside patches.

In this step, $\mathbf{C}_J$ obtained in the ULSCF step is reinitialized as $\mathbf{C}_0$, which is an approximate shape to the desired boundary. The neighborhood of $\mathbf{C}_0$ is also redefined as N_0, which is the union of both inside and outside patches connected to $\mathbf{C}_0$. Unlike the ULSCF, both inward and outward movement is allowed in this step. It is an iterative process of $1 + log_2(w)$ iterations. In each iteration, the value of w is made half of the value of it in the previous iteration. Therefore, after $1 + log_2(w)$ iterations, it turns into a pixel-based processing. The following steps are performed in each iteration

BLSCF

1: Input w

2: **while** $w > 1$ **do**

3: set k=0

4: Determine the bidirectional neighborhood N_k of $\mathbf{C}_k$. $\mathbf{C}_k$ is the union of the 8-connected patches of size $w \times w$. First, all patches having a transition of ϕ from -1 to +1 or vice versa. and shared at least one side with $\mathbf{C}_k$, are identified. For each of these patches, the bidirectional neighborhoods are included in N_k.

5: Divide each $w \times w$ patch $\mathbf{P}(i,j)$ in N_k that are connected to $\mathbf{C}_k$ into four subpatches of size $\frac{w}{2} \times \frac{w}{2}$. Assign each sub-patch with the ϕ value of the $w \times w$ patch. Update the values of $\phi(k,i,j)$ as follows

6: Set $w = \frac{w}{2}$

7: Compute $\mu(i,j)$ using (5.7)

8: **if** $\phi(k,i,j) = 1$ **then**

9:

$$\phi(k,i,j) \quad = \quad \phi(k,i,j) \quad - \quad 2\left[1 - H(\mu(i,j) - T_l)H(T_h - \mu(i,j))\right]$$

10: **else** $\phi(k+1,i,j) = \phi(k,i,j) + 2H(\mu(i,j) - T_l)H(T_h - \mu(i,j))$

11: **end if**

12: $k \leftarrow k + 1$

13: **end while**

Fig. 5.5.(b) shows the modification to the existing boundary patches shown by blue colors. All the blue-colored patches satisfy the threshold conditions and are included in the inside region. Similarly, if a sub-patch does not satisfy the threshold condition, $\phi(k,i,j)$ becomes -1.

The boundary generated by the ULSCF comprises $w \times w$ patches, which indicated poor resolution at the boundary. The actual boundary remains hidden within this patch of $w \times w$. During BLSCF, the patch size reduces in a dyadic scale that ensures to reach to the actual boundary within $log_2(w)$ iterations. It is illustrated graphically in Fig. 5.6.

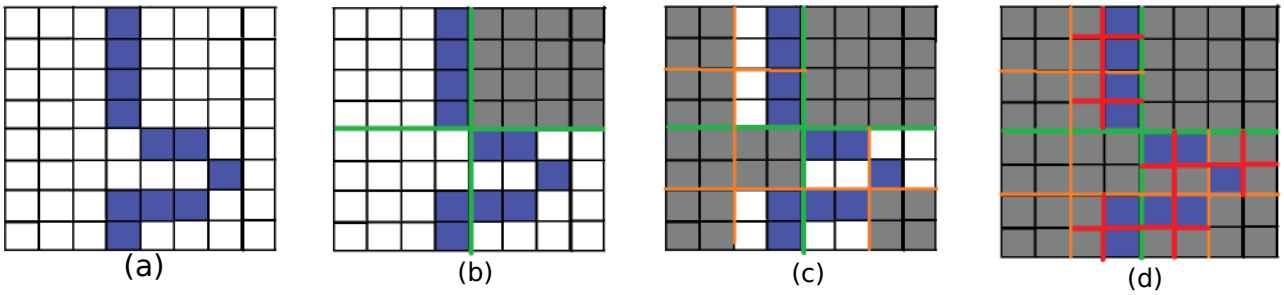


Fig. 5.6: (a) shows a 8×8 boundary patch after ULSCF. The connected blue-coloured pixels show the one-pixel-wide actual boundary. A gray sub-patch indicates that there is no boundary pixel within it. (b) shows the top right 4×4 sub-patch does not main threshold condition and turned gray. (c) shows the resulting boundary with a patch size of 2×2, and (d) the final boundary is obtained after boundary correction with patch size 1×1.

Fig. 5.6.(a) represents an 8×8 ($w = 8$) patch connected to $\mathbf{C}_0$ that resulted after the ULSCF. Initially, the whole patch represents a boundary patch. The actual boundary is somewhere within the patch. Suppose the connected blue-colored pixels represent the actual boundary as shown in Fig. 5.6. In the first iteration of BLSCF, the 8×8 patch is divided into four 4×4 patches. If any 4×4 patch does not satisfy the threshold condition, it is marked gray as shown in Fig. 5.6. (b). This step is repeated for the patch size of 2×2, and the corresponding result is shown in Fig. 5.6. (c). Fig. 5.6.(d) is the last step with $w = 1$.
It is observed that after $log_2(8) = 3$ iterations, the boundary matches with the actual boundary. The improvement of the boundary shape can also be seen on an ultrasound image shown in Fig. 5.7. Fig. 5.7.(a) shows a breast tumor, and Fig. 5.7.(b) shows the ground truth boundary as provided in the database [83]. Fig. 5.7.(c) shows the boundary after the ULSCF. In Fig. 5.7 (d-g), boundaries are produced for $w = 8, 4, 2$ and 1 respectively. The results clearly show that this step increases smoothness of the boundary with the decrease in w.

5.2.5 Contour smoothing

After BLSCF, ϕ values of the inside and outside regions are 1 and -1 respectively. We define a boundary contour as a one-pixel-wide curve along the inside region. This boundary suffers from irregularity in shapes due to its slope and curvature discontinuities. The smoothness of the boundary is an important requirement satisfied by other level-set-based segmentation methods like AC[87], ACWE [28], WLSE [90], the method by Shi *et al.* [182], and the method by Li *et al.* [98]. These methods include the length and the curvature of the level-set curve as

the regularization terms in the energy function that ensures the smoothness of the boundary. Similar to the earlier method in Chapter 4, this work also uses a cubic B-spline-based contour smoothing approach. This step approximates the noisy boundary curve to a regular-shaped B-spline curve. This step requires a curvature threshold and a minimum distance threshold between two consecutive nodes which were kept the same as used in Chapter 4. A detailed description of this step can be found in Section 4.2.1.

5.3 Results on the proposed level-set-based segmentation

The proposed method segments HE and anechoic regions of BUS images from the BITE and the RESECT datasets, because these regions are clinically important. The performance of the proposed method is compared with various level-set-based methods. We used this method for segmenting two anechoic regions, namely ventricles and RC, and three HE regions LF, CP, and tumors from over1000 BUS images.

For quantitative evaluation, we prepared a total of 192 different ground truth brain US image sections of desired regions from 120 US images from the BITE and 72 images from the RESECT datasets. Images were prepared using an open-source MATLAB tool *imageannotationBot* [39] in coordination with the radiologist. Images with better visibility of regions were chosen for preparing the ground truth images.

For performance comparison, six different level-set-based methods, namely GAC [24], ACWE [28], DRLSE [98], the method of Shi-Karl [182], WLSE [90] and the method by Liu *et al.* [101] were chosen. The outputs of BUS image segmentation with GAC, ACWE, DRLSE, and the method of Shi *et al.* are found by open-source tool CREASEG [50]. The implementation of the methods by Liu.*et al.* [101] and WLSE [90] was done in MATLAB. The results are presented both graphically and parametrically in terms of region-based quality metrics like accuracy, precision, Dice, and some edge-based metrics like the asymmetric squared contour distance (ASCD), and the *root mean squared contour distance* (RMSCD) [173]. The proximity of these boundaries with their respective ground truths are measured by ASCD and RMSCD values. The formulation of the said quality metrics is described in Section 4.2.5.

The results for the five different regions LF, CP, tumor, ventricles, and RC are shown in Table 5.1, Table 5.2, Table 5.3, Table 5.4, and Table 5.5 respectively. Table 5.1 shows the

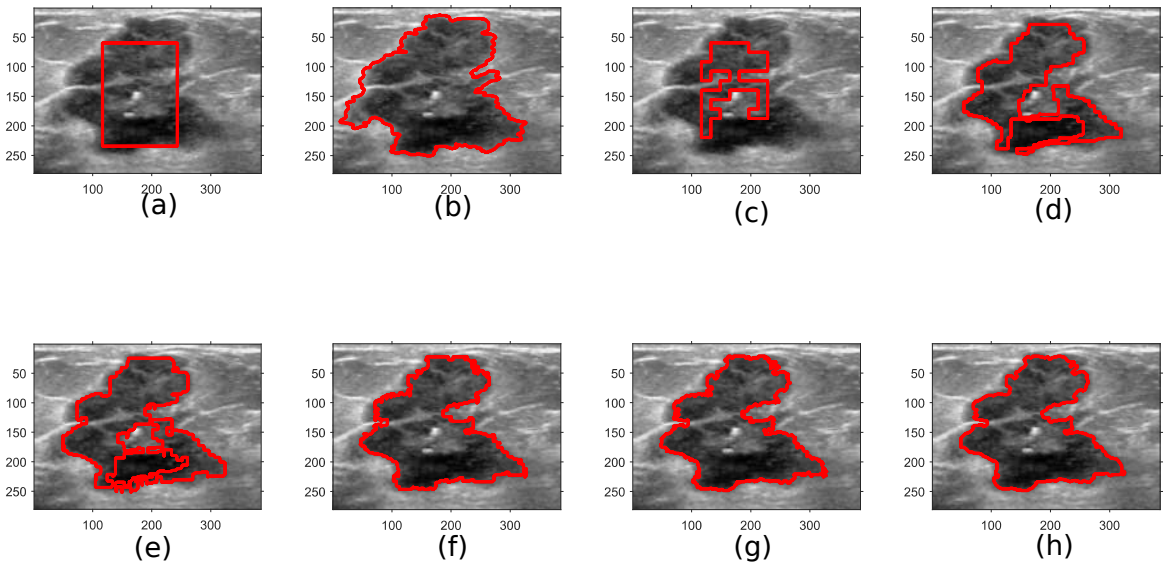


Fig. 5.7: (a) Original image with initial boundary selected by the user. (b) Ground truth labeled boundary. (c) Resulting boundaries after ULSCF. Resulting boundary after boundary correction with (d) $w = 8$. (e) $w = 4$. (f) $w = 2$, and (g) $w = 1$. (h) Final contour after B-spline contour smoothing.

average values of the quality metrics for the segmentation of LF. It shows that the proposed method outperforms other methods in terms of the quality metrics other than RMSCD. Table 5.2 shows that the proposed method segments CP better than the other methods in terms of the average precision, ASCD, and RMSCD.

The proposed method also resulted better segmentation of the tumor regions as shown in Table 5.3. Table 5.1, Table 5.2, Table 5.3, and Table 5.5 show that in most cases, the proposed method show better metrics compared to other methods. The method by Nitsch *et al.* resulted better accuracy for the segmentation of LF and better average Dice values for the segmentation of CP. ACWE showed better accuracy for the segmentation of the tumor regions. But accuracy is not a good measure always to examine the performance of segmentation. Higher accuracy can also generate for over-segmentation and under-segmentation. ASCD and RMSCD are comparatively better measures, and the proposed method resulted in better values for these metrics.

The proposed method is equally efficient in segmenting the anechoic regions like a ventricle and RC as shown in Table 5.4 and Table 5.5 respectively. Barring precision, all other quality metrics were better for the proposed method of ventricle segmentation. Similar results can be

Method	Precision	Accuracy	Dice	ASCD (mm)	RMSCD (mm)
GAC	0.365	0.551	0.525	4.044	6.593
ACWE	0.39	0.692	0.538	5.471	9.390
DRLSE	0.398	0.567	0.558	5.320	12.851
Shi-Karl's method	0.353	0.654	0.496	6.649	12.552
WLSE	0.367	0.59	0.599	5.057	12.101
Method by Liu *et al.*	0.416	0.605	0.568	2.950	4.974
Method by Nitsch.*et al.*	0.396	**0.728**	0.599	4.531	5.581
Proposed Method	**0.496**	0.708	**0.649**	**2.794**	**4.508**

Table 5.1: Performance comparison of different methods for the segmentation of LF

Method	Precision	Accuracy	Dice	ASCD(mm)	RMSCD(mm)
GAC	0.355	0.470	0.496	5.666	8.299
ACWE	0.653	0.709	0.783	2.107	3.623
DRLSE	0.409	0.661	0.565	6.0465	8.541
Shi-Karl's method	0.409	0.661	0.565	6.0125	8.624
WLSE	0.517	0.774	0.707	5.7425	7.876
Method by Liu *et al.*	0.393	0.493	0.545	2.792	5.126
Method by Nitsch *et al.*	0.608	0.652	**0.848**	2.170	3.566
Proposed Method	**0.717**	**0.789**	0.765	**1.484**	**2.421**

Table 5.2: Comparison of the different methods for segmentation of CP

Method	Precision	Accuracy	Dice	ASCD (mm)	RMSCD(mm)
GAC	0.437	0.741	0.571	6.538	9.377
ACWE	0.585	**0.906**	0.717	3.291	4.788
DRLSE	0.362	0.640	0.511	8.545	15.958
Shi-Karl's method	0.493	0.885	0.644	6.781	13.407
WLSE	0.579	0.976	0.733	6.280	12.721
Method by Liu *et al.*	0.510	0.847	0.661	5.272	7.466
Method by Nitsch *et al.*	0.571	0.746	0.718	3.84	7.162
Proposed Method	**0.620**	0.813	**0.753**	**2.291**	**3.288**

Table 5.3: Comparison of the different methods for the segmentation of the brain tumor

Method	Precision	Accuracy	Dice	ASCD (mm)	RMSCD (mm)
GAC	0.612	0.740	0.773	2.164	5.318
ACWE	**0.653**	0.682	0.769	2.031	4.378
DRLSE	0.436	0.578	0.593	5.587	14.57
Shi-Karl's method	0.622	0.675	0.735	3.312	12.122
WLSE	0.485	0.634	0.653	5.998	11.986
Method by Liu *et al.*	0.473	0.541	0.591	4.207	6.115
Method by Nitsch. *et al.*	0.620	0.813	0.753	13.166	23.152
Proposed Method	0.608	**0.852**	**0.848**	**2.17**	**3.565**

Table 5.4: Comparison of different methods for the segmentation of the ventricles

Method	Precision	Accuracy	Dice	ASCD (mm)	RMSCD(mm)
GAC	0.672	0.740	0.773	2.164	5.318
ACWE	0.653	0.712	0.769	2.031	4.633
DRLSE	0.436	0.578	0.593	5.587	14.579
Shi-Karl's method	0.622	0.675	0.735	3.312	12.122
WLSE	0.490	0.621	0.649	5.977	14.895
Method by Liu *et al.*	0.473	0.541	0.591	4.20	6.115
Method by Nitsch *et al.*	0.758	0.766	0.718	1.754	3.649
Proposed Method	**0.670**	**0.765**	**0.796**	**1.392**	**2.982**

Table 5.5: Comparison of different methods for the segmentation of the RC

observed in Table 5.5 where all five quality metrics are better than the other methods. The above results show that the proposed method outperforms other level-set based methods.

The comparative performances of the proposed method in segmenting LF, CP, tumor, ventricle, and RC are graphically presented in Fig. 5.8, Fig. 5.9, Fig. 5.10, Fig. 5.11, and Fig. 5.12 respectively. In each of these figures, (a) shows the targeted section along with the initial rectangular regions marked in red,(b) shows the manually segmented ground truth, (c) - (g) are the outputs of all the competing methods, and (h) shows the segmented boundary resulted by the proposed method.

We compared performance of the proposed method with one ML-based method by Nitsch *et al.* [141]. They proposed a random forest(RF) based BUS image segmentation method on the RESECT dataset that train with a few images only. It focused on the segmentation of hyper-echoic regions like Falx Cerebri and Tentorium cerebelli as a pre-work for the registration

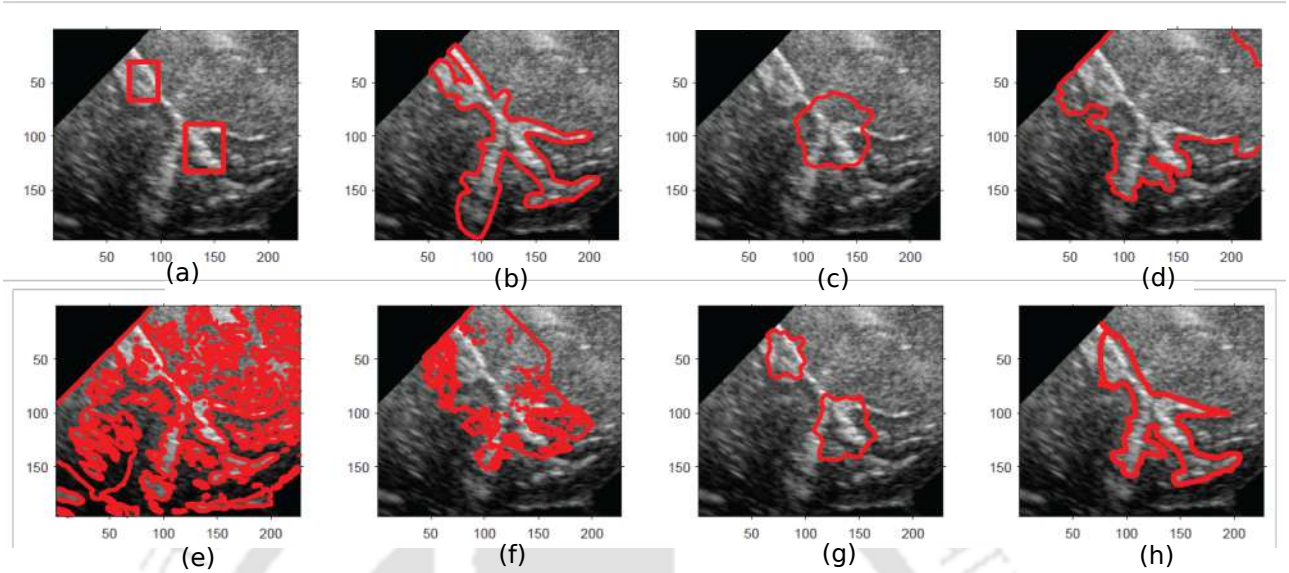


Fig. 5.8: (a) Original image with the initial boundary marked in red of LF by the user, (b) manually segmented ground truth LF. Resulting boundaries obtained by GAC, ACWE, the method by Shi *et al.*, the method by Liu *et al.*, WLSE, and the proposed method are shown in (c),(d),(e),(f),(g), and (h) respectively.

