Speckle Patch Similarity for Echogenicity based Multi-Organ Segmentation in Ultrasound Images of the Thyroid Gland

Nikhil S. Narayan, Pina Marziliano, Jeevendra Kanagalingam, MD and Christopher G.L. Hobbs, MD

Abstract—**Ultrasound (US) imaging deals with forming a brightness image from the amplified back-scatter echo when an ultrasound wave is triggered at the region of interest. Imaging artefacts and speckles occur in the image as a consequence of back-scattering and subsequent amplification. We demonstrate the usefulness of speckle related pixels and imaging artefacts as sources of information to perform multi-organ segmentation in US images of the thyroid gland. The speckle related pixels are clustered based on a similarity constraint to quantize the image. The quantization results are used to locate useful anatomical landmarks that aid the detection of multiple organs in the image which are the thyroid gland, the carotid artery, the muscles and the trachea. The spatial locations of the carotid artery and the trachea are used to estimate the boundaries of the thyroid gland in transverse US scans. Experiments performed on a multi-vendor dataset yield good quality segmentation results with Probabilistic Rand Index (PRI) > 0.83 and Boundary Error (BE) < 1mm, and an average accuracy greater than 94%. Analysis of the results using the Dice co-efficient as the metric shows that the proposed method performs better than the state-of-the-art methods. Further, experiments conducted on 971 images of a publicly available dataset prove the effectiveness of the algorithm to track the carotid artery for guided interventions. In addition to US guided interventions, the algorithm can be used as a general framework in applications pertaining to volumetric analysis and computer aided diagnosis.**

Index Terms—**Echogenicity, ultrasound image segmentation, thyroid gland, multi-organ segmentation, local phase.**

I. INTRODUCTION

ULTRASOUND (US) image segmentation has been a hot topic of research for over two decades. Attempts at segmenting US images can be dated back to the 1980's when US imaging started to gain importance as a diagnostic device in a clinical setting [1]. A survey by Noble and Boukerroui [2] conducted in 2005 gives an insight into the amount of research that has gone into developing US segmentation algorithms. The field has not run out of steam even after ten years since the survey was conducted. This can be attributed to the ever

N. S. Narayan is a PhD student at the Department of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, e-mail: (nikhilna001@e.ntu.edu.sg).

P. Marziliano is with the Department of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, e-mail: (epina@ntu.edu.sg).

J. Kanagalingam is with Lee Kong Chian School of Medicine, Nanyang Technological University and is a visiting consultant at Tan Tock Seng Hospital, Singapore, e-mail: (Jeevendra_K@ntu.edu.sg).

C. G. L. Hobbs is with the Nobel ENT Head Neck & Thyroid Surgery Centre, Gleneagles Medical Centre and is a visiting consultant at Tan Tock Seng Hospital, Singapore, e-mail: (drchrishobbs@hotmail.com).

changing US imaging technology. The advances in transducer technology, beam forming technology etc., have led to the introduction of 2D+T, 3D and 3D+T US image acquisition systems along with improvements made to the baseline 2D US imaging system [3]–[5]. Such progress in the imaging systems has created a need to upgrade the segmentation algorithms in order to keep up with the technological growth.

This has to be done in a way to provide the best possible information for a clinician to make a diagnosis. The desirable features (information inferred from the output of a segmentation algorithm) that help a clinician are described in detail in [6] and have been concisely stated by Quistgaard [3] as: (a) Like tissues look alike and unlike tissues look different; (b) Cystic tissues look cystic and solid structures look solid; (c) Low attenuators show acoustic enhancement and high attenuators show shadowing; and (d) Boundaries of organs and structures are visible. In the words of Quistgaard [3], "one might decide that it would be a good idea to process an image to remove acoustic shadowing and enhancement behind cystic structures, though it is extremely valuable information to the clinician. A speckle-reduction algorithm might obscure subtle textural differences in tissue that may indicate pathology. Spatially variant contrast-enhancement procedures may make identical tissues look different". The take away message here is that a segmentation method should not alter the characteristics of a tissue in the image under the pretext of providing valuable information to the clinician. Hence, a segmentation algorithm is considered clinically relevant if it outputs the desirable features listed above.

Research in medical imaging in general is headed in the direction of developing techniques to perform multi-organ segmentation [7]. The goal of this research work is to perform multi-organ segmentation in images of the thyroid gland obtained from a freehand 2D US imaging system. The segmentation method proposed in this research work adheres to the standards set by [3] and [6] in order to contribute to the ongoing research on developing clinically relevant US image segmentation algorithms. Here, the tissues that look alike are grouped together based on tissue echogenicity and the acoustic enhancements produced by low attenuators are used as landmarks to segment the different tissues so that the boundaries between different organs are clearly outlined. The input image is not pre-processed for contrast-enhancement or speckle-reduction as is the case in a majority of segmentation algorithms developed for US images.

The proposed algorithm automatically segments the trans-

verse and longitudinal US scans (Fig. 1) into the thyroid gland and muscles. In addition to this, the transverse scans are also segmented into the carotid artery and trachea. The highlight of this research work is the usage of speckle related pixels to quantize the input US image. The quantization levels correspond to the regions of similar reflectivity which are identified based on tissue echogenicity. The motivation to use speckles for tissue characterization and a brief literature review are presented in Section II. The multi-organ segmentation algorithm is presented in Section III. Experimental results, Discussion and Concluding Remarks are presented in Sections IV, V and VI, respectively.

II. LITERATURE REVIEW

Segmentation algorithms for US images of the thyroid gland are needed to: (a) Automatically detect nodules in the gland; (b) Automatically estimate the volume of the gland; and (c) Perform guided interventions. Keramidas et al. [8] and Maroulis et al. [9]–[11] propose methods to segment nodules within the thyroid gland for computer aided diagnosis. Chang et al. [12] and Garg et al. [13] treat the segmentation of the thyroid gland as a binary classification problem and use neural networks to classify the pixels of the image as belonging to the non-thyroid or thyroid gland. Kollorz et al. [14] propose the use of geodesic active contours to segment the thyroid gland. Authors in [12] and [14] use the segmentation results to estimate the volume of the thyroid gland in 2D US images and 3D US volumes, respectively.

Segmentation methods in the field of guided interventions mainly focus on segmenting and tracking the common carotid artery in 2D and 2D+T US images. To date, there have been three attempts to segment the carotid artery in transverse US scans of the thyroid gland. They are: (a) The Star Kalman algorithm [15]; (b) The Star Extended Kalman algorithm [16]; and (c) The Spokes Ellipse algorithm [17]. These are variants of the Star algorithm that was proposed by Friedland et al. [18] to segment and track the ventricular cavity boundaries in 2D+T US images. The methods are semi-automatic and model the carotid artery as an ellipse whose parameters are estimated along pre-defined radial lines emanating from a seed point manually provided by expert sonographers.