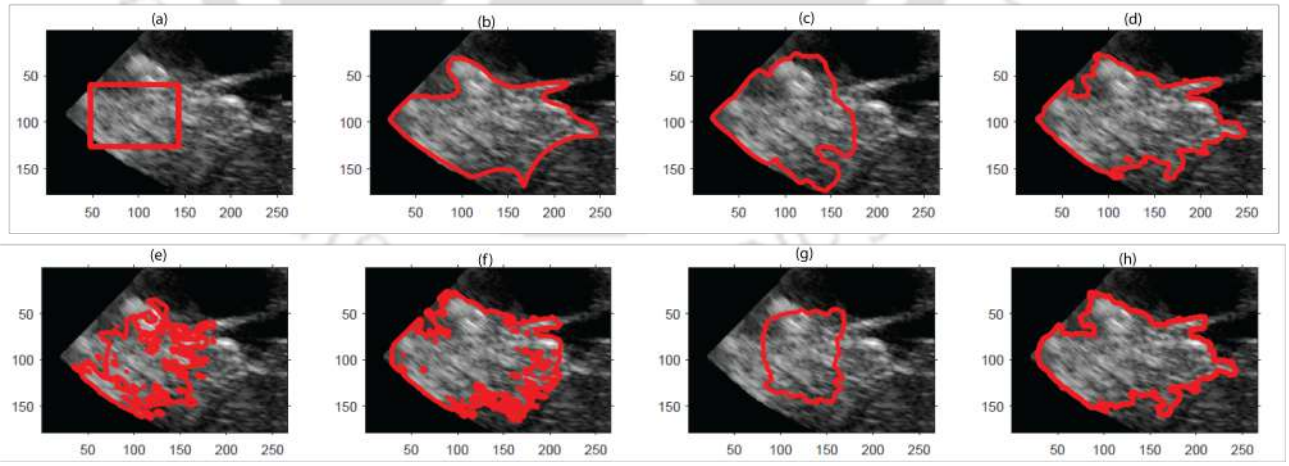


Fig. 5.9: (a) Original image with the initial boundary marked in red of CP by the user, (b) manually segmented ground truth CP. Resulting boundaries obtained by GAC, ACWE, the method by Shi *et al.*, the method by Liu *et al.*, WLSE, and the proposed method are shown in (c),(d),(e),(f),(g), and (h) respectively.

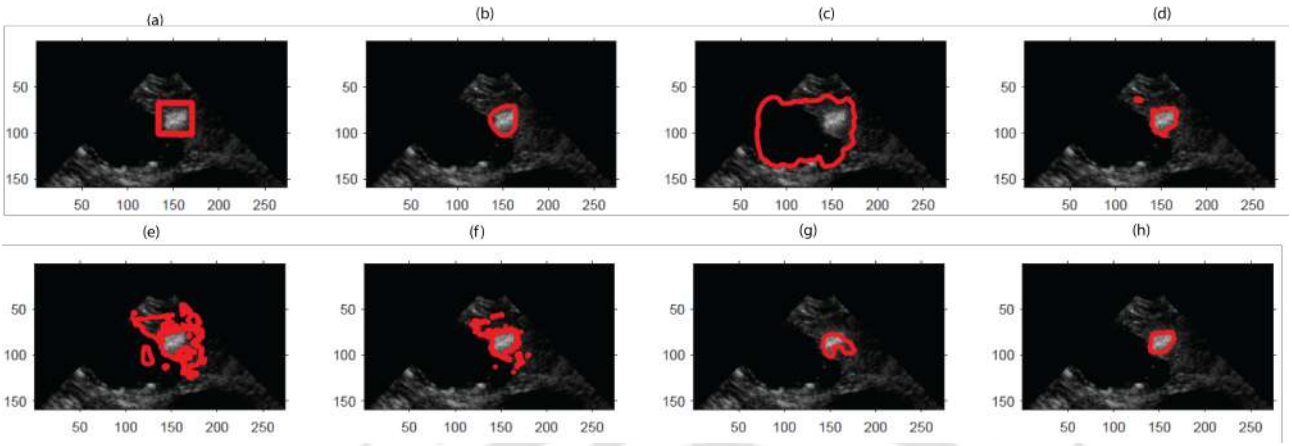


Fig. 5.10: (a) Original image with the initial boundary of the tumor marked by the user, (b) manually segmented ground truth tumor. Resulting boundaries obtained by GAC, ACWE, the method by Shi *et al.*, the method by Liu *et al.*, WLSE, and the proposed method are shown in (c),(d),(e),(f),(g), and (h) respectively.

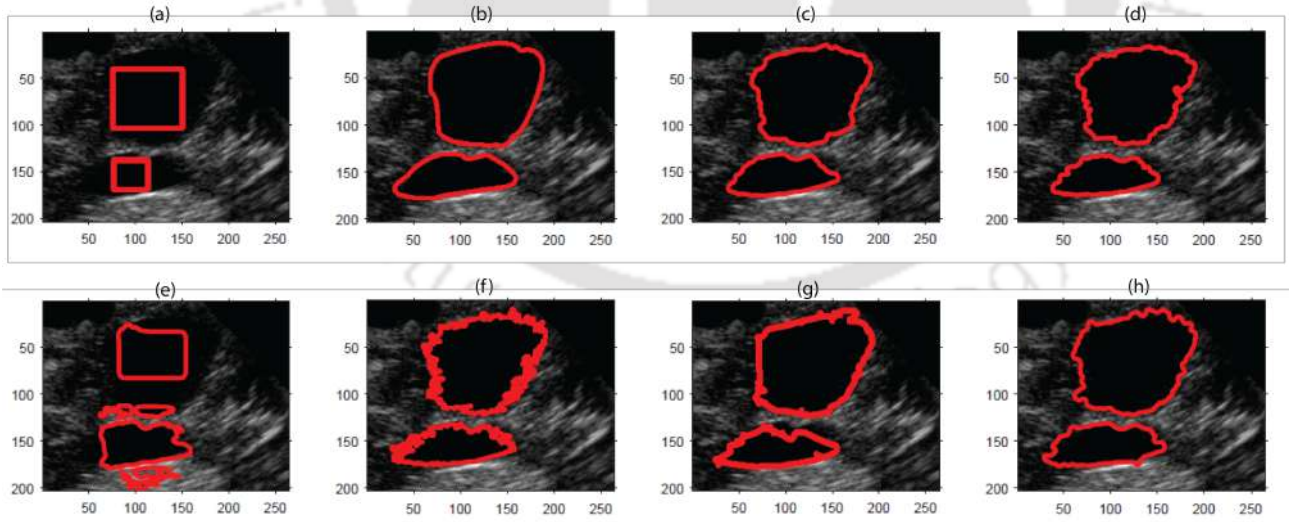


Fig. 5.11: (a) Original image with the initial boundary of ventricles marked in red by the user, (b) manually segmented ground truth ventricles. Resulting boundaries obtained by GAC, ACWE, the method by Shi *et al.*, the method by Liu *et al.*, WLSE, and the proposed method are shown in (c),(d),(e),(f),(g), and (h) respectively.

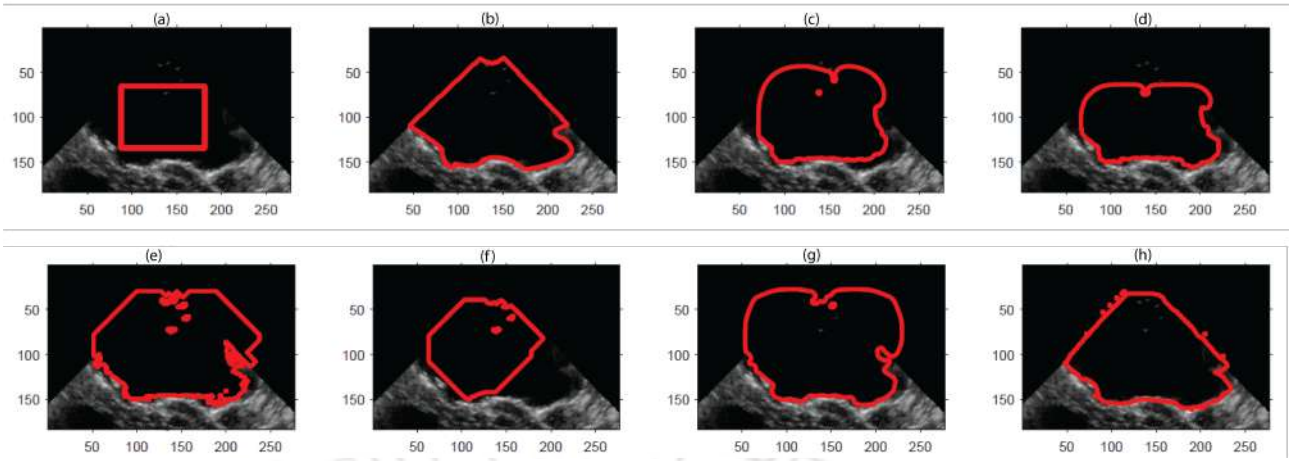


Fig. 5.12: (a) Original image with the initial boundary of the RC marked in red by the user, (b) manually segmented ground truth RC. Resulting boundaries obtained by GAC, ACWE, the method by Shi *et al.*, the method by Liu *et al.*, WLSE, and the proposed method are shown in (c),(d),(e),(f),(g), and (h) respectively.

of iBUS and the pre-operative MR images. US image slices are preprocessed with two stick filters as described in [142]. The stick filters strengthen and sustain the line-type structures and perform smoothing of the homogeneous regions by reducing speckle. After stick filtering, the entire image is divided into *simple linear iterative clustering* (SLIC) of superpixels as described in [2]. Each superpixel is classified by the trained RF classifier, which provides a probabilistic measure (λ_i) between 0 and 1 to be an element within the segmenting region. The region under segmentation grows by a three-stage multi-level super-pixel thresholding in three successive stages. For obtaining the segmented regions, superpixels are concatenated based on their λ_i scores. First, the concatenated superpixels having $\lambda \geq 0.59$ form an initial segmented region. The segmented region grows further attaching the the neighboring superpixels having $0.59 > \lambda_i \geq 0.45$. In a similar manner, the third step attaches the neighboring superpixels with $0.45 > \lambda_i \geq 0.3$. Fig. 5.13 compares the performance of the RF based method and the proposed method. For initiating the segmentation, initial rectangular regions are marked by red dotted boundaries in sub-figures (a1-a3) of Fig. 5.13. Fig. 5.13.(b1-b3) show a maroon coloured center region resulted from the first threshold comparison. Similarly, after the second threshold comparison, the orange colored superpixels get added to the segmented region, and thus the area of the segmenting region grows. Similarly, the third threshold comparison attaches the green superpixels to the previously grown region. The combined region after the third step shows the the final result. In Fig.5.13 (c1-c3), the green coloured boundaries represent the ground

truth. The red coloured boundaries are the results produced by the proposed method. The blue colored boundaries represent the output of the RF-based method. For the implementation of this method, we considered five classes (LF, CP, tumor, ventricle, and RC) and a total of 192 image sections were annotated for training and testing. Among them, 166 sets were used for training, and 26 sets were used for testing and evaluating the trained model. The comparative results based-on quantitative data for five different regions are shown in Table 5.1, Table 5.2, Table 5.3, Table 5.4, and Table 5.5. Fig. 5.13 shows that the performance of the proposed method is comparatively better, because it produces boundaries close to the actual boundaries.

Apart from the better segmentation performance, the proposed method is computationally faster compared to the other methods. The methods such as GAC [24], ACWE [28], DRLSE [98], and WLSE [90] minimize an energy function which is computationally expensive. The ML-based method of requires converting the image into a combination of superpixels, which has also significant computational burden. In addition to that, various rule-based concatenations of the superpixels require additional complexities. For comparing convergence speed of different algorithms, all experiments were conducted in the same computer system with a ninth-generation Intel Core-i5 processor (2.40 GHz), and 16 GB DDR4 RAM. We recorded the times it takes for the various steps of the proposed methods and the time taken by other algorithms with the best possible parameter setups. The comparative results shown in Fig.5.14 show that the proposed method takes 0.48 sec which is faster than the other methods. The bar diagrams in Fig.5.14 show the average time of segmentation of all 192 image sections used for the parametric study. The time taken by the method of Nitsch *et al.* is only due to the classification and rule-based concatenation of the superpixels. From this comparative study, it is evident that the proposed method is faster compared to other methods and has the potential of being used as an assisting tool to the neurosurgeon during brain surgery.

5.4 Registration of the pBUS and the iBUS images based on segmented regions

We have already discussed two methods of registration in Chapters 3 and 4. Both these methods do not need user intervention. Choice of various parameters such as the thresholds and various optimization parameters play crucial role for obtaining better results. In reality,

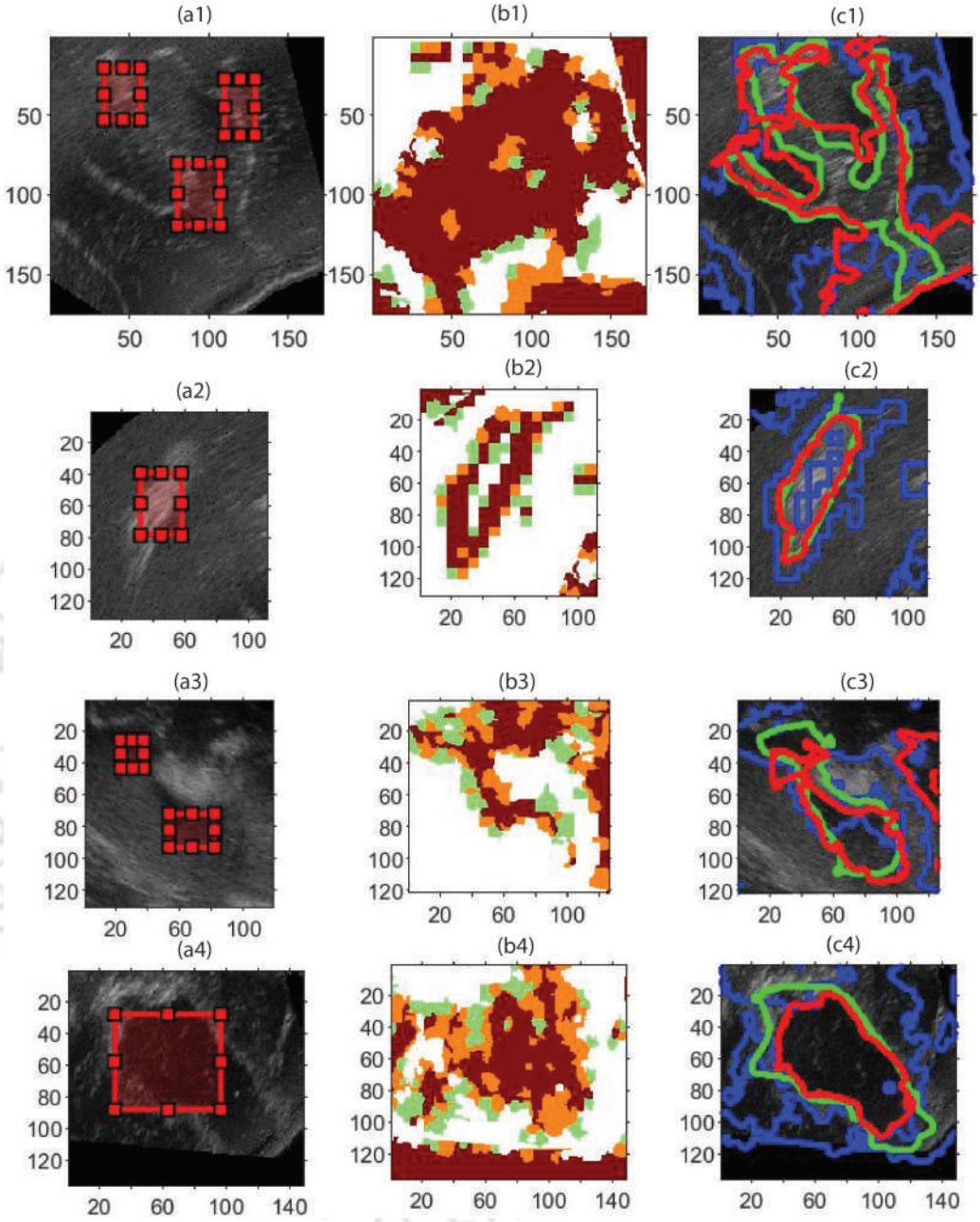


Fig. 5.13: (a1-a4) represent the original brain US image section of LF, CP ventricles, and RC respectively. The dotted red colored boundaries in (a1-a4) represent the initial region identification by the user. The maroon-colored regions in (b1-b4) are the group of superpixels having ($\lambda_i \geq 0.59$). Similarly, the orange-colored superpixels are neighboring to the maroon region and have $0.59 > \lambda_i \geq 0.45$. The green-colored superpixels are neighboring to the combined maroon and orange region resulting after the second thresholding and have $0.45 \geq \lambda_i > 0.3$. The green-colored boundaries in figures (c1-c4) are the corresponding ground truths of figures (a1-a4). The red-coloured boundaries are the results of the proposed methods and the blue-coloured boundaries are results by the method of Nitsch *et al.*

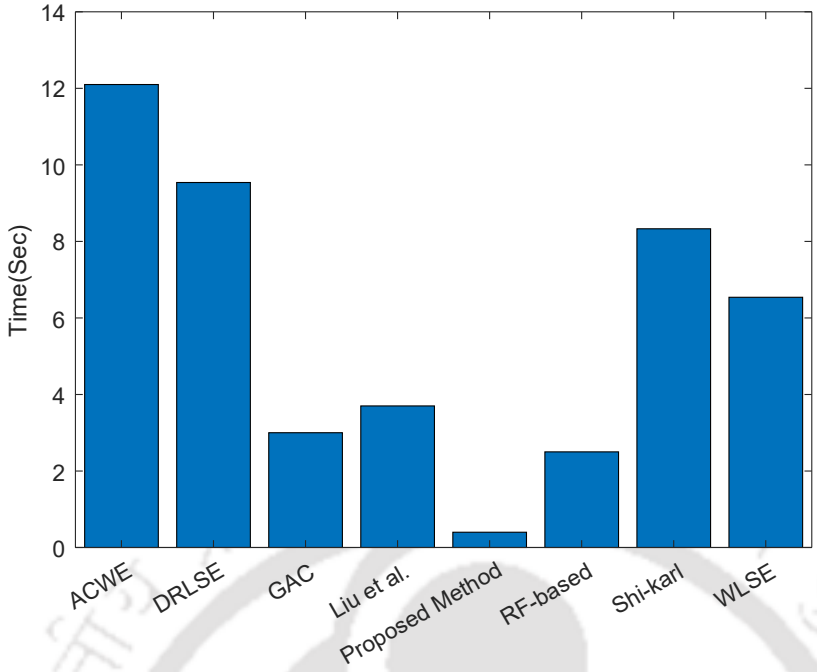


Fig. 5.14: Mean execution time of various segmentation methods

both these methods also need certain expertise for selecting these parameters. Both of these methods described in Chapter 3 and 4 require significant common area between the registering image pairs. Unfortunately, the iBUS images include various artifacts due to the presence of saline water and blood clotting agents in the resection cavity [123, 124]. The method in Chapter 3 performs patch-based thresholding of the absolute intensity gradient of each local patches. It can not consider bigger patch-size because it produces blocky edges of the segmented extracted region. For a better result, edge profiles in both the images should have considerable similarity. Unfortunately, the complete edge profiles in many iBUS images get lost due to the disturbed acoustic impedance. It reduces the pixel-to-pixel correlation between the pBUS and iBUS images.

Similarly, in Chapter 4 the HE regions of the registering images are extracted. It also adopts patch-based thresholding of the average patch intensities. The edge correction method and cubic B-spline-based contour smoothing in Chapter 4 restore the degraded edges, and thus captures better similarity between the registering image pairs. Out of the 20 difficult image pairs, 7 pairs could be registered by the methods in Chapter 4. However, there is scope of improvement in the performance by exploiting the knowledge of a user in a semi-automatic manner.

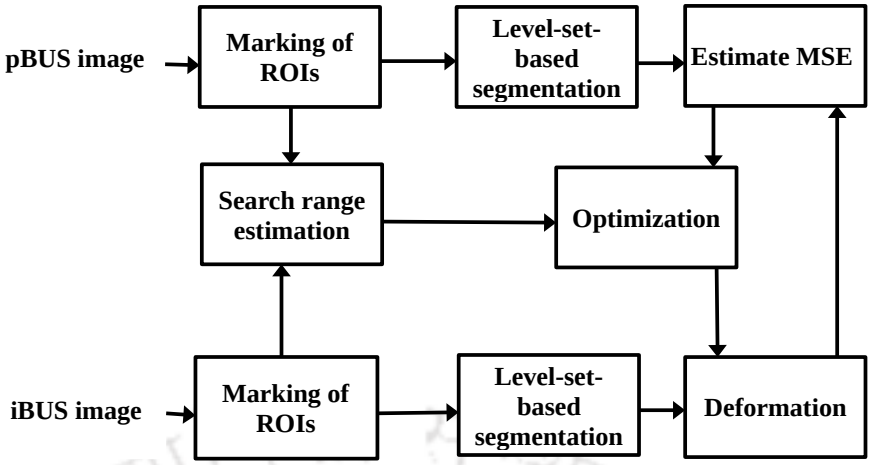


Fig. 5.15: Block diagram of the registration of pBUS and iBUS images using the proposed level-set-based segmentation

This chapter presented a semi-automatic method for the segmentation of the desired regions. The important steps involved in this registration are shown in Fig. 5.15. One or multiple corresponding ROIs are marked by the user on the pBUS -iBUS image pairs. Depending on the region's intensity profile, the algorithm identifies echogenicity type. The segmentation is carried out as per the method described in Section 5.2.

5.4.1 Marking of the ROIs

It was discussed earlier that the image pairs that are highly affected by artifacts possess low similarity between them. Therefore the similar regions can be identified better by human visual inspection. For the selection of these regions, registering image pairs are kept side-by side, and the same ROIs are marked by the user. Fig. 5.16(a) and (b) show a pBUS and iBUS image pair that have low similarity between them. Five corresponding ROIs are marked in both images and segmentation of the regions starts from these regions as described in the Section 5.2.

5.4.2 Search range estimation

The registration methods adopted in Chapter 3 and 4 applied various heuristic optimization methods. Gradient-based methods do not perform well because of the multiple local optima. This situation often arises when the initial search point is far from the global solution. Heuristic optimization methods perform well because they search for a better solution over a wide search

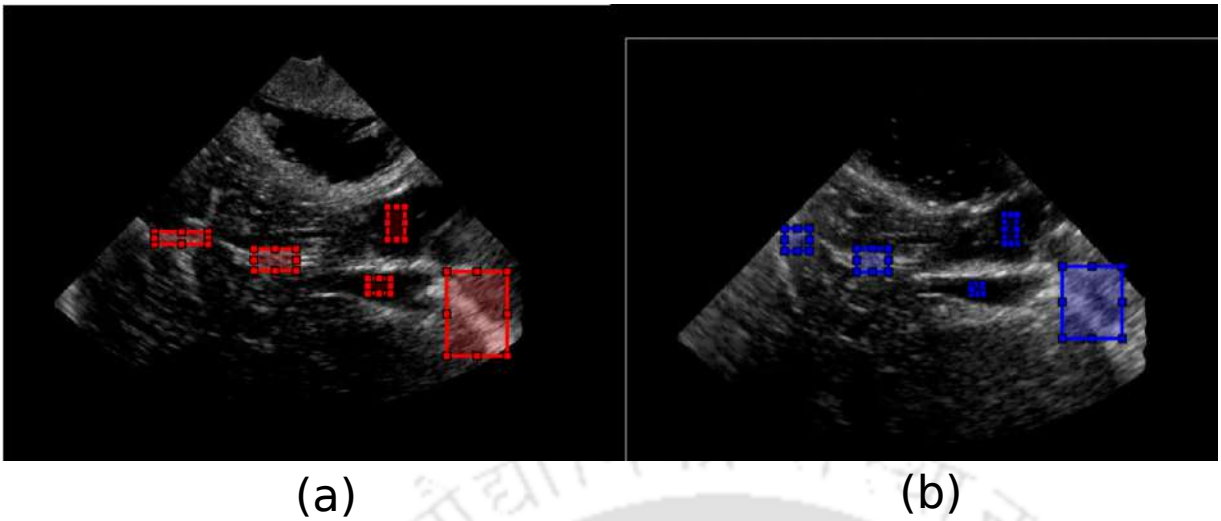


Fig. 5.16: Marked rectangles on (a) pBUS image and (b) iBUS image

range. But the computation of heuristic optimization methods are high. The registration methods used in Chapter 3 and 4 do not use any knowledge idea of the location of the probable global solution. Therefore, the bounds of the search-variables had to be kept much higher. If the searching is carried out in a big search-space, it is very likely that the solution converges to a local optimum by a gradient-based method. In the proposed method, the user marks the corresponding ROIs on the pBUS and the iBUS images. From the coordinates of the centers of the corresponding rectangles, we can obtain a better initial solution.

Let $\mathbf{X}^k = \begin{bmatrix} x_1^k \\ y_1^k \end{bmatrix}$ and $\mathbf{Y}^k = \begin{bmatrix} x_2^k \\ y_2^k \end{bmatrix}, k \in \{1, 2, ..., s\}$ be the corresponding centers of the marked rectangles on $\mathbf{V}_S$ and $\mathbf{V}_T$ respectively. Fig. 5.17 shows an example of the rectangular regions chosen from the registering image pairs. Fig. 5.17.(a) overlays the pBUS and iBUS images and the ROIs marked on them. The initial distances between the rectangles show misalignments between them. For correcting this, the transformation parameters are given by

$$\hat{a}, \hat{b} = \underset{a,b}{argmin} \sum_{k=1}^{s} \left| \mathbf{X}^k - T\mathbf{Y}^k \right|^2 \tag{5.11}$$

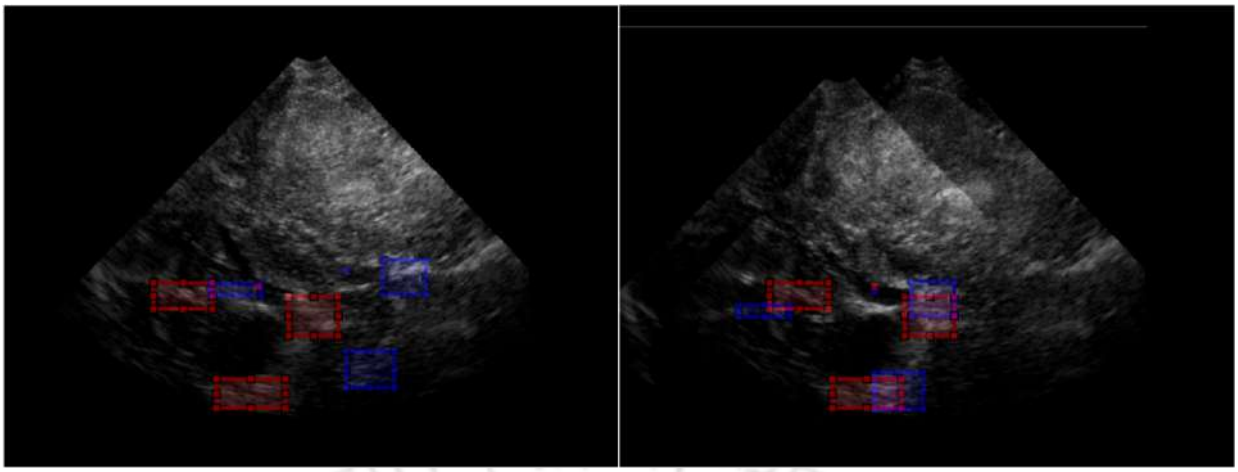


Fig. 5.17: Rectangles as shown in Fig. 5.16.(a), and (b) in a common frame in (a) before registration and (b) after registration

where $T\mathbf{Y}^k = \begin{bmatrix} x_2^k - a \\ y_2^k - b \end{bmatrix}$.

It can be shown that the optimum translation vector $\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}$ in (5.11) is

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \frac{1}{s} \sum_{i=1}^{s} \left| \mathbf{X}^k - \mathbf{Y}^k \right| \tag{5.12}$$

Estimation of $\hat{a}$ and $\hat{b}$ solely depends on the positions of the rectangles, and does not require the image intensity information. Fig. 5.17.(a) shows the rectangular regions of pBUS and iBUS images overlaid in a common frame. The average distance between the rectangles is bigger in this image. After aligning the ROIs by translating $\mathbf{V}_S$ by T, the average distances are reduced, and the corresponding rectangles come closer that can be seen in Fig. 5.17.(b). This improvement is shown by the position of tag points before and after the registration in Fig. 5.18.(a) and (b). The average distance between that tag-points appears higher before registration, which reduces significantly after this step which is shown in Fig. 5.18.(b).

The above step grossly aligns the registering image pairs. However, the above alignment does not consider rotation as a transformation parameter. We discussed earlier that the MSE of the error hypersurface is not convex, and the gradient-based methods often converge to a local

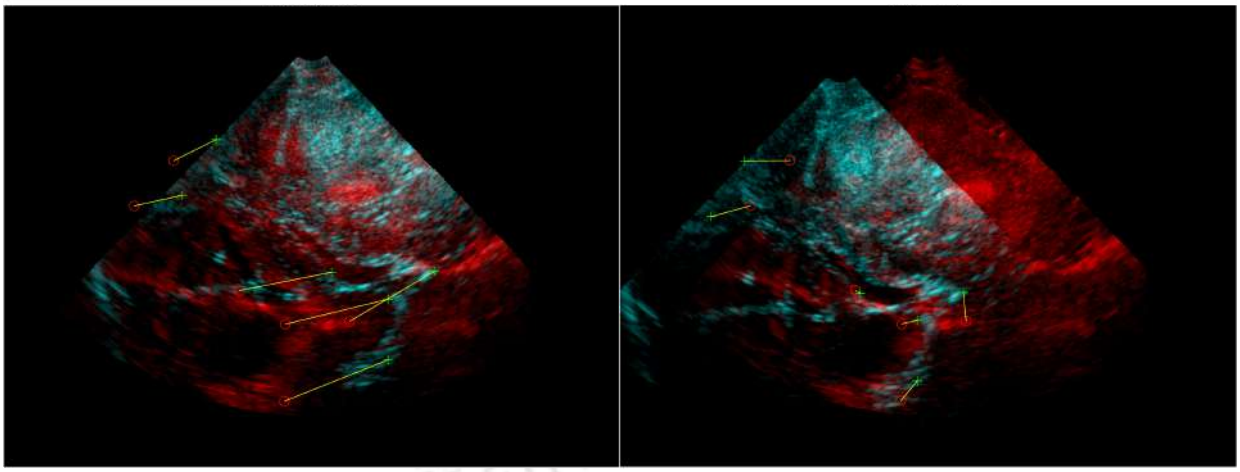


Fig. 5.18: Both the pBUS and iBUS images are overlaid in a common frame. The corresponding tag-points between pBUS and iBUS images are connected in (a) and (b). (a) shows the positions of the corresponding tag-points before the adjustment of the rectangle's positions, and (b) shows the corresponding tag-points after the adjustment of the rectangle's positions.

minimum. Heuristic optimization performs better in such a scenario, but they involve high computational complexity. The above step registers images grossly based on the positions of the corresponding ROIs marked by the user. In the next step, the MSE between the pBUS and the deformed iBUS image pairs is minimized for obtaining the required shift and rotation. The relative misalignment between the image pairs becomes less after aligning the ROIs; therefore, the search range can be reduced considerably. It makes the optimization more accurate, and the overall registration performance improves significantly. The registration model considers three parameters: i) The horizontal shift (x_0), ii) the vertical shift (y_0), and rotation θ_0. The search ranges for x_0, y_0, and θ_0 are kept $\hat{a} \pm 20$, pixels, $\hat{b} \pm 20$ pixels, and $\pm 20^\circ$ respectively.

5.4.3 Segmentation of the desired regions

The desired regions marked by the ROIs are segmented by the method described in Section 5.2. Fig. 5.19.(c) and (d) show the segmented regions of the images in Fig. 5.19.(a) and (b) respectively. During segmentation, the segmented regions are preserved with unaltered intensity, and the remaining part of the image is added to the background with zero intensity. But this methodology can be applied only for HE and hypo-echoic regions. The anechoic regions have low intensity values and therefore, it is hard to separate them from the background. For

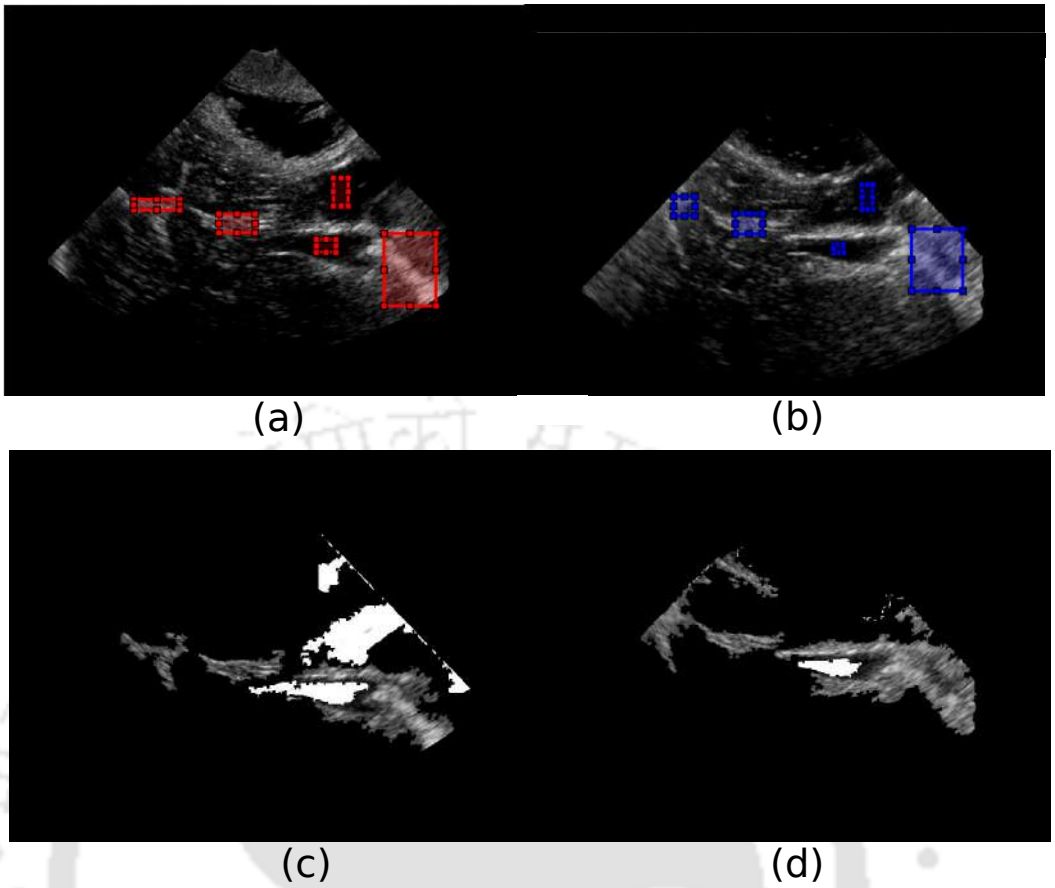


Fig. 5.19: (a) A pBUS image with marked ROIs (b) Corresponding iBUS image and marked ROIs. (c) The segmented regions of the image (a). (d) The segmented region of the image in (b)

mitigating this issue, the intensity values of the anechoic regions are inverted to its negative, which helps to differentiate it from the background region. Fig. 5.19.(a) and (b) show the segmented ventricles, which is the negative of its original intensity values. They can be easily identified and compared.

5.4.4 Optimization

Population-based heuristic optimizations have more complexity, and they search over a wide search range; hence, they often reach the global solution. In an automatic registration method, we do not have any previous assumption about the solution point and therefore, we need to consider a wide search range. Gradient-based methods fail because of multiple local optima within the wide search range. The alignment of ROIs does not consider image information, and it provides an approximate translation between $\mathbf{V}_S$ and $\mathbf{V}_T$. For obtaining

better registration, we optimize further the objective function in the neighborhood of $\begin{bmatrix}\hat{a}\\ \hat{b}\end{bmatrix}$. In the proposed registration, we applied three heuristic optimization methods: PSO, SA, and BBO, and two gradient-based optimization methods: Levenberg-Marquardt Algorithm (LMA) and LBFGS-B [17] for performance comparison.