A. Multi-Organ Segmentation

A majority of multi-organ segmentation algorithms proposed in literature are learning based and employ classifiers based on decision trees that are trained using Haar like features to segment the images [19]–[22]. Multi-organ segmentation algorithms have also been developed using atlas based registration and learning of shape contexts to segment CT volumes [23], [24]. Extension of shape based methods to US imaging requires a fixed landmark or a co-ordinate system to register the images first and then perform segmentation by classification. This is done by making use of markers attached to the probe and an external tracking device to map the movement of the probe in 3D. There is a chance that the shape-based supervised segmentation algorithms may perform sub-optimally even after registering the images using
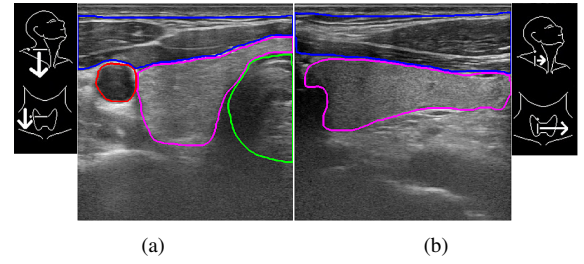

Fig. 1. Ultrasound image of the right lobe of the thyroid gland in (a) transverse and (b) longitudinal scans, depicting the boundaries of the thyroid (in purple), the carotid artery (in red), the muscles (in blue) and the trachea (in green) that have been manually delineated by a trained sonographer.
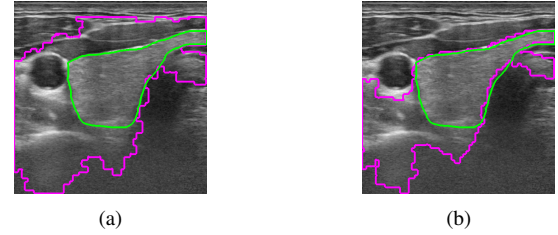

Fig. 2. Results of applying the segmentation method of [12] on the input US image of Fig. 1(a) when the negative class includes samples from : (a) the carotid only ; and (b) both muscles and the carotid. In both cases, the positive training set comprises of samples from the thyroid gland.

the information provided by the external markers. This is because organs in a US scan undergo deformation under probe-pressure. Deformation of organs under probe pressure might result in the shape of organs in the training set to be different from the shape of organs in the testing set. Shape information for supervised segmentation can be used if the deformation can be successfully modelled and corrected. But modelling the deformation is a daunting task as it is subject to inter-observer variation. Different sonographers may apply different probe pressures and the degree of organ deformation might vary among the acquired images. Hence, a shape-based supervised segmentation algorithm trained using the dataset of one sonographer may not perform well on the dataset of another sonographer, thereby, rendering the algorithm to be of little use in a clinical setting.

The performance of a supervised segmentation algorithm also depends on the modelling bias. This is illustrated in Fig. 2, where the supervised thyroid segmentation method of [12] produces two different segmentation results when two different sets of training samples are used for the negative class. Moreover, supervised segmentation algorithms require large databases of manually annotated datasets to train the classifiers. Acquiring such a large database of datasets and having each image manually labelled for a number of different anatomical structures is expensive and time consuming.

Unsupervised segmentation algorithms, on the other hand, do not suffer from modelling bias and can be used for exploratory data analysis. Unsupervised algorithms have previously been employed to perform tissue characterization and binary segmentation in CT and US images [25]–[27]. Linguraru et al., [28] make use of graph cuts to perform
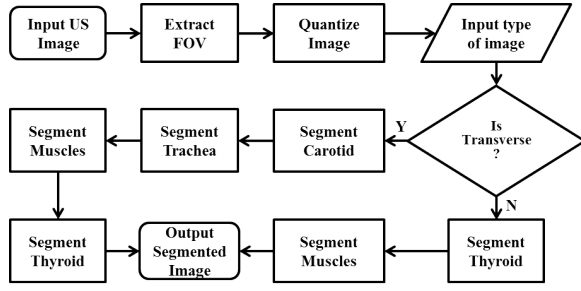
Fig. 3. Segmentation framework

Fig. 4. Flow diagram of the proposed quantization algorithm.

multi-organ segmentation in CT images of the abdomen. Other than the initial feasibility study undertaken in [29], to the best of the authors' knowledge, an unsupervised multi-organ segmentation scheme has not been proposed so far for US images.

B. Speckle in US images

Speckles in US images have always been treated as noise and in a majority of segmentation algorithms, the first step is to filter the image using variants of edge preserving filters [14], [30]–[32] to remove the speckle noise. Attempts have been made to model the noise statistics [33]–[35] and filter the RF signal for speckle noise removal at the envelope detection stage of US image formation. Methods have also been proposed to estimate the amplitude of the RF signal from the gray-scale image for speckle detection and filtering when the log compression parameters are unknown [36], [37]. But a closer look at the physics behind the formation of speckles in US images reveals that speckles are formed as a result of US waves being subjected to constructive or destructive interference after being scattered by cells in a tissue that are of size comparable to the wavelength of US waves [38]. Therefore, treating speckles as noise would be wrong since they are formed in the image due to the interaction of US waves with the tissues.

Thijssen et al. [39] analyze speckles in US images by taking into consideration the imaging physics involved in their formation. An in depth explanation is provided about the different statistical models (Rayleigh, K or Rician distribution) of speckle noise. The authors have also discussed the rationale behind the use of 1st and 2nd order statistics (Signal to Noise Ratio (SNR), auto-correlation, co-occurrence matrices etc.) to analyze the texture of US images. Statistical analysis of speckles and processing of it is valid only when a linear processing of the detected US envelopes is assumed. But in reality, all US images that are displayed and stored would have undergone some sort of logarithmic compression, the parameters of which are proprietary to the manufacturer [3]. The authors in [39] conclude with a note that the future of US imaging rests on clinical applications developed by making use of speckle related tissue characterization. In the spirit of the work by Thijssen et al. [39], speckle related pixels are used in this research work to characterize tissue based on echogenicity and perform multi-organ segmentation of US images of thyroid gland in an unsupervised manner.
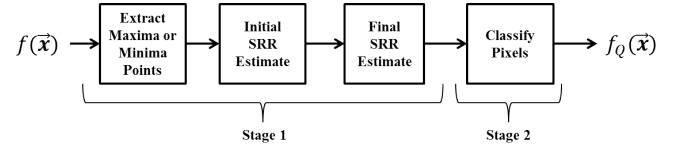
III. MULTI-ORGAN SEGMENTATION OF ULTRASOUND IMAGES OF THYROID GLAND

The proposed segmentation framework is depicted in the flow diagram of Fig. 3. The protocol for scanning a thyroid gland involves scanning the gland in two orthogonal planes called the transverse scan and longitudinal scans, respectively as shown in Fig. 1 [40]. The nature of anatomical structures observed under each of these scans is different and hence the landmarks chosen to perform the segmentation vary with the type of scan making it necessary for two algorithms for the two types of scans. Common to both the segmentation algorithms are the Field of View (FOV) extraction and quantization steps. The FOV, which is the rectangular or conical imaging window in the input US image, is extracted using the scheme proposed in [41]. Once the FOV is extracted, the quantization step follows as described in Section III-A.