The MSE between the segmented image pairs was considered as the objective function. Let $\mathbf{V}_S$ and $\mathbf{V}_T$ be the registering image pair. If $\mathbf{V}_S$ undergoes the transformation $T_{\theta_0}(x_0, y_0)$ the cost function can be defined as,

$$J = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}|\mathbf{V}_T(x,y) - \mathbf{V}_S(T_{\theta_0}(x-x_0, y-y_0)|^2 \tag{5.13}$$

where $\begin{bmatrix}x_0\\ y_0\end{bmatrix}$ is located about $\begin{bmatrix}\hat{a}\\ \hat{b}\end{bmatrix}$. The deformation matrix $T_{\theta_0}(x-x_0, y=y_0)$ considers only three parameters x_0, y_0, and θ_0, and

$$\mathbf{T}_{\theta_0}(x-x_0, y-y_0) = \begin{bmatrix} cos\theta_0 & -sin\theta_0 & 0\\ sin\theta_0 & cos\theta_0 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & -x_0\\ 0 & 1 & -y_0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x\\ y\\ 1\end{bmatrix}$$

In many cases, the cost function in (5.13) may diverge from the actual solution point, during iterative optimization. For stopping the large displacement of the solution points, a regularizer is added to the objective function to get the modified cost function as

$$J = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}(\mathbf{V}_T(x,y) - \mathbf{V}_S(\mathbf{T}_{\theta_0}(x-x_0, y-y_0))^2 + k((x_0-\hat{a})^2 + (y_0-\hat{b})^2 + \theta_0^2) \tag{5.14}$$

where k is a constant chosen between $1 \leq k \leq 5$. The value of k limits the spread of the solution and the search operation is carried out near $(\hat{a}, \hat{b})$ and chosen empirically.

Thus the optimum solution can be obtained as

$$\begin{bmatrix}x'\\ y'\\ \theta'\end{bmatrix} = \underset{(x_0,y_0,\theta_0)}{argmin} J \tag{5.15}$$

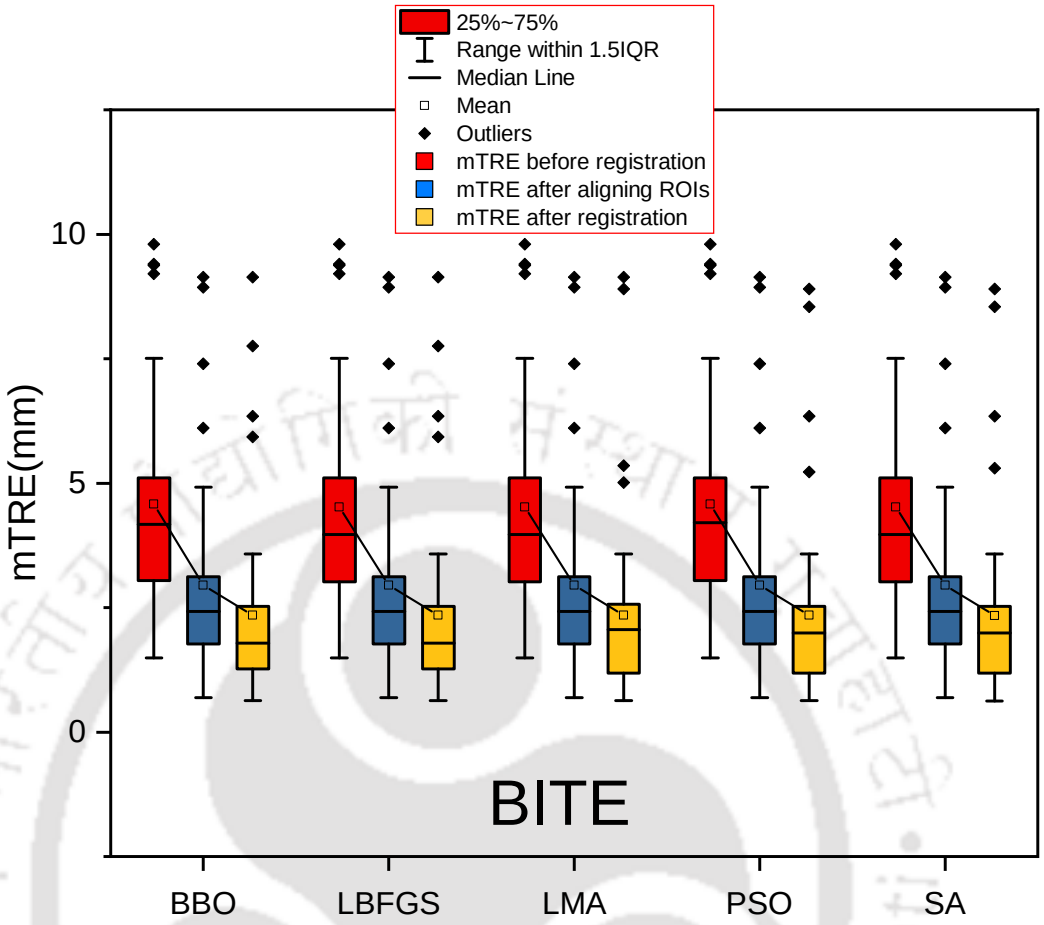


Fig. 5.20: Comparison of mTREs for the difficult image pairs of the BITE dataset. Five sets of results are provided for five optimization methods. The results of each optimization consists of 3 boxes. The red box represents the initial mTRE. The blue colored middle boxes represent the results after aligning the ROIs. The yellow boxes show the final result after the registration.

5.5 Results and discussion

The primary focus of the proposed method is to perform registration of the difficult image pairs, which could not be registered by the methods discussed in Chapter 3 and 4. Like in the previous chapters, we applied various optimization methods for each of the registering image pairs. First, we compare the average mTRE of all the difficult images before and after the aligning of the ROIs. Later, we apply segmentation-based registration using various optimization methods. We show the comparative results graphically for the difficult images.

We have added 4 pairs of difficult image pairs from the BITE and 3 pairs of difficult image

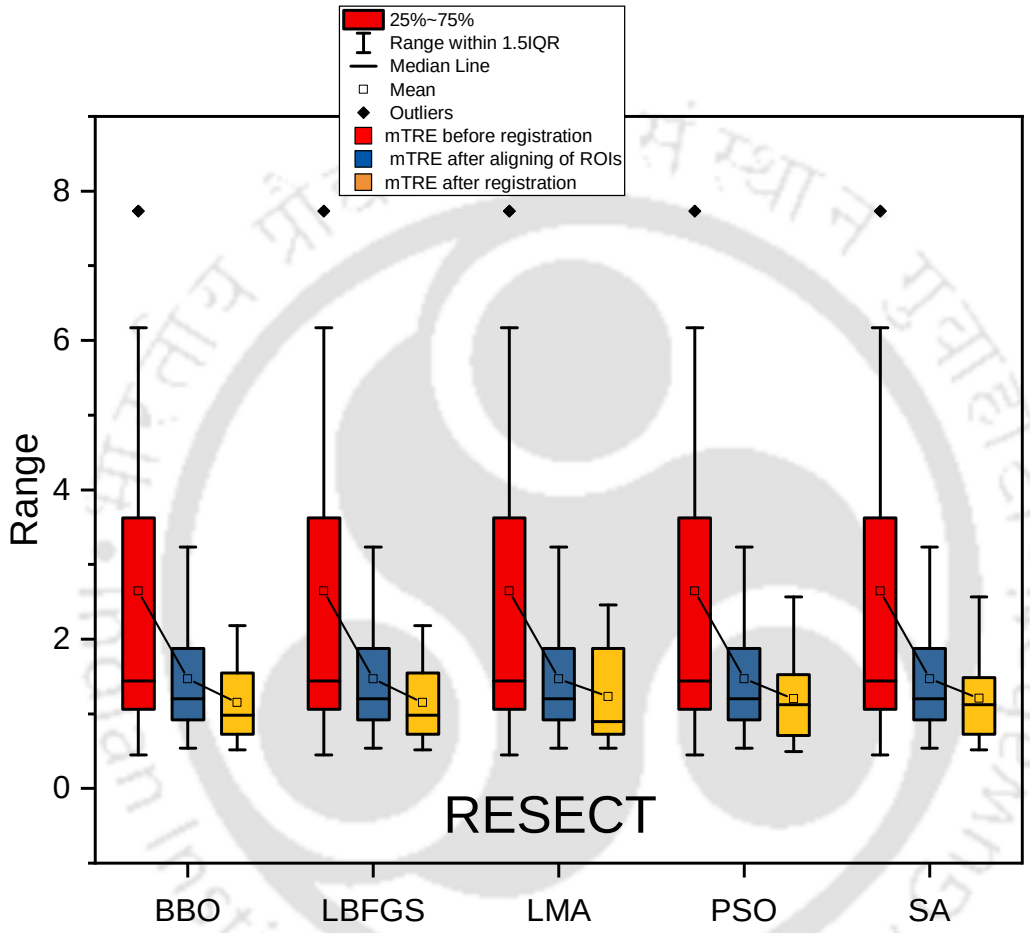


Fig. 5.21: Comparison of mTREs for the difficult image pairs of the RESECT dataset. Five sets of results are provided for five optimization methods. The results of each optimization consists of 3 boxes. The red box represents the initial mTRE. The blue colored middle boxes represent the results after aligning the ROIs. The yellow boxes show the final result after the registration.

Optimization algorithm	Average Initial mTRE(mm)	Average mTRE after aligning ROIs (mm)	Average final mTRE(mm)
BBO	4.577	2.949	2.352
PSO	4.577	2.949	2.367
SA	4.577	2.949	2.344
LMA	4.577	2.949	2.358
LBFGS-B	4.577	2.949	2.352

Table 5.6: Table shows the average mTRE for the registration of images of the BITE dataset. Three columns shows the initial mTRE, the average mTRE after aligning the ROIs, and the mTRE after the final registration.

pairs from the RESECT datasets during implementation of this method. Fig. 5.20 shows the average mTRE for 18 difficult image pairs of the BITE dataset. Similarly Fig. 5.21 show the average mTRE for 12 difficult image pairs from RESECT datasets. Each figure shows the results for five different optimization methods. For each optimization methods, three boxplots are shown. The first box plot in red color shows the initial average mTRE between the tag-points. This distance reduces the mTRE after aligning the ROIs, that is shown in the second box plots with blue color. The final mTREs are shown in the third box plot shown in yellow color. Because the search range during the registration is less, all optimization methods perform better, and reduces the mTRE. The parametric results for the segmentation of the images of the BITE and the RESECT datasets are summarized in Table 5.6 and Table 5.7 respectively.

On an average the initial mTRE for the BITE dataset images was 4.577 mm, which was improved after aligning the ROIs into 2.949 mm, which was further reduced to 2.367 mm by the LBFGS optimization method. Though all optimization methods obtain almost the same results, the LBFGS is the fastest among them; hence, we recommend to use this algorithm over others. Compared to the method in Chapter 4, this method is faster because both LBFGS algorithm is gradient-based.

The introduction of aligning of ROIs and the segmentation of desired regions improve the registration performance significantly. Out of 18 difficult images pairs of the BITE dataset, 12 image pairs could be successfully registered with an average mTRE less than 2.2 mm. Similarly,

Optimization algorithm	Average Initial mTRE(mm)	Average mTRE after aligning ROIs(mm)	Average final mTRE(mm)
BBO	2.645	1.465	1.149
PSO	2.645	1.465	1.2
SA	2.645	1.465	1.206
LMA	2.645	1.465	1.232
LBFGS-B	2.645	1.465	1.149

Table 5.7: Table shows the average mTRE for the registration of images from the RESECT dataset. The three columns show the initial mTRE, the average mTRE before aligning the ROIs, and the mTRE after final registration.

out of 12 difficult image pairs of the RESECT dataset, 8 image pairs were successfully registered with average resulting error less than 1.2 mm. Fig. 5.22 and 5.23 show the results of 2 pairs of images, which could not be registered by the methods described in Chapter 3 and 4, but the proposed method registers them successfully with a small error. Fig. 5.22, Fig. 5.23, Fig. 5.24, Fig. 5.25, and Fig. 5.26, (a) and (b) show the image rectangular regions marked by the user. The segmented region initiated from those rectangles, are shown in (c) and (d). The initial positions of the corresponding tag-points on pBUS and iBUS images are shown in (e). The positions of the tag-points after registration are shown in (f). The images in Fig. 5.22 and Fig. 5.23 are taken from the BITE datasets and Fig. 5.24, Fig. 5.25, and Fig.5.26 show the results for images from the RESECT dataset.

The proposed method produces results better than the methods described in Chapter 3 and 4 due to the following reasons:

- This method is semi-automatic, the user provides an idea of the actual solution by marking the ROIs in both images. The alignment of ROIs brings the images closer to the solution.
- We add a regularizer term in the objective function as shown in (5.14), that restricts large drift of the solutions from the initial points.
- LBFGS shows performance comparable to the heuristic methods. Considering the fast convergence of the gradient-based methods, it is preferred optimization scheme.

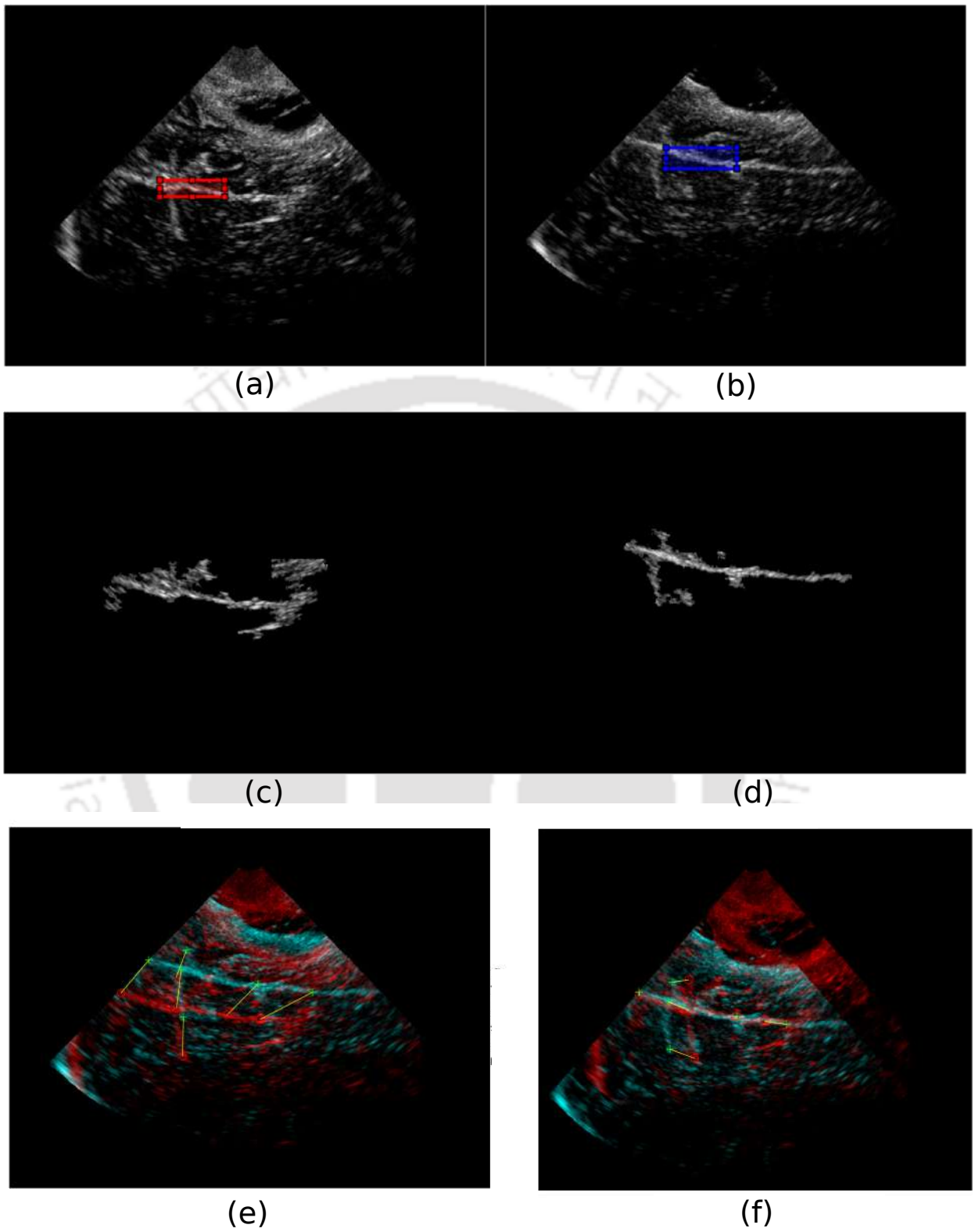


Fig. 5.22: (a)-(b) show the pBUS and iBUS images from the BITE dataset with the marked ROIs on the corresponding regions. (c)-(d) are segmented regions from the pBUS and iBUS images respectively. (e) shows the corresponding tag-points before registration on the overlaid pBUS and iBUS images. (f) shows corresponding tag-points after final registration on the overlaid pBUS and iBUS image.

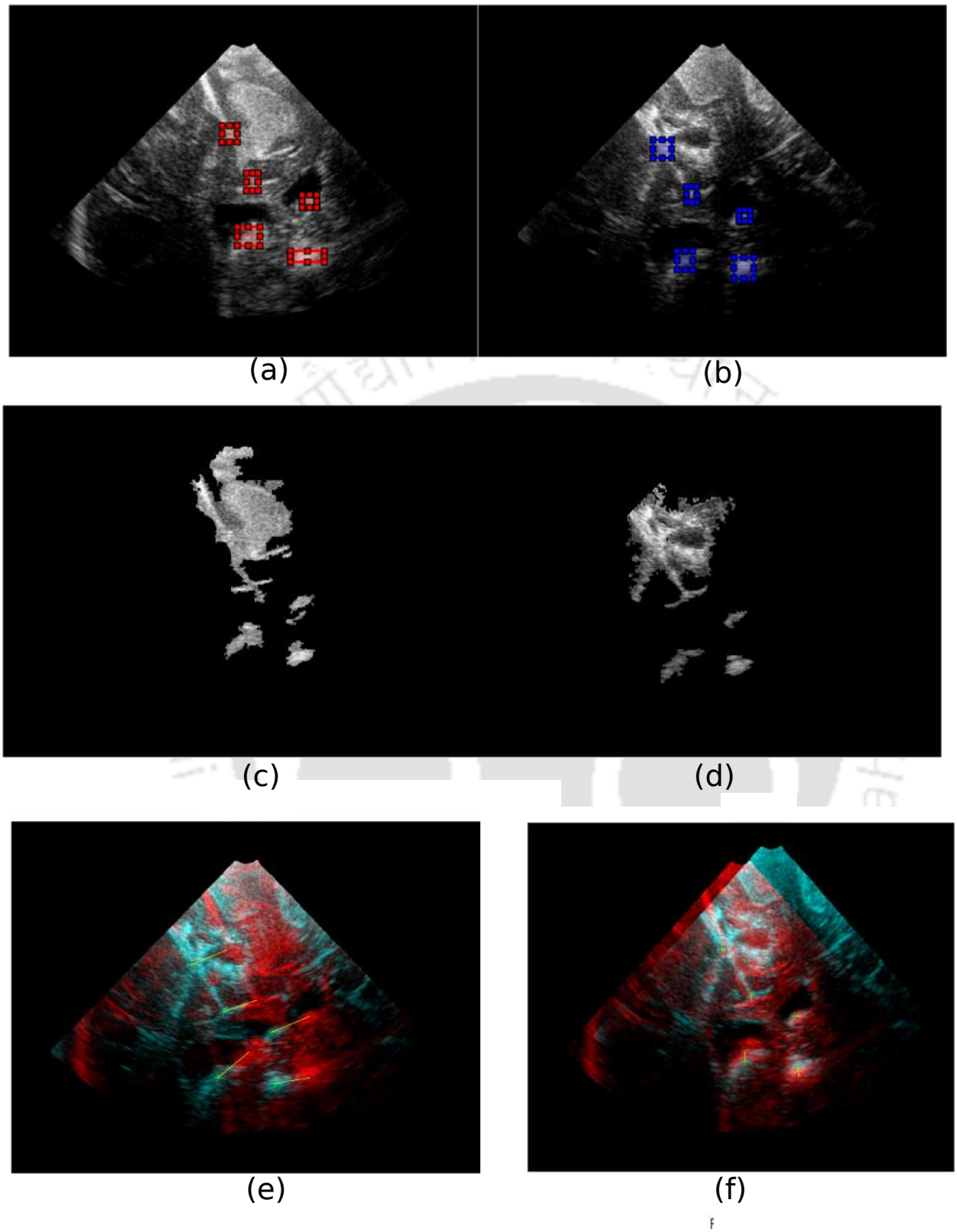


Fig. 5.23: (a)-(b) show the pBUS and iBUS images from the BITE dataset with the marked ROIs on the corresponding regions. (c)-(d) are segmented regions from the pBUS and iBUS images respectively. (e) shows the corresponding tag-points before registration on the overlaid pBUS and iBUS image. (f) shows corresponding tag-points after final registration on the overlaid pBUS and iBUS image.