A. Echogenicity Based Quantization

The term echogenicity refers to the relative brightness between different tissues in the US image [40]. Hyperechoic tissues appear brighter than hypoechoic tissues and iso-echoic tissues have the same brightness in the US image. A normal thyroid gland is hyperechoic compared to the muscles (in transverse and longitudinal scans) and hypoechoic compared to the enhancement artefact cast by the carotid artery. The carotid artery, jugular vein and the trachea are anechoic structures (transverse scans only) [42]. The omohyoid and strap muscles are iso-echoic to each other, the carotid artery and the trachea are isoechoic to each other (excluding artefacts in trachea). Hence, the US image of the thyroid gland can be viewed as a pattern of two or more isoechoic regions. In the scope of this research work, quantization is formulated as the grouping of the pixels in the US image into the different isoechoic regions. Since the isoechoic tissues reflect the incident US waves to the same extent, we term each region in the US image containing isoechoic tissue as a Similar Reflective Region (SRR's).

Each SRR is a quantization level and the number of SRR's in the input image depends on the imaging plane that is under consideration. The flow diagram of Fig. 4 shows the steps followed to quantize the image. The number of SRR's is the number of clusters formed by the hierarchical (agglomerative) clustering of patches of the image around speckle related pixels using the speckle patch similarity constraint. It is assumed here that a small patch of the image surrounding a speckle is similar to another patch around another speckle in the same tissue or any other iso-echoic tissue.

1) Estimating the number of SRR's: Let the input ultrasound image of size $M \times N$ be denoted by $f(\mathbf{x})$, where $\mathbf{x}$ is the vector of the 2D pixel coordinates (x, y) and $x \in [1, M]$

and $y \in [1, N]$. Let all the pixels in the image be represented by a set $\mathcal{F} = \{\mathbf{x}_k\}$, where $k \in [1, |\mathcal{F}|]$. Assuming that the formation of a speckle at a point in the US image corresponds to the occurrence of local maxima or minima at that point, the set of pixels $\mathcal{F}_s$ associated with speckle in the image is obtained by applying the second order partial derivative test [43] given by Eq. (1) as:

$$\mathcal{F}_s = \{\mathbf{x} : |H(f)| > 0 \text{ and } \frac{\partial^2 f(\mathbf{x})}{\partial^2 x} > 0\} \tag{1}$$

where,

$$H(f) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial^2 x} & \frac{\partial^2 f(\mathbf{x})}{\partial x \partial y} \\ \frac{\partial^2 f(\mathbf{x})}{\partial y \partial x} & \frac{\partial^2 f(\mathbf{x})}{\partial^2 y} \end{pmatrix} \tag{2}$$

is the Hessian matrix and $|H(f)|$ its determinant.

For every pixel $\mathbf{x}_m$ in $\mathcal{F}_s$, an $M_{R1} \times M_{R1}$ Region of Interest (ROI) is defined with $\mathbf{x}_m$ at its centre. The brightness values $f(\mathbf{x})$ at each of the M_{R1}^2 pixel locations in the ROI are denoted by a $1 \times M_{R1}^2$ vector $\mathbf{r}$. Two pixels $\mathbf{x}_m$ and $\mathbf{x}_n$, belong to the same SRR if their respective ROI's $\mathbf{r}_m$ and $\mathbf{r}_n$ ($m \neq n$ and $m, n \in [1, |\mathcal{F}_s|]$) satisfy the constraint:

$$s_{mn} = S(\mathbf{r}_m, \mathbf{r}_n) \leq \tau, \tag{3}$$

where,

$$S(\mathbf{r}_m, \mathbf{r}_n) = \sum_{k=1}^{M_{R1}^2} (\mathbf{r}_m(k) - \mathbf{r}_n(k))^2 \quad, m \neq n, \tag{4}$$

and τ is a pre-defined threshold. The threshold τ largely decides which SRR the ROI belongs to and is chosen in such a way that it is more sensitive to the regions that are hypo-echoic to the thyroid gland. The larger the value of τ, the higher is the sensitivity of the algorithm to the hypo-echoic regions. A detailed discussion on the selection of τ is provided in Section IV.

An agglomerative clustering scheme is used to obtain an initial estimate on the number of SRR's in the image. A pixel $\mathbf{x}_m$ is chosen at random from $\mathcal{F}_s$ and pairwise similarities with the rest of the pixels in the set are computed using Eq.(4). The pixels that satisfy the constraint of Eq.(3) are grouped under a single class label. The process is repeated, by choosing another unlabelled pixel from $\mathcal{F}_s$, until all the pixels are classified. The presence of shadow artefacts results in pixels of the same tissue being classified into multiple SRR's thereby leading to the overestimation of the number of SRR's in the image. This is corrected by merging the groups by the method outlined in Algorithm 1. The number of unique class labels that remain in the image after the merging process is the final estimate of the number of SRR in the input ultrasound image.

The remaining pixels that are not a part of $\mathcal{F}_s$ are then classified into the newly estimated SRR's using morphological operations. Given an SRR, a binary image is formed by retaining the pixels at that level as foreground pixels and resetting the remaining pixels to background. The holes in the binary image whose spatial locations do not overlap with the pixels belonging to other SRR's are morphologically filled and the pixels are assigned the label of the SRR that is

---

Algorithm 1 Merging pixels to obtain final SRR estimate

1: **initialize** Construct an empty $l \times l$ accumulator matrix $\mathbf{A}$ with each element represented by a_{mn}, $m, n \in [1, l]$.
2: **repeat**
3:     **for** $m = 1$ to l **do**
4:         For every pixel having a label λ_m , define a $M_{R1} \times M_{R1}$ ROI around the pixel.
5:         Determine the class label of the all pixels in the ROI excluding the pixel under consideration.
6:         Increment the value of a_{mn} if there exists at least one pixel with the class label λ_n, where $n \in [1, l]$, in the ROI.
7:         If $a_{mm} < a_{mn}$, re-assign the labels of all the pixels having class label λ_m with λ_n.
8:     **end for**
9: **until** there are no more class re-assignments.

---

under consideration. The process is repeated until all the SRR's are covered. Let L denote the final estimate of the number of clusters and $I_{\text{avg\_}m}$ denote the average intensity of the mth SRR, then intensity of the pixels in the quantized image $f_Q(\mathbf{x})$ takes on values in the range $[I_{\text{avg\_1}}, I_{\text{avg\_L}}]$ with $I_{\text{avg\_1}} < I_{\text{avg\_2}}... < I_{\text{avg\_L}}$.

B. Detecting and Segmenting the Carotid in Transverse Scans

The carotid artery being anechoic, exhibits an enhancement artefact directly underneath it. Regions of enhancement arte-facts directly correspond to the pixels with intensity value $I_{\text{avg\_L}}$ in $f_Q(\mathbf{x})$. So the possible candidates for the carotid artery are the pixels with the lowest $I_{\text{avg\_}m}$ directly above the pixels with values $I_{\text{avg\_L}}$. A binary image $f_{\text{bC}}(\mathbf{x})$ is formed with the pixels above the enhancement artefacts that belong to the carotid candidates as foreground pixels and the rest as background pixels. Since the carotid appears as an ellipse, for every component which is a carotid candidate an ellipse is constructed using linear least squares method. [44]. From the list of possible candidates in the binary image, the carotid artery is detected and segmented by the application of local phase based methods.

The local phase of an image $f(\mathbf{x})$ is calculated from the monogenic signal formed from the output of three filters [45], [46]. The image is first bandpass filtered using quadrature filters to obtain a filtered image $f_e(\mathbf{x})$. This represents the even component of the monogenic signal. The odd component of the monogenic signal, $f_o(\mathbf{x})$ is obtained by convolving $f_e(\mathbf{x})$ with two anti-symmetric filters $o_1(x, y)$ and $o_2(x, y)$:

$$f_o(x, y) = \sqrt{(o_1(\mathbf{x}) \otimes f_e(\mathbf{x}))^2 + (o_2(\mathbf{x}) \otimes f_e(\mathbf{x}))^2} \tag{5}$$

where $\otimes$ refers to the convolution operator,

$$o_1(\mathbf{x}) = \frac{-x}{2\pi(x^2 + y^2)^{\frac{3}{2}}} \text{ and } o_2(\mathbf{x}) = \frac{-y}{2\pi(x^2 + y^2)^{\frac{3}{2}}}. \tag{6}$$

The local phase of the image is then obtained as:

$$\phi(\mathbf{x}) = \tan^{-1}\left(\frac{f_e(\mathbf{x})}{f_o(\mathbf{x})}\right) \tag{7}$$
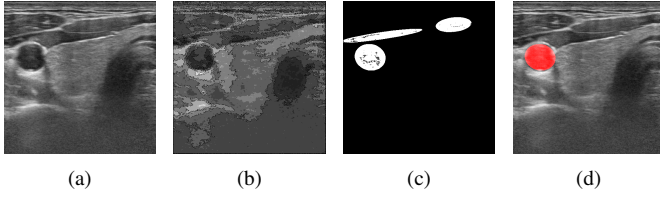
(a)   (b)   (c)   (d)

Fig. 5. Sequence of images showing the segmentation of carotid in transverse US scans where: (a) is the input US image;(b) is the quantized image $f_Q(\mathbf{x})$; (c) is the image of the carotid candidates with pixels that have FC > 0.45 highlighted (in gray) and (d) is the segmented carotid (in red).

The choice of the quadrature bandpass filter is purely application dependant and a detailed comparative account can be found in [47]. Log Gabor filters have previously been used in [45] and [46] to filter the image. In this research work, scale and orientation invariant filter proposed by Mellor and Brady [48] is used to bandpass filter the image. The impulse response of the filter is given by:

$$h(\mathbf{x}) = \frac{A}{(\sqrt{x^2 + y^2})^{\alpha + \beta}} - \frac{B}{(\sqrt{x^2 + y^2})^{\alpha - \beta}}, \qquad (8)$$

where, $\beta \ll \alpha$ with A and B chosen such that the integral of the fractions over x, y is unity.

With the local phase calculated, the phase congruency or feature assymetry is then obtained as:

$$\text{FC}(\mathbf{x}) = \frac{\lfloor [|f_e(\mathbf{x})| - |f_o(\mathbf{x})|] - T_s \rfloor}{\sqrt{f_e(\mathbf{x})^2 + f_o(\mathbf{x})^2 + \epsilon}}, \qquad (9)$$

where $\lfloor . \rfloor$ denotes the flooring function to zero and

$$T_s = e^{\frac{1}{MN} \sum_{\mathbf{x}} \log(\sqrt{f_e(\mathbf{x})^2 + f_o(\mathbf{x})^2})}. \qquad (10)$$

In [46], it was noted that FC values close to 1 indicated the presence of structurally significant boundaries in an ultrasound image. So the candidate with the highest number of pixels having $\text{FC}(\mathbf{x}) > 0.45$ within it is chosen as the most likely candidate for the carotid. The pixels of this candidate are retained as foreground pixels in $f_{\text{bC}}(\mathbf{x})$, with the rest reset to background pixels. The selected candidate is indeed the carotid and segmented in the image $f(\mathbf{x})$ if:

$$\sum_{\mathbf{x}} (f_{\text{bC}}(\mathbf{x}))(1 - \mathbb{H}(\phi(\mathbf{x}))) > 0. \qquad (11)$$

and

$$\text{Ecc} = \sqrt{1 - \frac{b^2}{a^2}} > 0.75, \qquad (12)$$

where, $\mathbb{H}(.)$ represents the Heaviside function and Ecc is the eccentricity of the candidate ellipse with major axis a and minor axis b. The constants of Eq. (9) and Eq. (12) are empirically chosen. Figure 5 shows the process of segmenting the carotid in transverse US scans of the thyroid gland.

C. Segmenting the Trachea, Muscles and the Thyroid in Transverse Scans

The trachea is segmented by applying connected component analysis to the binary image $\phi_b(\mathbf{x})$ given by:

$$\phi_b(\mathbf{x}) = (1 - \mathbb{H}(\phi(\mathbf{x}))). \qquad (13)$$

The largest component of $\phi_b(\mathbf{x})$ attached to the component at the bottom of the image is segmented as the Trachea.

Muscles in an US image of the thyroid gland are found above the carotid and thyroid and below layers of fatty tissue that appear hyperechoic. The muscles are segmented by making use of the carotid as the landmark. Similar to the segmentation of the trachea, the muscles are segmented by applying connected component analysis on the binary image $\phi_b(\mathbf{x}) - f_{\text{bTrc}}(\mathbf{x})$. The largest component above the carotid belongs to the Omohyoid muscles. The components that lie in between the upper and lower bounds of a bounding box around the Omohyoid muscles belong to the strap muscles. The pixels in the components belonging to the Omohyoid and Strap muscles are together labelled as muscles in the segmented image.