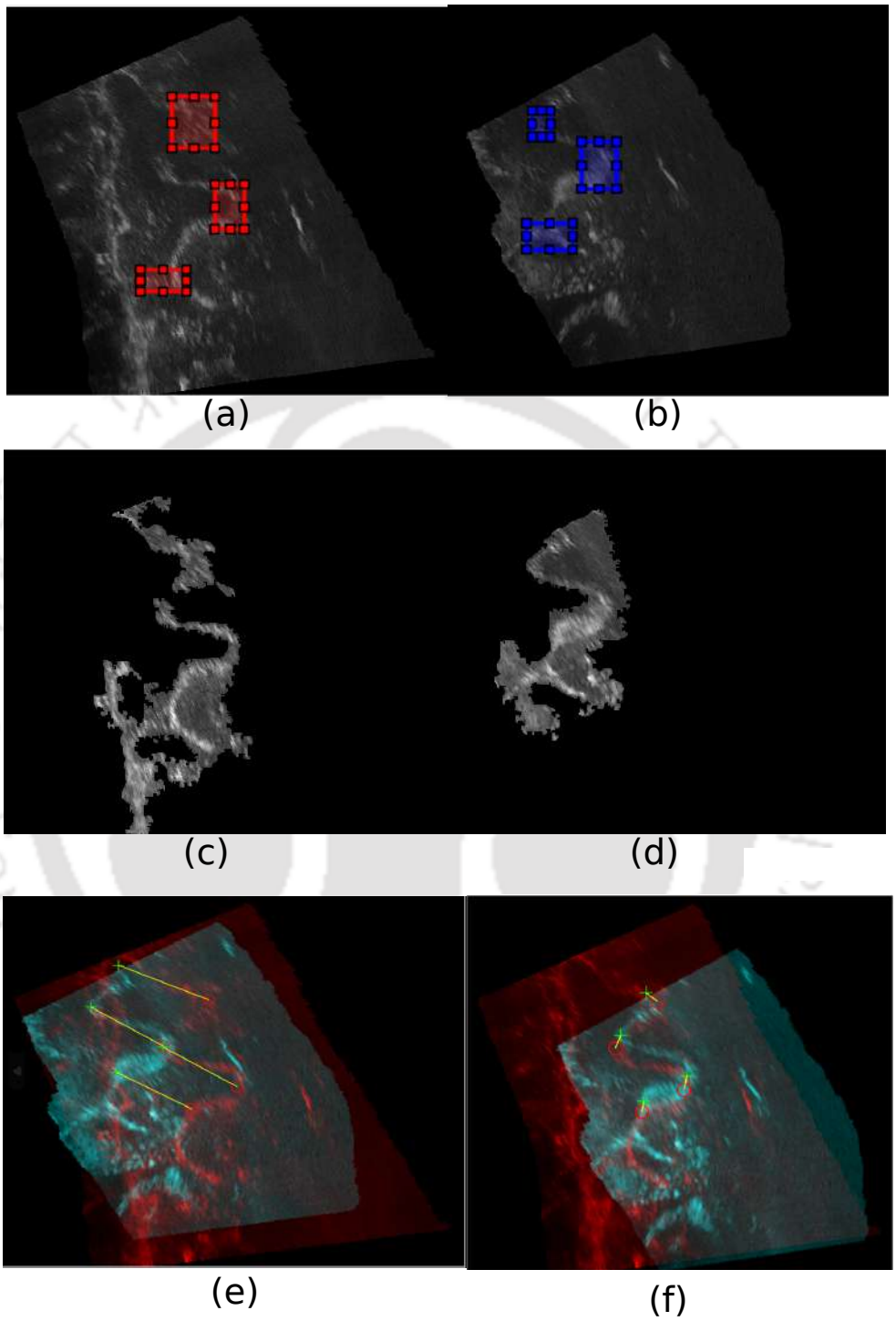


Fig. 5.24: (a)-(b) show the pBUS and iBUS images from the BITE dataset with the marked ROIs on the corresponding regions. (c)-(d) are segmented regions from the pBUS and iBUS images respectively. (e) shows the corresponding tag-points before registration on the overlaid pBUS and iBUS image. (f) shows corresponding tag-points after final registration on the overlaid pBUS and iBUS image.

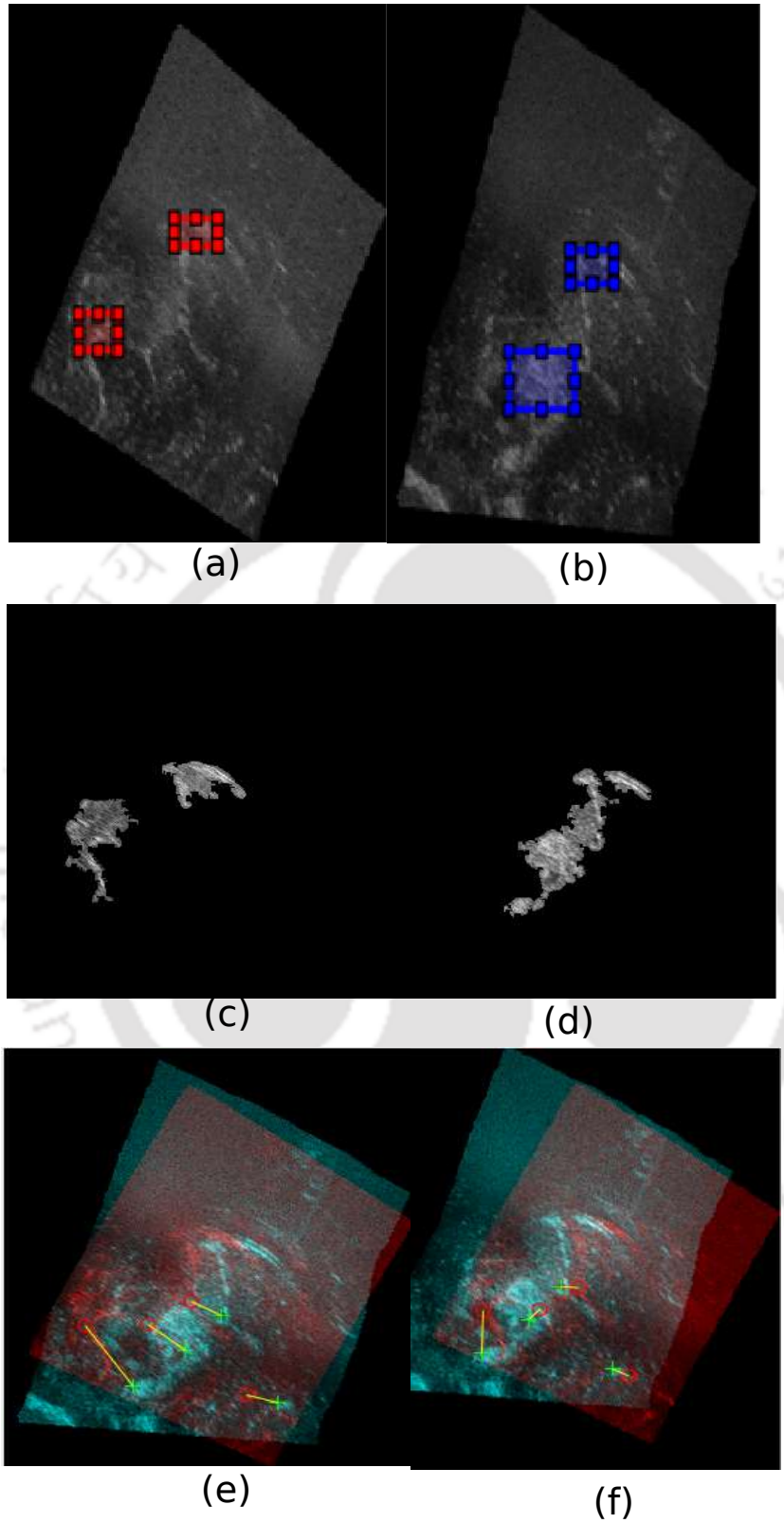


Fig. 5.25: (a)-(b) show the pBUS and iBUS images from the RESECT dataset with the marked ROIs on the corresponding regions. (c)-(d) are segmented regions from the pBUS and iBUS images respectively. (e) shows the corresponding tag-points before registration on the overlaid pBUS and iBUS image. (f) shows corresponding tag-points after final registration on the overlaid pBUS and iBUS image.

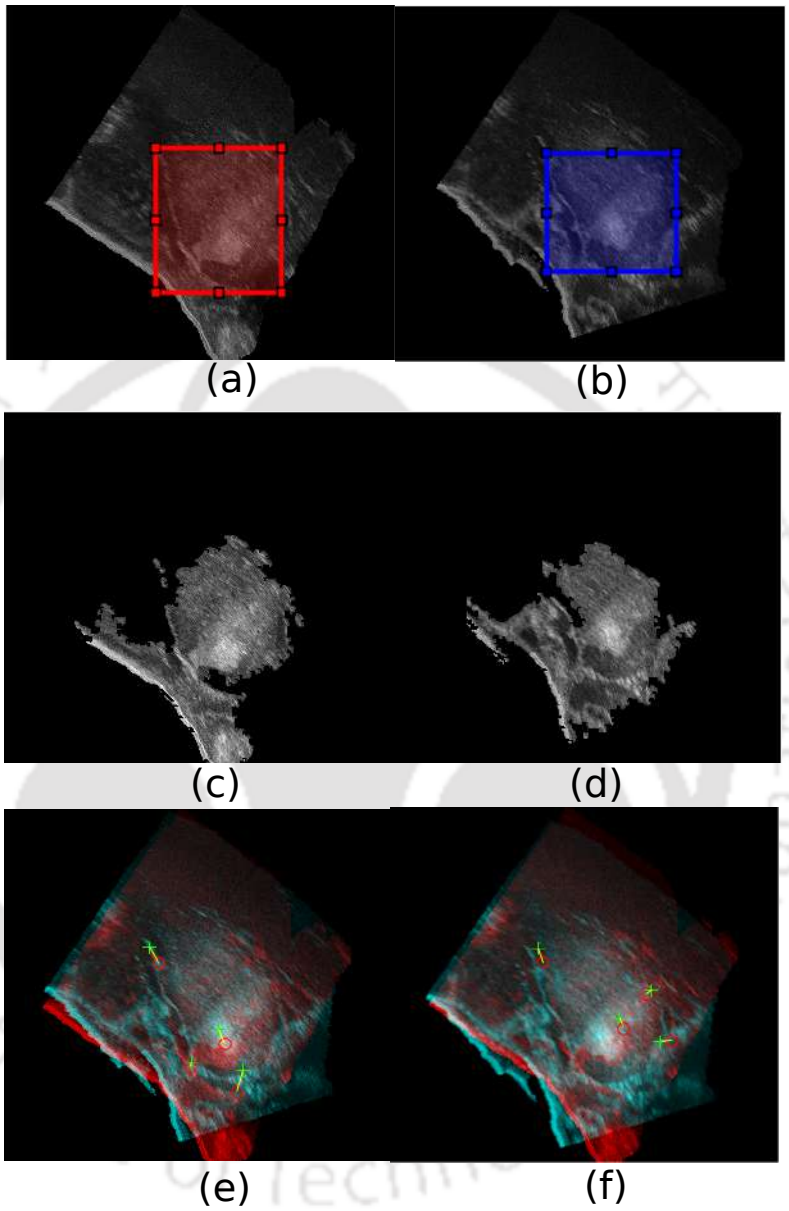


Fig. 5.26: (a)-(b) show the pBUS and iBUS images from the RESECT dataset with the marked ROIs on the corresponding regions. (c)-(d) are segmented regions from the pBUS and iBUS images respectively. (e) shows the corresponding tag-points before registration on the overlaid pBUS and iBUS image. (f) shows corresponding ta-points after final registration on the overlaid pBUS and iBUS image.

5.5.1 Overall registration performance of the proposed method

We compare the overall improvements of the performance of the fast level-set segmentation-based method with the methods presented in Chapter 3 and 4. The registration method discussed in Chapter 3 could not register the difficult image pairs. The success rate of the method in Chapter 4 was poor for both the BITE and the RESECT datasets. It was able to perform only 33 % difficult image pairs of the RESECT dataset and 29 % of different pairs in the BITE dataset. The patch-based method in this chapter registers 8 pairs out of 12 pairs of difficult images from the RESECT dataset with a small mTRE values. Similarly, for the BITE dataset, 13 pairs of images could be registered out of 18 pairs of difficult images from the BITE dataset. Fig. 5.27.(a) and (b) show graphically the success rates of the methods on three categories of easy, medium, and difficult image pairs of the RESECT and the BITE datasets respectively. The registration is considered as a success when the post-registration mTRE is less than 2.5 mm. The method of Canalini *et al.*, Machado *et al.* are also included. These figures show that the proposed method has a success rate 67 % for the difficult images of the RESECT dataset and 71 % for the difficult images of the BITE datasets. All three proposed methods are able to register easy and medium BUS image pairs. The proposed fast level-set-based method in this chapter shows a significant improvement registering difficult image pairs over the other methods and the methods described in Chapter 3 and 4.

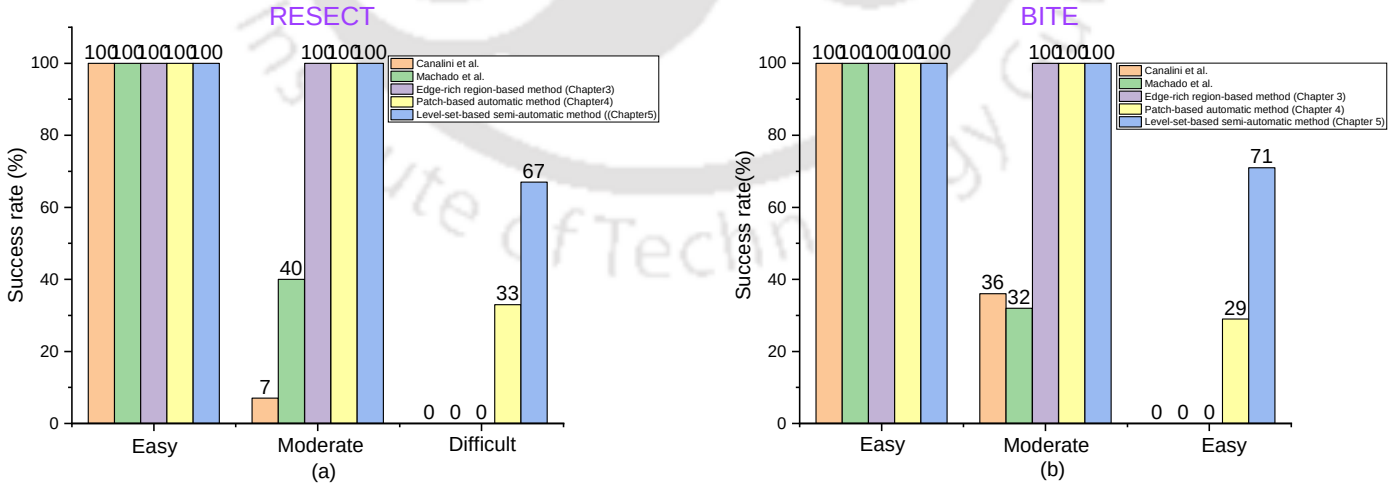


Fig. 5.27: (a) Comparative success rates of five different methods for registering easy, medium and difficult image pairs.

5.6 Conclusions

Aa semi-automatic approach is proposed for improvement of registration of the pBUS and the iBUS images. The proposed method allows user to identify one or many corresponding regions from the pBUS and iBUS images and mark them. Corresponding regions are segmented from the pBUS and iBUS image pairs. The proposed method uses patch-based level-set curve evolution to segment different regions in pBUS and iBUS images based on their echogenicity. Five different image regions were segmented, and comparison of the results of the proposed method to the other state-of-the-art methods were presented region wise. An upper and a lower threshold are determined from each of the rectangular region chosen by the user. A patch-based ULSCF step is applied followed by BLSCF. This ULSCF step creates approximate blocky boundaries by segmenting the target region with bigger block sizes. The BLSCF requires a finite number of iterations and returns the actual shape of the boundaries. This step focuses only on the boundary region and iteratively refines the shape of segmented region in the boundary area by reducing the patch size in a dyadic scale. This procedure begins with a patch size of 8×8 and continues until the patch size is reduced to 1×1. A curvature controlled cubic B-spline interpolation method adds to the smoothness of the boundaries. The performances of other level-set based approaches and a super-pixel based RF classifier were compared in this chapter. The proposed method outperforms previous level-set based algorithms and the ML-based RF classifier. Because of the patch-based processing, it is faster than other approaches, and the resulting boundaries are free of slope and curvature discontinuities.

The registration process is divided into two steps. In the first step, the position of the marked ROIs on pBUS and iBUS images are used to bring those regions closer, and thus an initial registration is obtained. In this step no image data is used for registration. In the second step, registration is performed again based on the segmented regions of pBUS and iBUS images with a reduced search range. We developed a constrained optimization framework and applied three heuristic optimization methods and two gradient-based optimization methods, and.