The spatial locations of carotid, trachea and muscles are used as landmarks to segment the thyroid. The location of the trachea is used to determine the lobe of the thyroid gland that is in the image. The image is of the right lobe of the thyroid gland if the ordinate of the centroid of the trachea is greater than $\frac{N}{2}$, else it belongs to the left lobe. In order to detect and segment the lower boundary of the thyroid gland, the binary image $\bar{\phi}_b(\mathbf{x})$ given by:

$$\bar{\phi}_b(\mathbf{x}) = \lceil 1 - \phi_b(\mathbf{x}) \rceil = \lceil \mathbb{H}(\phi(\mathbf{x})) \rceil, \qquad (14)$$

where, $\lceil . \rceil$ stands for the ceiling operator, is considered. All the pixels from the first row of the image up to the lower boundary of the muscles are set as background pixels in $\bar{\phi}_b(\mathbf{x})$. The co-ordinate of the highest point of the trachea, $\mathbf{x}_{\text{Tr}}$ and the centroid of the carotid $\mathbf{x}_{\text{Cr}}$ are then determined from the segmented carotid and trachea, respectively. All the edge pixels in $f(\mathbf{x})$ that lie immediately under a digital line running from $\mathbf{x}_{\text{Tr}}$ to $\mathbf{x}_{\text{Cr}}$ form a part of the lower boundary of the thyroid. The digital line is given by [49]:

$$n = \lceil r \times m \rceil + c, \forall m \in [1, M] \text{ and } |r| \le 1, \quad \text{or} \qquad (15)$$

$$m = \lceil \frac{n}{r} \rceil + c, \forall n \in [1, N] \text{ and } |r| > 1, \qquad (16)$$

where,

$$r = \frac{y_{\text{Cr}} - y_{\text{Tr}}}{x_{\text{Cr}} - x_{\text{Tr}}} \qquad (17)$$

and

$$c = y_{\text{Cr}} - r x_{\text{Cr}} \text{ or } c = y_{\text{Tr}} - r x_{\text{Tr}}. \qquad (18)$$

The lower boundary of the thyroid is obtained by fitting a spline to the edge pixels just determined. The thyroid gland comprises of the pixels in $\bar{\phi}_b(\mathbf{x})$ that are above the estimated lower boundary; in between the carotid, lower boundary of the muscles and the upper boundary of the trachea. Figure 6 shows the segmentation of the transverse US scans of thyroid gland along with the estimated anterior boundary of the thyroid gland.

D. Segmenting the Thyroid and Muscles in Longitudinal Scans

In longitudinal scans, the thyroid gland is segmented directly by making use of the composite image $f_Q(\mathbf{x})$ of Section
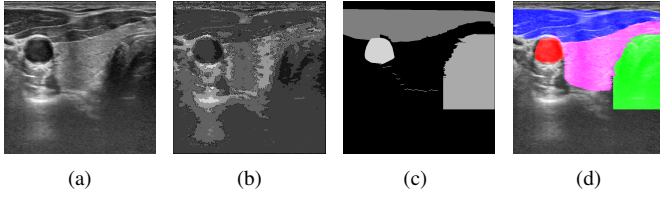
(a) (b) (c) (d)

Fig. 6. Estimating the lower boundary of the thyroid gland. (a) Input image in transverse scan; (b) Quantized image $f_Q(\mathbf{x})$; (c) Estimated lower boundary of thyroid gland and (d) Segmentation results with the thyroid in Magenta, the carotid in Red, muscles in Blue and the trachea in Green.

III-B and the local phase $\phi(\mathbf{x})$ information. All the pixels of the image that are associated with SRR's whose $I_{\mathrm{avg\_}m}$'s lie the range $[I1, I2]$, such that $I1 < I_{\mathrm{avg\_}m} < I2$, are labelled as foreground pixels with the rest labelled as background pixels and stored as a binary segmentation mask $f_{\mathrm{bMask}}(\mathbf{x})$. Although the bulk of the thyroid is present in $f_{\mathrm{bMask}}(\mathbf{x})$, the segmentation is incomplete as the mask contains leaks into muscles and the lower non-thyroid region of the US image. A local phase based method is employed to segment the thyroid gland without any segmentation leaks.

The local phase image $\phi(\mathbf{x})$ and a binary image $\bar{\bar{\phi}}_b(\mathbf{x})$ are obtained using Eq.(14) and Eq. (7), respectively. Connected component labelling is then applied on the binary image and the pixels of the largest component in $\bar{\bar{\phi}}_b(\mathbf{x})$ are retained as the foreground pixels. The component is that of the thyroid gland if:

$$\sum_{\mathbf{x}} f_{\mathrm{bMask}}(\mathbf{x})\bar{\bar{\phi}}_b(\mathbf{x}) > 0. \tag{19}$$

In order to segment the muscles in the longitudinal scan, the binary segmentation $\phi_b(\mathbf{x})$ mask of Eq. (13) is considered and all the pixels below the upper boundary of the thyroid are set as background pixels. Connected component labelling is then applied to $\phi_b(\mathbf{x})$ to determine the components of the binary image that belong to the muscles. For each row in the labelled image, the histogram of class labels is obtained and the label at which the histogram peaks is stored in an accumulator array. The total number of pixels belonging to each of the labels in the accumulator array is then determined and the array is sorted in descending order. The top two labels in the sorted accumulator array belong to the Omohyoid and Strap muscles. The pixels that belong to these labels are labelled as muscles in the segmented image.

IV. RESULTS

Figure 7 shows the results on transverse and longitudinal scans of the thyroid gland. The dataset and metrics used to evaluate the performance of the segmentation algorithm are described in detail in the subsections that follow.

A. Datasets and Ground Truth

Two sets of anonymized US images in B-mode were retrospectively collected from two hospitals':

- *Set 1 (S1)*: consists of 34 images (20 transverse and 14 longitudinal) from 6 patients and 6 volunteers acquired
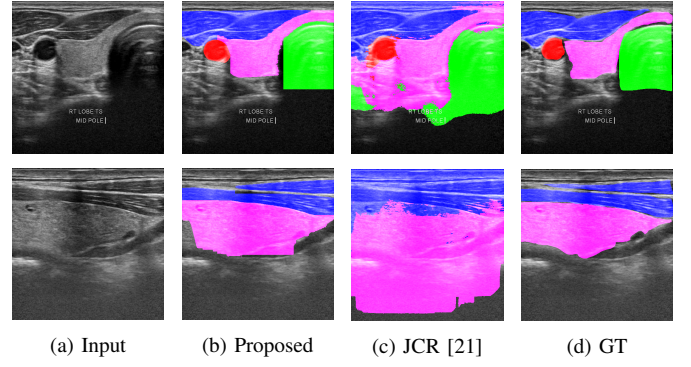

(a) Input (b) Proposed (c) JCR [21] (d) GT

Fig. 7. Segmentation results in transverse (1st row) and longitudinal US scans (2nd row) of thyroid gland where the thyroid, carotid, muscles and trachea are colour coded in Magenta, Red, Blue and Green, respectively, for illustrative purposes. (a) Input US image; (b) Results of the proposed algorithm on the input in (a); (c) Results of applying the supervised multi-organ segmentation algorithm of [21]; and (d) Ground truth segmentation.