Indian Institute of Technology Guwahati

6

Conclusions

6.1 Conclusions

Brain surgery is performed with the assistance of an image-guided NN system. The position of the surgical equipment is mapped to a pre-operative MR image through a fiducial marker-based calibration. However, often the NN system cannot track the diseased location due to the brain-shift. A BUS imaging system is used to monitor the surgical process and to mitigate the effect of brain-shift, which can be solved by an effective image registration method. pBUS images are captured after opening the skull and iBUS images are captured during the surgery. Similarity between pBUS and the corresponding iBUS images gets affected due to various reasons, which makes the registration challenging. The effect of brain-shift is neutralized by registering the pBUS and the iBUS images.

This thesis investigated the problems of registration of pBUS and iBUS images. The primary difficulties in this registration is the poor similarity between pBUS and iBUS images. Three different methods were proposed for improving the similarity between the registering images and registration is carried out. The methods are summarized below:

- In the first approach, we extracted prominent edge-rich regions in the pBUS and iBUS images. We adopted a patch-based approach to reduce the effect of noise and thresholded the gradient of the average patch intensities. Rigid registration was performed on the edge-rich regions by minimizing the MSE. The limitations of gradient-based optimizations were highlighted and heuristic optimizations like PSO, BBO, and SA were applied. For performance evaluation, common tag-points were marked on both of the images. The performance of this method was compared with two other recent methods. The mTREs before and after registration were noted for all competing methods. Results show that the PSO performs better than other optimization methods for registering image pairs of easy and medium categories. This method was unable to register the images having very less similarity between them.

- For improving the similarity between the registering images, we proposed a segmentation-based approach which extracts the HE regions from the registering image pairs. This is a fully automatic method, and needs no user intervention. We applied a simple patch-based segmentation method for extracting the HE regions from the registering image pairs. The proposed method employed the patch-based thresholding, boundary correction,

cubic B-spline contour smoothing to mitigate the effects of artifacts and to avoid complex optimization. Rigid registration of the segmented HE regions of the registering image pairs is performed by minimizing the MSE with various heuristic optimization methods. The said method outperformed other level-set-based methods, and a random-forest-based machine-learning method [141]. Among the various optimization methods, PSO was found to be the best performing algorithm. This method is capable of registering some difficult images, thus showing improvement over the previous method.

- There are some anechoic and hypoechoic regions like ventricle, resection cavity, etc. that appear prominently in BUS images. They may also work as good landmarks for registration. Therefore, the patch-based method was modified to a patch-based level-set method for curve evolution. The method is semi-automatic. Like the other level-set-based methods, initial level-set curve is defined by the user. Using a patch-based threshoding approach, the initial level-sets are estimated in one step. Two types of level-set curve movements called the ULSCF and the BLSCF were followed. The ULSCF step grows rapidly towards the outside region. The BLSCF step allows bidirectional movements, and in each iteration, the sides of the patch are reduced to half of its previous size. This step refines the boundary and drives the growth of segmentation toward the actual boundary. The resulting segmentation undergoes the next step of B-spline contour smoothing. The method performs segmentation of the longitudinal fissure, tumor, choroid plexus, ventricles and the resection cavity.

  During registration, both the pBUS and the iBUS images are placed side by side. The user identifies the common HE or NHE regions as mentioned above, and marks by drawing rectangles on them. The registration method consists of two steps: first, the search range is estimated using the positions of the rectangles marked by the user. The second step performs registration of the segmented regions of pBUS and iBUS images in the estimated search range minimizing the MSE. Various optimization methods were applied and their results were compared. All the methods applied here performed better because of narrowed search range. As LBFGS showed performance compared to the PSO and other heuristic methods, may be used for this registration method. This method showed significant improvement in registering difficult image pairs compared to the previous methods.

6.2 Future directions

This thesis proposed approaches for registering pBUS and iBUS images. The success of the methods was demonstrated on selected image pairs from the BITE and the RESECT datasets. The methods developed in the thesis leave some room for further research and a few research directions are outlined below.

1 We have modeled the deformation as a rigid transformation. The brain is made of soft tissues; a non-rigid transformation is a more practical consideration for this problem. The proposed methods of registration can be extended for the non-rigid deformation models.

2 We have not applied any speckle suppression method before segmentation and registration. There are plenty of methods available for speckle removal methods like speckle reducing anisotropic diffusion (SRAD), stochastic distance-based block-matching 3D (BM3D), and CNN-based speckle reduction techniques. Recently Feng *et al.* [60] proposed a CNN-based hybrid neural network for speckle reduction. The noise model was generated using a pre-trained VGGNet using natural images. Another VGGNet extracts the structural boundaries and enhance them. This method may be applied for de-speckling the BUS images prior registration.

3 Deep learning approaches for ultrasound image segmentation are popular and effective. Segmentation of an image greatly differs depending on acoustic shadows. Xu *et al.* [206] developed a Shadow-consistent Semi-supervised Learning (SCO-SSL) method for prostate US image segmentation. It introduces two novel mechanisms, namely shadow augmentation (Shadow-AUG) and shadow dropout (Shadow-DROP) for dealing with the shadow artifact. Shadow-AUG adds simulated shadow artifacts to the images and learn its patterns. Shadow-DROP network learns the prostate boundary using the neighboring shadow-free pixels. This method may be applied for the segmentation of various HE and NHE brain image regions and the registration performance based on these regions may be studied.

4 Formation of a brain ultrasound image is completely operator-dependent. Varying the types and orientation of probe, scan depth, frequencies, different images can be generated from the same scene. The BITE and RESECT datasets are not exhaustive. For 2D image registration, an exhaustive dataset with the annotated ground truths may be prepared.

Bibliography

[1] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels. Technical report, 2010.

[2] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. IEEE transactions on pattern analysis and machine intelligence, 34(11):2274–2282, 2012.

[3] Mohammad Alkhatib, Adel Hafiane, Pierre Vieyres, and Alain Delbos. Deep visual nerve tracking in ultrasound images. Computerized Medical Imaging and Graphics, 76:101639, 2019.

[4] Tal Arbel, Tal Arbel, Xavier Morandi, Roch M Comeau, and D Louis Collins. Automatic non-linear mri-ultrasound registration for the correction of intra-operative brain deformations. Computer Aided Surgery, 9(4):123–136, 2004.

[5] Babak A Ardekani, Stephen Guckemus, Alvin Bachman, Matthew J Hoptman, Michelle Wojtaszek, and Jay Nierenberg. Quantitative comparison of algorithms for inter-subject registration of 3d volumetric brain mri scans. Journal of neuroscience methods, 142(1):67–76, 2005.

[6] John Ashburner. A fast diffeomorphic image registration algorithm. Neuroimage, 38(1):95–113, 2007.

[7] M Adeel Azam, K Bahadar Khan, Muhammad Ahmad, and Manuel Mazzara. Multimodal medical image registration and fusion for quality enhancement. Computers, Materials & Continua, 68(1):821–840, 2021.

[8] Sophia Bano, Seong Young Ko, and Ferdinando Rodriguez y Baena. Smooth path planning for a biologically-inspired neurosurgical probe. In 2012 Annual International

Conference of the IEEE Engineering in Medicine and Biology Society, pages 920–923. IEEE, 2012.

[9] Giuseppe MV Barbagallo, Stefano Palmucci, Massimiliano Visocchi, Sabrina Paratore, Giancarlo Attinà, Giuseppe Sortino, Vincenzo Albanese, and Francesco Certo. Portable intraoperative computed tomography scan in image-guided surgery for brain high-grade gliomas: analysis of technical feasibility and impact on extent of tumor resection. Operative Neurosurgery, 12(1):19–30, 2016.

[10] E Bermejo, Manuel Chica, Sergio Damas, Sancho Salcedo-Sanz, and Oscar Cordón. Mono-modal medical image registration with coral reef optimization. In International Conference on Hybrid Artificial Intelligence Systems, pages 222–234. Springer, 2018.

[11] Enrique Bermejo, Manuel Chica, Sancho Salcedo Sanz, and Oscar Cordón. Coral reef optimization for intensity-based medical image registration. In 2017 IEEE Congress on Evolutionary Computation (CEC), pages 533–540. IEEE, 2017.

[12] Olivier Bernard, Denis Friboulet, Philippe Thévenaz, and Michael Unser. Variational b-spline level-set: a linear filtering approach for fast deformable model evolution. IEEE Transactions on Image Processing, 18(6):1179–1191, 2009.

[13] Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In Sensor fusion IV: control paradigms and data structures, volume 1611, pages 586–606. Spie, 1992.

[14] Peter McL Black, Thomas Moriarty, Eben Alexander, Philip Stieg, Eric J Woodard, P Langham Gleason, Claudia H Martin, Ron Kikinis, Richard B Schwartz, and Ferenc A Jolesz. Development and implementation of intraoperative magnetic resonance imaging and its neurosurgical applications. Neurosurgery, 41(4):831–845, 1997.

[15] Richard D Bucholz, David D Yeh, Jason Trobaugh, Leslie L McDurmont, Christopher D Sturm, Carol Baumann, Jaimie M Henderson, Ari Levy, and Paul Kessman. The correction of stereotactic inaccuracy caused by brain shift using an intraoperative ultrasound device. In CVRMed-MRCAS'97, pages 459–466. Springer, 1997.

[16] Silvia Budday, Richard Nay, Rijk de Rooij, Paul Steinmann, Thomas Wyrobek, Timothy C Ovaert, and Ellen Kuhl. Mechanical properties of gray and white matter brain tissue

by indentation. Journal of the mechanical behavior of biomedical materials, 46:318–330, 2015.

[17] Richard H Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. A limited memory algorithm for bound constrained optimization. SIAM Journal on scientific computing, 16(5):1190–1208, 1995.

[18] Emilio F Campana, Giovanni Fasano, and Antonio Pinto. Dynamic analysis for the selection of parameters and initial population, in particle swarm optimization. Journal of Global Optimization, 48(3):347–397, 2010.

[19] Luca Canalini, Jan Klein, Dorothea Miller, and Ron Kikinis. Registration of ultrasound volumes based on euclidean distance transform. In Large-Scale Annotation of Biomedical Data and Expert Label Synthesis and Hardware Aware Learning for Medical Imaging and Computer Assisted Intervention: International Workshops, LABELS 2019, HAL-MICCAI 2019, and CuRIOUS 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13 and 17, 2019, Proceedings 4, pages 127–135. Springer, 2019.

[20] Luca Canalini, Jan Klein, Dorothea Miller, and Ron Kikinis. Segmentation-based registration of ultrasound volumes for glioma resection in image-guided neurosurgery. International journal of computer assisted radiology and surgery, 14(10):1697–1713, 2019.

[21] Luca Canalini, Jan Klein, Dorothea Miller, and Ron Kikinis. Enhanced registration of ultrasound volumes by segmentation of resection cavity in neurosurgical procedures. International journal of computer assisted radiology and surgery, 15:1963–1974, 2020.

[22] DE Carl. A practical guide to splines. Springer, 2001.

[23] Kristin Carmody, Christopher Moore, and David Feller-Kopman. Handbook of critical care and emergency ultrasound. McGraw Hill Professional, 2011.

[24] Vicent Caselles, Ron Kimmel, and Guillermo Sapiro. Geodesic active contours. International journal of computer vision, 22(1):61–79, 1997.

[25] Philippe C Cattin, Herbert Bay, Luc Van Gool, and Gábor Székely. Retina mosaicing using local features. Lecture Notes in Computer Science, 4191:185, 2006.

[26] AG Chacko, NKS Kumar, G Chacko, R Athyal, and V Rajshekhar. Intraoperative ultrasound in determining the extent of resection of parenchymal brain tumours–a comparative study with computed tomography and histopathology. Acta neurochirurgica, 145(9):743–748, 2003.

[27] Joe Chalfoun, Michael Majurski, Adele Peskin, Catherine Breen, Peter Bajcsy, and Mary Brady. Empirical gradient threshold technique for automated segmentation across image modalities and cell lines. Journal of microscopy, 260(1):86–99, 2015.

[28] Tony F Chan and Luminita A Vese. Active contours without edges. IEEE Transactions on image processing, 10(2):266–277, 2001.

[29] William F Chandler, James E Knake, John E McGillicuddy, Kevin O Lillehei, and Terry M Silver. Intraoperative use of real-time ultrasonography in neurosurgery. Journal of neurosurgery, 57(2):157–163, 1982.

[30] Samprit Chatterjee and Jeffrey S Simonoff. Handbook of regression analysis, volume 5. John Wiley & Sons, 2013.

[31] Haradhan Chel and PK Bora. A novel outlier detection based approach to registering pre- and post-resection ultrasound brain tumor images. In 2017 4th International Conference on Advances in Electrical Engineering (ICAEE), pages 476–481. IEEE, 2017.

[32] Haradhan Chel, PK Bora, and KK Ramchiary. A fast technique for hyper-echoic region separation from brain ultrasound images using patch based thresholding and cubic b-spline based contour smoothing. Ultrasonics, 111:106304, 2021.

[33] Haradhan Chel, Debashis Nandi, and PK Bora. Image registration in presence of multiplicative noise by particle swarm optimization. In 2015 Third International Conference on Image Information Processing (ICIIP), pages 135–140. IEEE, 2015.

[34] Jian Chen, Jie Tian, Noah Lee, Jian Zheng, R Theodore Smith, and Andrew F Laine. A partial intensity invariant feature descriptor for multimodal retinal image registration. IEEE Transactions on Biomedical Engineering, 57(7):1707–1718, 2010.

[35] Mingli Chen, Weiguo Lu, Quan Chen, Kenneth J Ruchala, and Gustavo H Olivera. A simple fixed-point approach to invert a deformation field a. Medical physics, 35(1):81–88, 2008.

[36] Yilin Chen, Fazhi He, Haoran Li, Dejun Zhang, and Yiqi Wu. A full migration bbo algorithm with enhanced population quality bounds for multimodal biomedical image registration. Applied Soft Computing, 93:106335, 2020.

[37] Rui Chibante. Simulated annealing: theory with applications. BoD–Books on Demand, 2010.

[38] Gary E Christensen and Hans J Johnson. Consistent image registration. IEEE transactions on medical imaging, 20(7):568–582, 2001.

[39] Marcelo Cicconet. Imageannotationbot, https://www.mathworks.com/matlabcentral/fileexchan imageannotationbot. 2021.

[40] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19, pages 424–432. Springer, 2016.

[41] Mayfield Clinic. Craniotomy, mayfield certified health info materials are written and developed by the mayfield clinic, https://mayfieldclinic.com/pe-craniotomy.htm. 2021.

[42] Mary In-Ping Huang Cobb, Jeffrey M Taekman, Ali R Zomorodi, L Fernando Gonzalez, and Dennis A Turner. Simulation in neurosurgery—a brief review and commentary. World neurosurgery, 89:583–586, 2016.

[43] Roch M Comeau, Aaron Fenster, and Terry M Peters. Intraoperative us in interactive image-guided neurosurgery. Radiographics, 18(4):1019–1027, 1998.

[44] Mohammad I Daoud, Ayman A Atallah, Falah Awwad, Mahasen Al-Najjar, and Rami Alazrai. Automatic superpixel-based segmentation method for breast ultrasound images. Expert Systems with Applications, 121:78–96, 2019.

[45] Dante De Nigris, D Louis Collins, and Tal Arbel. Multi-modal image registration based on gradient orientations of minimal uncertainty. IEEE transactions on medical imaging, 31(12):2343–2354, 2012.

[46] Kalyanmoy Deb. Multi-Objective Optimization using Evolutionary Algorithms. WILEY INDIA Pvt. Ltd., 2010.

[47] Kalyanmoy Deb. Optimization for engineering design: Algorithms and examples. PHI Learning Pvt. Ltd., 2012.

[48] Maria del Mar Vila, Beatriz Remeseiro, Maria Grau, Roberto Elosua, Àngels Betriu, Elvira Fernandez-Giraldez, and Laura Igual. Semantic segmentation with densenets for carotid artery ultrasound plaque segmentation and cimt estimation. Artificial Intelligence in Medicine, 103:101784, 2020.

[49] Hedifa Dida, Fella Charif, and Abderrazak Benchabane. Grey wolf optimizer for multi-modal medical image registration. In 2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS), pages 1–5. IEEE, 2020.

[50] Thomas Dietenbeck, Martino Alessandrini, Denis Friboulet, and Olivier Bernard. Creaseg: a free software for the evaluation of image segmentation algorithms based on level-set. In 2010 IEEE International Conference on Image Processing, pages 665–668. IEEE, 2010.

[51] Luke Dixon, Adrian Lim, Matthew Grech-Sollars, Dipankar Nandi, and Sophie Camp. Intraoperative ultrasound in brain tumor surgery: A review and implementation guide. Neurosurgical Review, 45(4):2503–2515, 2022.

[52] George J Dohrmann and Jonathan M Rubin. History of intraoperative ultrasound in neurosurgery. Neurosurgery Clinics of North America, 12(1):155–66, 2001.

[53] TS Douglas, SE Solomonidis, WA Sandham, and WD Spence. Ultrasound image matching using genetic algorithms. Medical and Biological Engineering and Computing, 40:168–172, 2002.

[54] Juan Du, Songyuan Tang, Tianzi Jiang, and Zhensu Lu. Intensity-based robust similarity for multimodal image registration. International Journal of Computer Mathematics, 83(1):49–57, 2006.

[55] Goretti Echegaray, Imanol Herrera, Iker Aguinaga, Carlos Buchart, and Diego Borro. A brain surgery simulator. IEEE computer Graphics and Applications, 34(3):12–18, 2014.

[56] E Efstathiou[1], TM Deserno, C Münzenmayer[1], T Wittenberg[1], and T Bergen[1]. Evaluation of 3d ultrasound image registration. 2011.

[57] Martin Engelhardt, Christian Hansen, Jens Eyding, Wilko Wilkening, Christopher Brenke, Christos Krogias, Martin Scholz, Albrecht Harders, Helmut Ermert, and Kirsten Schmieder. Feasibility of contrast-enhanced sonography during resection of cerebral tumours: initial results of a prospective study. Ultrasound in medicine & biology, 33(4):571–575, 2007.

[58] Carlos Faria, Wolfram Erlhagen, Manuel Rito, Elena De Momi, Giancarlo Ferrigno, and Estela Bicho. Review of robotic technology for stereotactic neurosurgery. IEEE reviews in biomedical engineering, 8:125–137, 2015.

[59] Parastoo Farnia, Alireza Ahmadian, Alireza Khoshnevisan, AmirHossein Jaberzadeh, Nasim Dadashi Serej, and Anahita F Kazerooni. An efficient point based registration of intra-operative ultrasound images with mr images for computation of brain shift; a phantom study. In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 8074–8077. IEEE, 2011.

[60] Xiangfei Feng, Qinghua Huang, and Xuelong Li. Ultrasound image de-speckling by a hybrid deep network with transferred filtering and structural prior. Neurocomputing, 414:346–355, 2020.

[61] Zishun Feng, Dong Nie, Li Wang, and Dinggang Shen. Semi-supervised learning for pelvic mr image segmentation based on multi-task residual fully convolutional networks. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pages 885–888. IEEE, 2018.

[62] Enzo Ferrante and Nikos Paragios. Non-rigid 2d-3d medical image registration using markov random fields. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013: 16th International Conference, Nagoya, Japan, September 22-26, 2013, Proceedings, Part III 16, pages 163–170. Springer, 2013.

[63] Karl J Friston, John Ashburner, Christopher D Frith, J-B Poline, John D Heather, and Richard SJ Frackowiak. Spatial registration and normalization of images. Human brain mapping, 3(3):165–189, 1995.

[64] Arnaud Gelas, Olivier Bernard, Denis Friboulet, and Rmy Prost. Compactly supported radial basis functions based collocation method for level-set evolution in image segmentation. IEEE Transactions on Image Processing, 16(7):1873–1887, 2007.

[65] Soumya Ghose, Arnau Oliver, Jhimli Mitra, Robert Martí, Xavier Lladó, Jordi Freixenet, Désiré Sidibé, Joan C Vilanova, Josep Comet, and Fabrice Meriaudeau. A supervised learning framework of statistical shape and probability priors for automatic prostate segmentation in ultrasound images. Medical image analysis, 17(6):587–600, 2013.

[66] Giuseppe Roberto Giammalva, Gianluca Ferini, Sofia Musso, Giuseppe Salvaggio, Maria Angela Pino, Rosa Maria Gerardi, Lara Brunasso, Roberta Costanzo, Federica Paolini, Rina Di Bonaventura, et al. Intraoperative ultrasound: emerging technology and novel applications in brain tumor surgery. Frontiers in Oncology, 12:46, 2022.

[67] Rémi Giraud, Vinh-Thong Ta, Aurélie Bugeau, Pierrick Coupé, and Nicolas Papadakis. Superpatchmatch: An algorithm for robust correspondences using superpixel patches. IEEE Transactions on Image Processing, 26(8):4068–4078, 2017.

[68] Rafael C Gonzalez. Digital image processing. Pearson education india, 2009.

[69] Rafael C Gonzalez, Richard E Woods, et al. Digital image processing. Pearson Education Inc, US, 2002.

[70] Alain Goriely, Marc GD Geers, Gerhard A Holzapfel, Jayaratnam Jayamohan, Antoine Jérusalem, Sivabal Sivaloganathan, Waney Squier, Johannes AW van Dommelen, Sarah Waters, and Ellen Kuhl. Mechanics of the brain: perspectives, challenges, and opportunities. Biomechanics and modeling in mechanobiology, 14(5):931–965, 2015.

[71] A Gronningsaeter, G Unsgård, S Ommedal, and BAJ Angelsen. Ultrasound-guided neurosurgery: A feasibility study in the 3-30 mhz frequency range. British journal of neurosurgery, 10(2):161–168, 1996.

[72] Houssem-Eddine Gueziri, Oded Rabau, Carlo Santaguida, and D Louis Collins. Evaluation of an ultrasound-based navigation system for spine neurosurgery: a porcine cadaver study. Frontiers in Oncology, 11:619204, 2021.

[73] Maarouf A Hammoud, B Lee Ligon, Rabih Elsouki, Wei Ming Shi, Donald F Schomer, and Raymond Sawaya. Use of intraoperative ultrasound for localizing tumors and determining the extent of resection: a comparative study with magnetic resonance imaging. Journal of neurosurgery, 84(5):737–741, 1996.

[74] Luyi Han, Yunzhi Huang, Haoran Dou, Shuai Wang, Sahar Ahamad, Honghao Luo, Qi Liu, Jingfan Fan, and Jiang Zhang. Semi-supervised segmentation of lesion from breast ultrasound images with attentional generative adversarial network. Computer Methods and Programs in Biomedicine, 189:105275, 2020.

[75] Xiaohui Hao, Charles J Bruce, Cristina Pislaru, and James F Greenleaf. Segmenting high-frequency intracardiac ultrasound images of myocardium into infarcted, ischemic, and normal regions. IEEE transactions on medical imaging, 20(12):1373–1383, 2001.

[76] Renjie He and Ponnada A Narayana. Global optimization of mutual information: application to three-dimensional retrospective registration of magnetic resonance images. Computerized medical imaging and graphics, 26(4):277–292, 2002.

[77] Mattias P Heinrich, Mark Jenkinson, Manav Bhushan, Tahreema Matin, Fergus V Gleeson, Michael Brady, and Julia A Schnabel. Mind: Modality independent neighbourhood descriptor for multi-modal deformable registration. Medical image analysis, 16(7):1423–1435.

[78] Gerardo Hermosillo and Olivier Faugeras. Dense image matching with global and local statistical criteria: a variational approach. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, volume 1, pages I–I. IEEE, 2001.

[79] Danying Hu, Yuanzheng Gong, Blake Hannaford, and Eric J Seibel. Path planning for semi-automated simulated robotic neurosurgery. In 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 2639–2645. IEEE, 2015.

[80] Yujin Hu, Bei Xia, Muyi Mao, Zelong Jin, Jie Du, Libao Guo, Alejandro F Frangi, Baiying Lei, and Tianfu Wang. Aidan: An attention-guided dual-path network for pediatric echocardiography segmentation. Ieee Access, 8:29176–29187, 2020.

[81] Xishi Huang, Jing Ren, Gerard Guiraudon, Derek Boughner, and Terry M Peters. Rapid dynamic image registration of the beating heart for diagnosis and surgical navigation. IEEE transactions on medical imaging, 28(11):1802–1814, 2009.

[82] Yu-Len Huang and Dar-Ren Chen. Watershed segmentation for breast tumor in 2-d sonography. Ultrasound in medicine & biology, 30(5):625–632, 2004.

[83] Barys Ihnatsenka and André Pierre Boezaart. Ultrasound: Basic understanding and learning the language. International journal of shoulder surgery, 4(3):55, 2010.

[84] Songbai Ji, Ziji Wu, Alex Hartov, David W Roberts, and Keith D Paulsen. Mutual-information-based image to patient re-registration using intraoperative ultrasound in image-guided neurosurgery. Medical physics, 35(10):4612–4624, 2008.

[85] Bilge Karaçali. Information theoretic deformable registration using local image information. International journal of computer vision, 72(3):219–237, 2007.

[86] Davood Karimi, Qi Zeng, Prateek Mathur, Apeksha Avinash, Sara Mahdavi, Ingrid Spadinger, Purang Abolmaesumi, and Septimiu E Salcudean. Accurate and robust deep learning-based segmentation of the prostate clinical target volume in ultrasound images. Medical image analysis, 57:186–196, 2019.

[87] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. International journal of computer vision, 1(4):321–331, 1988.

[88] Henrich Kele. Ultrasonography of the peripheral nervous system. Perspectives in Medicine, 1(1-12):417–421, 2012.

[89] James Kennedy and Russell Eberhart. Particle swarm optimization. In Proceedings of ICNN'95-international conference on neural networks, volume 4, pages 1942–1948. IEEE, 1995.

[90] Alaa Khadidos, Victor Sanchez, and Chang-Tsun Li. Weighted level set evolution based on local edge features for medical image segmentation. IEEE Transactions on Image Processing, 26(4):1979–1991, 2017.

[91] Stefan Klein, Uulke A Van Der Heide, Irene M Lips, Marco Van Vulpen, Marius Staring, and Josien PW Pluim. Automatic segmentation of the prostate in 3d mr images by atlas matching using localized mutual information. Medical physics, 35(4):1407–1417, 2008.

[92] Jochen F Krucker, Gerald L LeCarpentier, J Brian Fowlkes, and Paul L Carson. Rapid elastic image registration for 3-d ultrasound. IEEE Transactions on Medical Imaging, 21(11):1384–1394, 2002.

[93] Ankur N Kumar, Michael I Miga, Thomas S Pheiffer, Lola B Chambless, Reid C Thompson, and Benoit M Dawant. Persistent and automatic intraoperative 3d digitization of surfaces under dynamic magnifications of an operating microscope. Medical image analysis, 19(1):30–45, 2015.

[94] Jen-wei Kuo, Ziming Qiu, Orlando Aristizabal, Jonathan Mamou, Daniel H Turnbull, Jeffrey Ketterling, and Yao Wang. Automatic body localization and brain ventricle segmentation in 3d high frequency ultrasound images of mouse embryos. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pages 635–639. IEEE, 2018.

[95] Shawn Lankton and Allen Tannenbaum. Localizing region-based active contours. IEEE transactions on image processing, 17(11):2029–2039, 2008.

[96] Marloes MJ Letteboer, Peter WA Willems, Max A Viergever, and Wiro J Niessen. Non-rigid registration of 3d ultrasound images of brain tumours acquired during neurosurgery. In Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003: 6th International Conference, Montréal, Canada, November 15-18, 2003. Proceedings 6, pages 408–415. Springer, 2003.

[97] Chunming Li, Chiu-Yen Kao, John C Gore, and Zhaohua Ding. Minimization of region-scalable fitting energy for image segmentation. IEEE transactions on image processing, 17(10):1940–1949, 2008.

[98] Chunming Li, Chenyang Xu, Changfeng Gui, and Martin D Fox. Distance regularized level set evolution and its application to image segmentation. IEEE transactions on image processing, 19(12):3243–3254, 2010.

[99] Haoming Li, Jinghui Fang, Shengfeng Liu, Xiaowen Liang, Xin Yang, Zixin Mai, Manh The Van, Tianfu Wang, Zhiyi Chen, and Dong Ni. Cr-unet: A composite network for ovary and follicle segmentation in ultrasound images. IEEE journal of biomedical and health informatics, 2019.

[100] Qi Li and Isao Sato. Multimodality image registration by particle swarm optimization of mutual information. In International Conference on Intelligent Computing, pages 1120–1130. Springer, 2007.

[101] Bo Liu, Heng-Da Cheng, Jianhua Huang, Jiawei Tian, Xianglong Tang, and Jiafeng Liu. Probability density difference-based active contour for ultrasound image segmentation. Pattern Recognition, 43(6):2028–2042, 2010.

[102] Dong C Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. Mathematical programming, 45(1-3):503–528, 1989.

[103] Qingyi Liu, Mingyan Jiang, Peirui Bai, and Guang Yang. A novel level set model with automated initialization and controlling parameters for medical image segmentation. Computerized Medical Imaging and Graphics, 48:21–29, 2016.

[104] Haiping Ma, Dan Simon, Minrui Fei, et al. On the convergence of biogeography-based optimization for binary problems. Mathematical Problems in Engineering, 2014, 2014.

[105] Jinlian Ma, Fa Wu, Tian'an Jiang, Qiyu Zhao, and Dexing Kong. Ultrasound image-based thyroid nodule automatic segmentation using convolutional neural networks. International journal of computer assisted radiology and surgery, 12:1895–1910, 2017.

[106] Lei Ma, Hidemichi Kiyomatsu, Keiichi Nakagawa, Junchen Wang, Etsuko Kobayashi, and Ichiro Sakuma. Accurate vessel segmentation in ultrasound images using a local-phase-based snake. Biomedical Signal Processing and Control, 43:236–243, 2018.

[107] Inês Machado, Matthew Toews, Elizabeth George, Prashin Unadkat, Walid Essayed, Jie Luo, Pedro Teodoro, Herculano Carvalho, Jorge Martins, Polina Golland, et al. Deformable mri-ultrasound registration using correlation-based attribute matching for brain shift correction: Accuracy and generality in multi-site data. NeuroImage, 202:116094, 2019.

[108] Inês Machado, Matthew Toews, Jie Luo, Prashin Unadkat, Walid Essayed, Elizabeth George, Pedro Teodoro, Herculano Carvalho, Jorge Martins, Polina Golland, et al. Non-rigid registration of 3d ultrasound for neurosurgery using automatic feature detection and matching. International journal of computer assisted radiology and surgery, 13(10):1525–1538, 2018.

[109] Junji Machi, B Sigel, JJ Jafar, R Menoni, JC Beitler, RA Bernstein, RM Crowell, JR Ramos, and DG Spigos. Criteria for using imaging ultrasound during brain and spinal cord surgery. Journal of Ultrasound in Medicine, 3(4):155–161, 1984.

[110] Anant Madabhushi and Dimitris N Metaxas. Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions. IEEE transactions on medical imaging, 22(2):155–169, 2003.

[111] Frederik Maes, Andre Collignon, Dirk Vandermeulen, Guy Marchal, and Paul Suetens. Multimodality image registration by maximization of mutual information. IEEE transactions on Medical Imaging, 16(2):187–198, 1997.

[112] Omid Haji Maghsoudi. Superpixel based segmentation and classification of polyps in wireless capsule endoscopy. In 2017 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), pages 1–4. IEEE, 2017.

[113] Ilias Maglogiannis and Elias Zafiropoulos. Automated medical image registration using the simulated annealing algorithm. In Hellenic Conference on Artificial Intelligence, pages 456–465. Springer, 2004.

[114] Dwarikanath Mahapatra. Semi-supervised learning and graph cuts for consensus based medical image segmentation. Pattern recognition, 63:700–709, 2017.

[115] JB Antoine Maintz and Max A Viergever. A survey of medical image registration. Medical image analysis, 2(1):1–36, 1998.

[116] Masatoshi Makuuchi, Guido Torzilli, and Junji Machi. History of intraoperative ultrasound. Ultrasound in medicine & biology, 24(9):1229–1242, 1998.

[117] Ezio Malis. Improving vision-based control using efficient second-order minimization techniques. In IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004, volume 2, pages 1843–1848. IEEE, 2004.

[118] Marcos Martın-Fernández and Carlos Alberola-Lopez. An approach for contour detection of human kidneys from ultrasound images using markov random fields and active contours. Medical image analysis, 9(1):1–23, 2005.

[119] Nima Masoumi, Yiming Xiao, and Hassan Rivaz. Marcel (inter-modality affine registration with correlation ratio): An application for brain shift correction in ultrasound-guided brain tumor resection. In Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers 3, pages 55–63. Springer, 2018.

[120] Matthew J McGirt, Frank J Attenello, Ghazala Datoo, Muraya Gathinji, April Atiba, Jon D Weingart, Benjamin Carson, and George I Jallo. Intraoperative ultrasonography as a guide to patient selection for duraplasty after suboccipital decompression in children with chiari malformation type i. Journal of Neurosurgery: Pediatrics, 2(1):52–57, 2008.

[121] Qingjie Meng, Matthew Sinclair, Veronika Zimmer, Benjamin Hou, Martin Rajchl, Nicolas Toussaint, Ozan Oktay, Jo Schlemper, Alberto Gomez, James Housden, et al. Weakly supervised estimation of shadow confidence maps in fetal ultrasound imaging. IEEE transactions on medical imaging, 38(12):2755–2767, 2019.

[122] Laurence Mercier, David Araujo, Claire Haegelen, Rolando F Del Maestro, Kevin Petrecca, and D Louis Collins. Registering pre-and postresection 3-dimensional ultrasound

for improved visualization of residual brain tumor. Ultrasound in medicine & biology, 39(1):16–29, 2013.

[123] Laurence Mercier, Rolando F Del Maestro, Kevin Petrecca, David Araujo, Claire Haegelen, and D Louis Collins. Online database of clinical mr and ultrasound images of brain tumors. Medical physics, 39(6Part1):3253–3261, 2012.

[124] Laurence Mercier, Rolando F Del Maestro, Kevin Petrecca, Anna Kochanowska, Simon Drouin, Charles XB Yan, Andrew L Janke, Sean Jy-Shyang Chen, and D Louis Collins. New prototype neuronavigation system based on preoperative imaging and intraoperative freehand ultrasound: system description and validation. International journal of computer assisted radiology and surgery, 6(4):507–522, 2011.

[125] Laurence Mercier, Vladimir Fonov, Claire Haegelen, Rolando F Del Maestro, Kevin Petrecca, and D Louis Collins. Comparing two approaches to rigid registration of three-dimensional ultrasound and magnetic resonance images for neurosurgery. International journal of computer assisted radiology and surgery, 7(1):125–136, 2012.

[126] Charles R Meyer, Jennifer L Boes, Boklye Kim, Peyton H Bland, Gerald L Lecarpentier, J Brian Fowlkes, Marilyn A Roubidoux, and Paul L Carson. Semiautomatic registration of volumetric ultrasound scans. Ultrasound in medicine & biology, 25(3):339–347, 1999.

[127] Seyedali Mirjalili, Seyed Mohammad Mirjalili, and Andrew Lewis. Grey wolf optimizer. Advances in engineering software, 69:46–61, 2014.

[128] Deepak Mishra, Santanu Chaudhury, Mukul Sarkar, and Arvinder Singh Soin. Ultrasound image segmentation: a deeply supervised network with attention to boundaries. IEEE Transactions on Biomedical Engineering, 66(6):1637–1648, 2018.

[129] Are Mjaavatten. Curvature of a 2d or 3d curve, 2019.

[130] Beadaa J Mohammed, Amin M Abbosh, Samah Mustafa, and David Ireland. Microwave system for head imaging. IEEE Transactions on Instrumentation and Measurement, 63(1):117–123, 2013.

[131] Aliasgar V Moiyadi. Objective assessment of intraoperative ultrasound in brain tumors. Acta neurochirurgica, 156(4):703, 2014.

[132] Aliasgar V Moiyadi, Sadhana Kannan, Prakash Shetty, et al. Navigated intraoperative ultrasound for resection of gliomas: predictive value, influence on resection and survival. Neurology India, 63(5):727, 2015.

[133] Aryan Mokhtari and Alejandro Ribeiro. Global convergence of online limited memory bfgs. The Journal of Machine Learning Research, 16(1):3151–3181, 2015.

[134] Shakiba Moradi, Mostafa Ghelich Oghli, Azin Alizadehasl, Isaac Shiri, Niki Oveisi, Mehrdad Oveisi, Majid Maleki, and Jan Dhooge. Mfp-unet: A novel deep learning based approach for left ventricle segmentation in echocardiography. Physica Medica, 67:58–69, 2019.

[135] Marius Muja and David G Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. VISAPP (1), 2(331-340):2, 2009.

[136] Miguel Mulet-Parada and J Alison Noble. 2d+ t acoustic boundary detection in echocardiography. Medical image analysis, 4(1):21–30, 2000.

[137] David Mumford and Jayant Shah. Optimal approximations by piecewise smooth functions and associated variational problems. Communications on pure and applied mathematics, 42(5):577–685, 1989.

[138] Wayne Niblack. An introduction to digital image processing. Strandberg Publishing Company, 1985.

[139] Dante De Nigris, Laurence Mercier, Rolando Del Maestro, D Louis Collins, and Tal Arbel. Hierarchical multimodal image registration based on adaptive local mutual information. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 643–651. Springer, 2010.