TABLE I
PERFORMANCE MEASURE OF THE PROPOSED SEGMENTATION SCHEME IN TERMS OF PRI, GCE, VOI AND BE. A GOOD SEGMENTATION ALGORITHM HAS PRI VALUES CLOSE TO ONE WITH GCE, VOI AND BE VALUES CLOSE TO ZERO, RESPECTIVELY.

Dataset	PRI	GCE	VOI	BE (pixels)
Transverse (Set 1)	0.8611 ±0.0260	0.1770 ±0.0310	1.1357 ±0.1453	11.42 ±2.193
Transverse (Set 2)	0.8345 ±0.0410	0.2300 ±0.0604	1.3592 ±0.2850	11.1729 ±3.121
Longitudinal (Set 1)	0.8352 ±0.0436	0.1692 ±0.0413	1.013 ±0.1870	14.07 ±3.051

TABLE II
QUANTITATIVE ANALYSIS ON THE TRANSVERSE SCANS OF SET 1 *.

Organ	GT	SE	SP	DSC	PPV	Avg Acc
Thyroid	E1	0.982 ±0.007	0.827 ±0.074	0.854 ±0.033	0.889 ±0.042	0.959 ±0.012
	E2	0.974 ±0.010	0.884 ±0.070	0.855 ±0.045	0.835 ±0.069	
Carotid	E1	0.998 ±0.002	0.856 ±0.073	0.879 ±0.035	0.912 ±0.064	0.995 ±0.001
	E2	0.997 ±0.001	0.916 ±0.069	0.895 ±0.037	0.883 ±0.062	
Muscles	E1	0.990 ±0.009	0.778 ±0.129	0.847 ±0.082	0.946 ±0.039	0.955 ±0.019
	E2	0.959 ±0.021	0.934 ±0.070	0.829 ±0.062	0.753 ±0.093	
Trachea	E1	0.980 ±0.017	0.874 ±0.083	0.863 ±0.036	0.864 ±0.072	0.968 ±0.013
	E2	0.974 ±0.017	0.885 ±0.100	0.834 ±0.062	0.804 ±0.108	

* E1 and E2 stand for the GT provided by expert 1 and expert 2, respectively.

at the Tan Tock Seng Hospital, Singapore, using Hitachi HI Vision Avius Ultrasound equipped with an L75 5-18MHz probe. All images have a resolution $M \times N$ equal to 511×510. Informed consent was obtained from the

TABLE III
QUANTITATIVE ANALYSIS ON THE TRANSVERSE SCANS OF SET 2 ⋆.

Organ	GT	SE	SP	DSC	PPV	Avg Acc
Thyroid	E1	0.964 ±0.013	0.829 ±0.074	0.810 ±0.026	0.804 ±0.063	0.945 ±0.008
	E2	0.951 ±0.018	0.917 ±0.056	0.808 ±0.034	0.730 ±0.076	
Carotid	E1	0.999 ±0.001	0.816 ±0.075	0.882 ±0.039	0.968 ±0.039	0.994 ±0.001
	E2	0.998 ±0.001	0.856 ±0.062	0.878 ±0.017	0.909 ±0.049	
Muscles	E1	0.993 ±0.003	0.653 ±0.136	0.767 ±0.104	0.946 ±0.036	0.944 ±0.024
	E2	0.975 ±0.004	0.805 ±0.123	0.805 ±0.088	0.808 ±0.054	
Trachea	E1	0.971 ±0.024	0.878 ±0.081	0.876 ±0.026	0.886 ±0.073	0.946 ±0.029
	E2	0.941 ±0.046	0.901 ±0.072	0.788 ±0.099	0.720 ±0.170	

⋆ E1 and E2 stand for the GT provided by expert 1 and expert 2, respectively.

TABLE IV
QUANTITATIVE ANALYSIS ON THE LONGITUDINAL SCANS OF SET 1 ⋆

		SE	SP	DSC	PPV	Avg Acc
Thyroid	E1	0.942 ±0.035	0.920 ±0.036	0.870 ±0.052	0.830 ±0.084	0.921 ±0.051
	E2	0.919 ±0.048	0.960 ±0.026	0.834 ±0.079	0.749 ±0.125	
Muscles	E1	0.996 ±0.005	0.611 ±0.180	0.732 ±0.159	0.973 ±0.030	0.935 ±0.030
	E2	0.964 ±0.028	0.766 ±0.259	0.691 ±0.236	0.673 ±0.273	

⋆ E1 and E2 stand for the GT provided by expert 1 and expert 2, respectively.

volunteers before acquiring their thyroid US scans.

- *Set 2 (S2)*: consists of 18 images in transverse scan from 10 patients acquired at Mt. Elizabeth Hospital, Singapore, using GE Logiq Ultrasound equipped with L6-12-RC probe. The images have a resolution $M \times N$ equal to 820×614 .
- *Supervised Set (SS)*: This dataset consists of 24 images (15 transverse and 9 longitudinal) to train and 16 images (11 transverse and 4 longitudinal) to test supervised learning-based segmentation algorithms. Of the 15 training images in transverse scan 12 are from Set 1 and 3 are from Set 2. All images (train and test) in the longitudinal scan are from Set 1.

The datasets consist of a mix of images that are of the right or left lobes of the thyroid gland. The images were acquired at the mid or lower pole with the patient in supine position. Images that were affected by abnormalities leading to organ occlusion are excluded from the study. All images were acquired at a centre frequency of 10 MHz. Since the datasets were retrospectively collected, the Time Gain Compensation (TGC) setting used during imaging is unknown.

Ground Truth (GT): Two expert Otorhinolaryngologists who are also trained sonographers at Tan Tock Seng hospital, Singapore, were asked to manually delineate the organs (Muscles, trachea, thyroid and carotid) using the ITK-SNAP software [50]. GT acquisition is an ongoing process and is available for 40 of the 52 images (20 Transverse and 14 Longitudinal in Set 1; 6 Transverse in Set 2). So the performance of the proposed segmentation method is validated on 40 images for which the GT is available.

B. Qualitative and Quantitative Assessment

The proposed segmentation algorithm is qualitatively assessed using Probabilistic Rand Index (PRI) [51], Global Consistency Error (GCE) [52], Variation of Information (VOI) [53] and Boundary Distance Error (BE) [54] metrics. The results of a segmentation algorithm are considered to be of good quality if the PRI $\in [0, 1]$ has a value that is close to one, GCE $\in [0, 1]$ has a value close to zero, VOI and BE $\in [0, \infty)$ have values as low as possible [55]. The performance of the multi-organ segmentation scheme in terms of PRI, GCE, VOI and BE is summarized in Table I. An extension to the PRI is the Normalized Probabilistic Rand (NPR) Index [56]. The NPR index is used to select the optimal parameters needed in the SRR estimation algorithm of Section III-A. The optimal parameters for the SRR estimation are those that have a high NPR index.