[140] S Nirmala and S Sridevi. Markov random field segmentation based sonographic identification of prenatal ventricular septal defect. Procedia Computer Science, 79:344–350, 2016.

[141] J Nitsch, J Klein, P Dammann, K Wrede, O Gembruch, JH Moltz, H Meine, U Sure, R Kikinis, and D Miller. Automatic and efficient mri-us segmentations for improving intra-

operative image fusion in image-guided neurosurgery. NeuroImage: Clinical, 22:101766, 2019.

[142] Jennifer Nitsch, Jan Klein, Dorothea Miller, Ulrich Sure, and Horst K Hahn. Automatic segmentation of the cerebral falx and adjacent gyri in 2d ultrasound images. pages 287–292, 2015.

[143] Alison Noble and Djamal Boukerroui. Ultrasound image segmentation: a survey. IEEE Transactions on medical imaging, 25(8):987–1010, 2006.

[144] Kevin D Oakeson, Hui Zhu, and Morton H Friedman. Quantification of cross-sectional artery wall motion with ivus image registration. In Medical Imaging 2004: Ultrasonic Imaging and Signal Processing, volume 5373, pages 119–130. SPIE, 2004.

[145] Johan Öfverstedt, Joakim Lindblad, and Nataša Sladoje. Fast computation of mutual information in the frequency domain with applications to global multimodal image alignment. Pattern Recognition Letters, 159:196–203, 2022.

[146] Stanley Osher and James A Sethian. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. Journal of computational physics, 79(1):12–49, 1988.

[147] Nobuyuki Otsu. A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics, 9(1):62–66, 1979.

[148] Niko Pagoulatos, Warren S Edwards, David R Haynor, and Yongmin Kim. Interactive 3d registration of ultrasound and magnetic resonance images based on a magnetic position sensor. IEEE Transactions on Information Technology in Biomedicine, 3(4):278–288, 1999.

[149] Nikos Paragios and Rachid Deriche. Geodesic active regions and level set methods for supervised texture segmentation. International Journal of Computer Vision, 46(3):223–247, 2002.

[150] Xavier Pennec, Nicholas Ayache, Alexis Roche, and Pascal Cachier. Non-rigid mr/us registration for tracking brain deformations. In Proceedings International Workshop on Medical Imaging and Augmented Reality, pages 79–86. IEEE, 2001.

[151] Graeme P Penney, Jane M Blackall, MS Hamady, T Sabharwal, A Adam, and David J Hawkes. Registration of freehand 3d ultrasound and magnetic resonance liver images. Medical image analysis, 8(1):81–91, 2004.

[152] Maria Angela Pino, Alessia Imperato, Irene Musca, Rosario Maugeri, Giuseppe Roberto Giammalva, Gabriele Costantino, Francesca Graziano, Francesco Meli, Natale Francaviglia, Domenico Gerardo Iacopino, et al. New hope in brain glioma surgery: the role of intraoperative ultrasound. a review. Brain Sciences, 8(11):202, 2018.

[153] Ioannis Pitas. Digital image processing algorithms and applications. 2000.

[154] Josien PW Pluim, JB Antoine Maintz, and Max A Viergever. Mutual-information-based registration of medical images: a survey. IEEE transactions on medical imaging, 22(8):986–1004, 2003.

[155] Francesco Prada, Massimiliano Del Bene, Riccardo Fornaro, Ignazio G Vetrano, Alberto Martegani, Luca Aiani, Luca Maria Sconfienza, Giovanni Mauri, Luigi Solbiati, Bianca Pollo, et al. Identification of residual tumor with intraoperative contrast-enhanced ultrasound during glioblastoma resection. Neurosurgical focus, 40(3):E7, 2016.

[156] Ruowei Qu, Guizhi Xu, Chunxia Ding, Wenyan Jia, and Mingui Sun. Deep learning-based methodology for recognition of fetal brain standard scan planes in 2d ultrasound images. IEEE Access, 8:44443–44451, 2019.

[157] Reza Moradi Rad, Parvaneh Saeedi, Jason Au, and Jon Havelock. Trophectoderm segmentation in human embryo images via inceptioned u-net. Medical Image Analysis, 62:101612, 2020.

[158] V Rajinikanth, Nilanjan Dey, Rajesh Kumar, John Panneerselvam, and N Sri Madhava Raja. Fetal head periphery extraction from ultrasound image using jaya algorithm and chan-vese segmentation. Procedia Computer Science, 152:66–73, 2019.

[159] I-A Rasmussen, F Lindseth, OM Rygh, EM Berntsen, T Selbekk, J Xu, TA Nagelhus Hernes, E Harg, A Håberg, and G Unsgaard. Functional neuronavigation combined with intra-operative 3d ultrasound: initial experiences during surgical resections close

to eloquent brain areas and future directions in automatic brain shift compensation of preoperative data. Acta neurochirurgica, 149(4):365–378, 2007.

[160] Qolamreza R Razlighi, Nasser Kehtarnavaz, and Siamak Yousefi. Evaluating similarity measures for brain image registration. Journal of visual communication and image representation, 24(7):977–987, 2013.

[161] Michael H Reid. Ultrasonic visualization of a cervical cord cystic astrocytoma. American Journal of Roentgenology, 131(5):907–908, 1978.

[162] Ingerid Reinertsen, Frank Lindseth, Christian Askeland, Daniel Høyer Iversen, and Geirmund Unsgård. Intra-operative correction of brain-shift. Acta neurochirurgica, 156(7):1301–1310, 2014.

[163] Robert Reisch, Axel Stadie, Ralf A Kockro, and Nikolai Hopf. The keyhole concept in neurosurgery. World neurosurgery, 79(2):S17–e9, 2013.

[164] Hassan Rivaz, Sean Jy-Shyang Chen, and D Louis Collins. Automatic deformable mr-ultrasound registration for image-guided neurosurgery. IEEE transactions on medical imaging, 34(2):366–380, 2014.

[165] Hassan Rivaz and D Louis Collins. Deformable registration of preoperative mr, pre-resection ultrasound, and post-resection ultrasound images of neurosurgery. International journal of computer assisted radiology and surgery, 10(7):1017–1028, 2015.

[166] Hassan Rivaz and D Louis Collins. Near real-time robust non-rigid registration of volumetric ultrasound images for neurosurgery. Ultrasound in medicine & biology, 41(2):574–587, 2015.

[167] Alexis Roche, Gregoire Malandain, and Nicholas Ayache. Unifying maximum likelihood approaches in medical image registration. International Journal of Imaging Systems and Technology, 11(1):71–80, 2000.

[168] Alexis Roche, Grégoire Malandain, Nicholas Ayache, and Sylvain Prima. Towards a better comprehension of similarity measures used in medical image registration. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 555–566. Springer, 1999.

[169] Alexis Roche, Grégoire Malandain, Xavier Pennec, and Nicholas Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 1115–1124. Springer, 1998.

[170] Alexis Roche, Xavier Pennec, Grégoire Malandain, and Nicholas Ayache. Rigid registration of 3-d ultrasound with mr images: a new approach combining intensity and gradient information. IEEE transactions on medical imaging, 20(10):1038–1049, 2001.

[171] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015.

[172] Jonathan M Rubin and GJ Dohrmann. Intraoperative neurosurgical ultrasound in the localization and characterization of intracranial masses. Radiology, 148(2):519–524, 1983.

[173] Sylvia Rueda, Sana Fathima, Caroline L Knight, Mohammad Yaqub, Aris T Papageorghiou, Bahbibi Rahmatullah, Alessandro Foi, Matteo Maggioni, Antonietta Pepe, Jussi Tohka, et al. Evaluation and comparison of current fetal ultrasound image segmentation methods for biometric measurements: a grand challenge. IEEE Transactions on medical imaging, 33(4):797–813, 2013.

[174] Leonardo Rundo, Andrea Tangherloni, Carmelo Militello, Maria Carla Gilardi, and Giancarlo Mauri. Multimodal medical image registration using particle swarm optimization: a review. In 2016 IEEE Symposium Series on Computational Intelligence (SSCI), pages 1–8. IEEE, 2016.

[175] CA Saether, M Torsteinsen, SH Torp, S Sundstrøm, G Unsgård, and O Solheim. Did survival improve after the implementation of intraoperative neuronavigation and 3d ultrasound in glioblastoma surgery? a retrospective analysis of 192 primary operations. Journal of Neurological Surgery Part A: Central European Neurosurgery, 73(02):073–078, 2012.

[176] Otmar Scherzer. Mathematical models for registration and applications to medical imaging. Springer, 2006.

[177] M Schwier, T Chitiboi, T Hülnhagen, and HK Hahn. Automated spine and vertebrae detection in ct images using object-based image analysis. International journal for numerical methods in biomedical engineering, 29(9):938–963, 2013.

[178] James A Sethian et al. Level set methods and fast marching methods, volume 98. Cambridge Cambridge UP, 1999.

[179] Reuben R Shamir, Leo Joskowicz, Idit Tamir, Elad Dabool, Lihi Pertman, Adam Ben-Ami, and Yigal Shoshan. Reduced risk trajectory planning in image-guided keyhole neurosurgery. Medical physics, 39(5):2885–2895, 2012.

[180] Raj Shekhar, Vladimir Zagrodsky, Mario J Garcia, and James D Thomas. Registration of real-time 3-d ultrasound images of the heart for novel 3-d stress echocardiography. IEEE Transactions on Medical Imaging, 23(9):1141–1149, 2004.

[181] Jonathan Richard Shewchuk et al. An introduction to the conjugate gradient method without the agonizing pain, 1994.

[182] Yonggang Shi and William Clem Karl. A real-time algorithm for the approximation of level-set-based curve evolution. IEEE transactions on image processing, 17(5):645–656, 2008.

[183] Dan Simon. Biogeography-based optimization. IEEE transactions on evolutionary computation, 12(6):702–713, 2008.

[184] Andrej Šteňo, Viktor Matejčík, and Juraj Šteňo. Intraoperative ultrasound in low-grade glioma surgery. Clinical neurology and neurosurgery, 135:96–99, 2015.

[185] George Stockman and Linda G Shapiro. Computer vision. Prentice Hall PTR, 2001.

[186] Colin Studholme, Valerie A Cardenas, and Michael W Weiner. Multiscale image and multiscale deformation of brain anatomy for building average brain atlases. In Medical Imaging 2001: Image Processing, volume 4322, pages 557–568. SPIE, 2001.

[187] Hassiba Talbi and Mohamed Batouche. Hybrid particle swarm with differential evolution for multimodal image registration. In 2004 IEEE International Conference on Industrial Technology, 2004. IEEE ICIT'04., volume 3, pages 1567–1572. IEEE, 2004.

[188] Neil WD Thomas and John Sinclair. Image-guided neurosurgery: history and current clinical applications. Journal of medical imaging and radiation sciences, 46(3):331–342, 2015.

[189] Helena R Torres, Sandro Queiros, Pedro Morais, Bruno Oliveira, Jaime C Fonseca, and Joao L Vilaca. Kidney segmentation in ultrasound, magnetic resonance and computed tomography images: A systematic review. Computer methods and programs in biomedicine, 157:49–67, 2018.

[190] Jason W Trobaugh, William D Richard, Kurt R Smith, and Richard D Bucholz. Frameless stereotactic ultrasonography: method and applications. Computerized Medical Imaging and Graphics, 18(4):235–246, 1994.

[191] Geirmund Unsgaard, Aage Gronningsaeter, Steinar Ommedal, and Toril A Nagelhus Hernes. Brain operations guided by real-time two-dimensional ultrasound: new possibilities as a result of improved image quality. Neurosurgery, 51(2):402–412, 2002.

[192] Geirmund Unsgaard, Steinar Ommedal, Tomm Muller, Aage Gronningsaeter, and Toril A Nagelhus Hernes. Neuronavigation by intraoperative three-dimensional ultrasound: initial experience during brain tumor resection. Neurosurgery, 50(4):804–812, 2002.

[193] Aleksandar Vakanski, Min Xian, and Phoebe E Freer. Attention-enriched deep learning model for breast tumor segmentation in ultrasound images. Ultrasound in Medicine & Biology, 2020.

[194] Andrea Valsecchi, Sergio Damas, Jose Santamaria, and Linda Marrakchi-Kacem. Genetic algorithms for voxel-based medical image registration. In 2013 Fourth International Workshop on Computational Intelligence in Medical Imaging (CIMI), pages 22–29. IEEE, 2013.

[195] Luminita A Vese and Tony F Chan. A multiphase level set framework for image segmentation using the mumford and shah model. International journal of computer vision, 50(3):271–293, 2002.

[196] Mark P Wachowiak, Renata Smolíková, Yufeng Zheng, Jacek M Zurada, and Adel Said Elmaghraby. An approach to multimodal biomedical image registration utilizing parti-

cle swarm optimization. IEEE Transactions on evolutionary computation, 8(3):289–301, 2004.

[197] Yi Wang, Haoran Dou, Xiaowei Hu, Lei Zhu, Xin Yang, Ming Xu, Jing Qin, Pheng-Ann Heng, Tianfu Wang, and Dong Ni. Deep attentive features for prostate segmentation in 3d transrectal ultrasound. IEEE transactions on medical imaging, 38(12):2768–2778, 2019.

[198] William M Wells III, Paul Viola, Hideki Atsumi, Shin Nakajima, and Ron Kikinis. Multi-modal volume registration by maximization of mutual information. Medical image analysis, 1(1):35–51, 1996.

[199] Susanne Winter, Bernhard Brendel, Ioannis Pechlivanis, Kirsten Schmieder, and Christian Igel. Registration of ct and intraoperative 3-d ultrasound images of the spine using evolutionary and gradient-based methods. IEEE Transactions on Evolutionary Computation, 12(3):284–296, 2008.

[200] Han-Ming Wu and Henry Horng-Shing Lu. Iterative sliced inverse regression for segmentation of ultrasound and mr images. Pattern Recognition, 40(12):3492–3502, 2007.

[201] Guofang Xiao, Michael Brady, J Alison Noble, and Yongyue Zhang. Segmentation of ultrasound b-mode images with intensity inhomogeneity correction. IEEE Transactions on medical imaging, 21(1):48–57, 2002.

[202] Yiming Xiao, Maryse Fortin, Geirmund Unsgård, Hassan Rivaz, and Ingerid Reinertsen. Re trospective evaluation of cerebral tumors (resect): A clinical database of pre-operative mri and intra-operative ultrasound in low-grade glioma surgeries. Medical physics, 44(7):3875–3882, 2017.

[203] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1492–1500, 2017.

[204] Lu Xu, Mingyuan Liu, Zhenrong Shen, Hua Wang, Xiaowei Liu, Xin Wang, Siyu Wang, Tiefeng Li, Shaomei Yu, Min Hou, et al. Dw-net: A cascaded convolutional neural net-

work for apical four-chamber view segmentation in fetal echocardiography. Computerized Medical Imaging and Graphics, 80:101690, 2020.

[205] Xiaoyan Xu and Robert D Dony. Differential evolution with powell's direction set method in medical image registration. In 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821), pages 732–735. IEEE, 2004.

[206] Xuanang Xu, Thomas Sanford, Baris Turkbey, Sheng Xu, Bradford J Wood, and Pingkun Yan. Shadow-consistent semi-supervised learning for prostate ultrasound segmentation. IEEE Transactions on Medical Imaging, 41(6):1331–1345, 2021.

[207] Yuan Xu, Yuxin Wang, Jie Yuan, Qian Cheng, Xueding Wang, and Paul L Carson. Medical breast ultrasound image segmentation by machine learning. Ultrasonics, 91:1–9, 2019.

[208] Ji Yang, Mehdi Faraji, and Anup Basu. Robust segmentation of arterial walls in intravascular ultrasound images using dual path u-net. Ultrasonics, 96:24–33, 2019.

[209] Parham Yashar and L Nelson Hopkins. Instrumentation in neurosurgery: nurturing the trend toward minimalism. World neurosurgery, 80(3-4):240–242, 2013.

[210] Shi Yin, Qinmu Peng, Hongming Li, Zhengqiang Zhang, Xinge You, Katherine Fischer, Susan L Furth, Yong Fan, and Gregory E Tasian. Multi-instance deep learning of ultrasound imaging data for pattern classification of congenital abnormalities of the kidney and urinary tract in children. Urology, 2020.

[211] Shi Yin, Qinmu Peng, Hongming Li, Zhengqiang Zhang, Xinge You, Katherine Fischer, Susan L Furth, Gregory E Tasian, and Yong Fan. Automatic kidney segmentation in ultrasound images using subsequent boundary distance regression and pixelwise classification networks. Medical image analysis, 60:101602, 2020.

[212] Jinhua Yu, Yuanyuan Wang, Ping Chen, and Yuzhong Shen. Fetal abdominal contour extraction and measurement in ultrasound images. Ultrasound in medicine & biology, 34(2):169–182, 2008.

[213] Vladimir Zagrodsky, Raj Shekhar, and J Fredrick Cornhill. Multifunction extension of simplex optimization method for mutual information-based registration of ultrasound

volumes. In Medical Imaging 2001: Image Processing, volume 4322, pages 508–515. SPIE, 2001.

[214] Yiqiang Zhan and Dinggang Shen. Deformable segmentation of 3-d ultrasound prostate images using statistical texture matching method. IEEE Transactions on Medical Imaging, 25(3):256–272, 2006.

[215] Jianhua Zhang, Lei Chen, Xiaoyan Wang, Zhongzhao Teng, Adam J Brown, Jonathan H Gillard, Qiu Guan, and Shengyong Chen. Compounding local invariant features and global deformable geometry for medical image registration. Plos one, 9(8):e105815, 2014.

[216] Di Zhou, Jun Sun, Choi-Hong Lai, Wenbo Xu, and Xiaoguang Lee. An improved quantum-behaved particle swarm optimization and its application to medical image registration. International Journal of Computer Mathematics, 88(6):1208–1223, 2011.

[217] Hang Zhou and Hassan Rivaz. Registration of pre-and postresection ultrasound volumes with noncorresponding regions in neurosurgery. IEEE journal of biomedical and health informatics, 20(5):1240–1249, 2016.

[218] Shujun Zhou, Hong Wu, Jie Gong, Ting Le, Hao Wu, Qin Chen, and Zenglin Xu. Mark-guided segmentation of ultrasonic thyroid nodules using deep learning. In Proceedings of the 2nd International Symposium on Image Computing and Digital Medicine, pages 21–26, 2018.

[219] Veronika A Zimmer, Miguel Ángel González Ballester, and Gemma Piella. Multimodal image registration using laplacian commutators. Information Fusion, 49:130–145, 2019.

[220] Orly Zvitia, Arnaldo Mayer, Ran Shadmi, Shmuel Miron, and Hayit K Greenspan. Co-registration of white matter tractographies by adaptive-mean-shift and gaussian mixture modeling. IEEE Transactions on Medical Imaging, 29(1):132–145, 2009.