The segmentation algorithm is also quantitatively evaluated by using metrics that measure the amount of overlap between the Ground Truth (GT) and the segmentation results. The overlap measures used for validation are: (1) Sensitivity (SE); (2) Specificity (SP); (3) Dice Coefficient (DSC) and (4) True Positive Rate (TPR) or Positive Predictive Value (PPV) and (5) Accuracy (Acc) given by:

$$SE = \frac{TN}{TN + FP}; SP = \frac{TP}{TP + FN}; Acc = \frac{TP + TN}{A_P + A_N} \quad (20)$$

$$PPV = \frac{TP}{TP + FP}; \quad DSC = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (21)$$

where TP, TN, FP, FN, A_P and A_N stand for True Positives, True Negatives, False Positives, False Negatives, All Positives and All Negatives, respectively. The performance of the algorithm in terms of overlap metrics of Eq. (21) is summarized in Tables II, III and IV, respectively.

C. Comparison with state-of-the-art

Segmentation results of the proposed method are compared with our implementation of the following state-of-the-art methods and the results are summarized in Table V.

1) Segmenting the carotid artery: The semi-automatic methods of [15], [16] and [17] mentioned in Section II are implemented for this comparative study. The centroid of the GT carotid segmentation (pixels of carotid that lie in the intersection of the two manually segmented regions) is chosen as the seed point to initialize the algorithms. Values of the parameters used in our implementation to execute the algorithms are the same as those reported in the respective
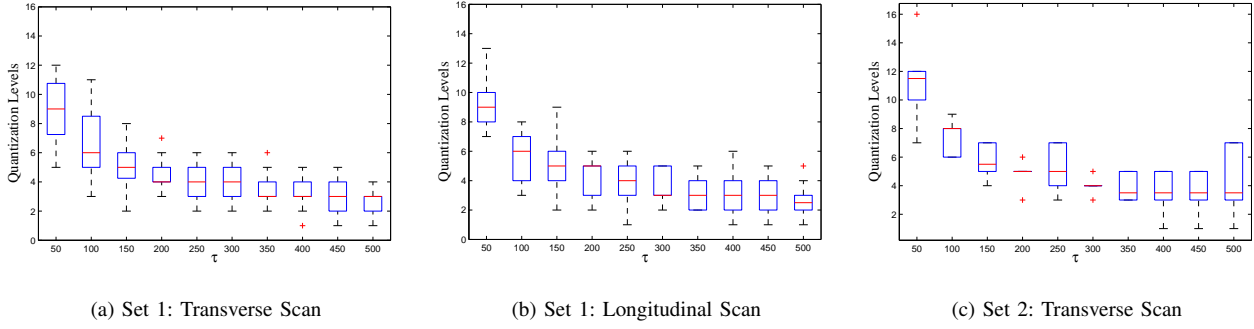
(a) Set 1: Transverse Scan   (b) Set 1: Longitudinal Scan   (c) Set 2: Transverse Scan

Fig. 8. Variation of estimated quantization levels with changes in threshold τ.

papers for optimal performance. In addition to these, the star algorithm of [18] has also been implemented as [15] and [16] are variants of this method.

2) Segmenting the thyroid gland: The thyroid segmentation algorithm of [12] is implemented by first determining the probable thyroid region followed by applying adaptive weighted median filtering and morphological closing of the image. Then, features such as: (a) Co-efficient of Local Variation; (b) Block Difference of Inverse Probabilities; (c) Normalized Multi-Scale Intensity Difference;(d) Haar wavelet based features; and (e) histogram features are extracted in a 16×16 rectangular window. An RBF neural network is trained using the features extracted from the images in the training set of SS. The parameters used at every stage in our implementation are the same as those reported in [12]. Each test image in SS is divided into blocks of size 16×16 pixels, with 50% overlap between the blocks. The blocks are then classified into the thyroid or non-thyroid region by the trained classifier. The resulting image is post-processed by the recommended morphological operations. A similar approach is followed in implementing the method proposed by Garg et al. [13] with the following changes: (a) feature set is changed to histogram based features and multi-scale Haar wavelet based features; and (b) the classifier changed to feed-forward neural networks. It should be noted here that the regions belonging to the thyroid gland form the positive class and regions not belonging to the thyroid gland (carotid, muscles and the trachea) form the negative class. Overall 66 samples (33 positive and 33 negative) are generated to train the neural network classifier.

3) Unsupervised multi-organ segmentation: An initial feasibility study was undertaken in [29] to study the effects of speckle related pixels on multi-organ segmentation. The proposed method is different from that of [29] in the following ways: (a) the number of echogenic levels are not restricted to three; (b) patches of image around speckles are clustered using a similarity constraint instead of clustering individual pixels based on intensity; (c) imaging artefacts are used as landmarks to detect the carotid artery instead of using an energy based approach; and (d) local phase based methods are used in the segmentation process.

4) Supervised multi-organ segmentation: The Joint Classification Regression (JCR) algorithm proposed by Glocker et al. [21] to perform multi-organ segmentation in CT images of the abdomen method is modified for use on US images so that

it can be compared with our method. In our implementation of [21], the adaptive weighted median filter of [12] is used with a window size of 5×5 to filter the image for noise removal. Five variants of the box features are extracted from the filtered image. The box sizes vary between 5mm and 10mm, and displacements of the boxes are drawn from a [0,10mm] interval. In all, 369 features are extracted per pixel in every organ from the training images in SS. Bagging is used to train 50 trees where each tree is trained on a random subset containing 10% of the total number of training samples. At each split node, 40 features are evaluated from the pool of 369 features and for each feature 10 different thresholds are employed that are uniformly distributed along the range of feature responses. The distance from every pixel in one organ to the nearest pixel of another organ obtained using signed distance maps is used for regression.

V. DISCUSSION

The SRR estimation method proposed in Section III-A is an agglomerative clustering algorithm and the number of clusters converges to the trivial solution of one level for a very large τ. This behavior can be observed in the plots of Fig. 8. The number of SRR's obtained remain more or less steady at 5 ± 2 levels for τ between 100 and 400 and converges to one level for $\tau > 10000$. Figure 9 shows the performance of the SRR estimation algorithm in terms of NPR at different τ and M_{R1}. From the plots it can be inferred that the SRR estimation algorithm performs at its best when $\tau = 100$ and $M_{R1} = 3$. Hence, all experiments are conducted by setting τ and M_{R1} equal to 100 and 3, respectively.

Three empirically set constants are used to segment the carotid artery and thyroid gland in transverse and longitudinal US scans, respectively. These are: (a) the phase congruency (FS) for the carotid artery; (b) the eccentricity of the ellipse (Ecc) for the carotid artery; and (c) the intensity range [I1, I2] for the thyroid gland. From the sensitivity analysis plots of Fig. 10, it can be inferred that the algorithm is robust to changes in Ecc and changes in FS up to 0.55. The best performance of the algorithm is seen at low values of FS and high values of Ecc. Accordingly, all experiments for US images in the transverse scan are conducted by setting FS=0.45 and ECC=0.77. The algorithm is robust to changes in the intensity range used in the segmentation of the thyroid gland and the best performance is seen when the intensity is chosen in the range [50,200].
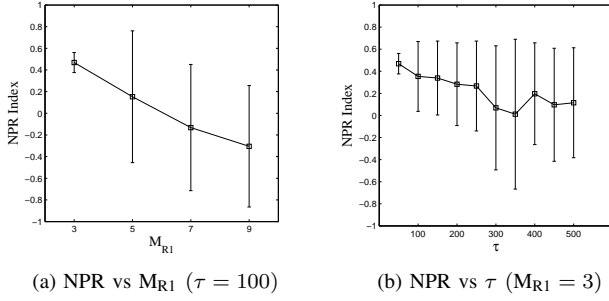
(a) NPR vs M_{R1} ($\tau = 100$) (b) NPR vs τ ($M_{R1} = 3$)

Fig. 9. Performance analysis of the proposed quantization scheme with respect to: (a) changes in the ROI size M_{R1} when the threshold τ is held constant at 100 and (b) changes in the threshold τ when the ROI size M_{R1} is held constant at 3.
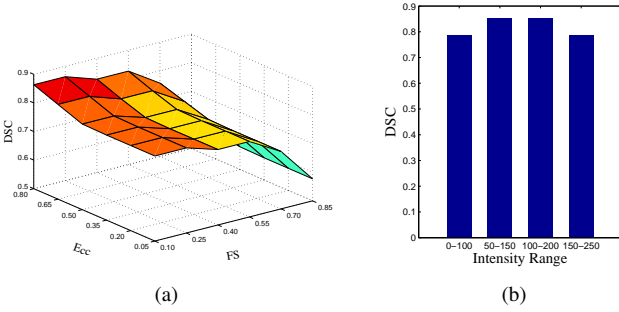
(a) (b)

Fig. 10. Sensitivity analysis for the constants used to segment: (a) the carotid artery; and (b) the thyroid gland in transverse and longitudinal scans, respectively.
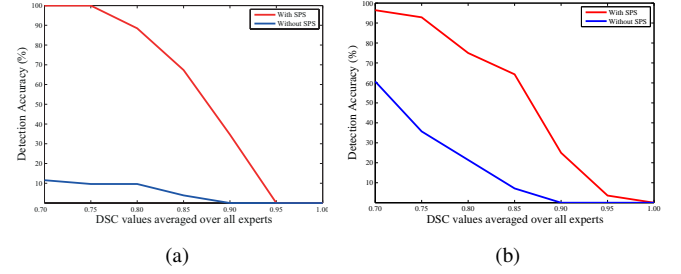
(a) (b)

Fig. 11. Performance analysis of (a) the carotid artery and (b) the thyroid gland detectors with (in red) and without (in blue) clustering based on speckle patch similarity (SPS).

From table I it can be seen that: the average PRI values are well above 0.8; average GCE values are close to zero with values under 0.25; average VOI is not more than 1.5 and BE is at around 12 pixels (under 1mm). This indicates that the proposed method is capable of producing good quality segmentation results with millimetre accuracy.

The average sensitivity for all the anatomical structures, in both transverse and longitudinal scans is well above 95% indicating that the possibility of false positive segmentation is less. The average specificity for all the anatomical structures except the muscles is well above 83% while that for muscles is about 73% in the case of transverse scans of Set 2 and 69% in longitudinal scans. The high sensitivity and low specificity values for the muscles suggest a possible under-segmentation of the muscles. The DSC values for all the organs, except for the muscles in longitudinal scans, are above 82%. The low value of DSC for the muscles in longitudinal scans is due to the under-segmentation of the organ by the algorithm. The average segmentation accuracy is above 92% for all the organs.

From the results of Table V, it can be observed that the proposed method performs better than the existing methods used to segment both individual and multiple organs. The thyroid segmentation method of Chang et al. [12] is found to under-segment the gland while that of Garg et al. [13] is found to over-segment the thyroid gland as evidenced by the respective sensitivity and specificity values in Table V. With respect to segmenting the carotid artery, the high sensitivity

and low specificity values for the star algorithm and the star kalman algorithm indicate the under-segmentation of the carotid artery by both methods. The star extended kalman and the spoke-ellipse algorithms on the other hand perform better with high sensitivity and specificity values. Of all the methods proposed to segment the carotid artery, the star extended kalman algorithm of [57] is the only algorithm that has a DSC above 80% but is still less than what is achieved by our method. The multi-organ segmentation method of [21] is found to over-segment the organs.

One of the key factors for the success of the proposed method is the usage of speckle related pixels and its subsequent clustering based on a similarity constraint. The plots in Fig. 11 show the difference in organ detection accuracies with and without the use of speckle related pixels. Since the carotid is used as the landmark in transverse scans and the thyroid is used as the landmark in longitudinal scans to detect other organs, it suffices to show the detection accuracies of the carotid and thyroid gland for the transverse and longitudinal scans, respectively. Without the use of speckle patch similarity, the carotid is successfully detected in 3 out of the 26 images in the transverse scan and the thyroid is successfully detected in 5 out of 14 images in longitudinal scans. Whereas, by using the speckle patch similarity, the carotid is successfully detected in all the images in transverse scans and the thyroid is successfully detected in 92% of the images in the longitudinal scans. Here, an organ is said to be detected successfully when the DSC values are greater than 0.70, hence the DSC values start from 0.70 in the plots of Fig. 11.

Overlap statistics of Section IV-B are used to determine the inter-observer variation and the results are summarized in Table VI. The low SP values of E12 when compared to E21 for all the organs suggest the possibility of expert 1 under-segmenting the organs when compared to expert 2. Comparing the DSC values in Tables II, III, IV and V to the inter-observer variation in Table VI, it can be seen that DSC values obtained by the proposed method are close to the inter-observer variation. Statistically speaking, there is no significant difference between the standardized DSC values:

$$\text{standardized DSC} = ln\frac{\text{DSC}}{1 - \text{DSC}}, \quad (22)$$

of E1 and E2 in Tables II, III and IV as determined by one-way ANOVA for the Thyroid ($F(1,78) = 3.963$, $p = 0.413$),

TABLE V
COMPARISON WITH EXISTING METHODS [*] [†].
S = SUPERVISED. U = UNSUPERVISED. T = TRANSVERSE. L = LONGITUDINAL. S1 = SET 1. S2 = SET 2. SS = SUPERVISED SET

Organ	Method	Automation	Supervision	Dataset	Scan	SE	SP	DSC	PPV
Thyroid	*Proposed*	Full	U	S1	T	0.977 ± 0.010	0.855 ± 0.078	0.854 ± 0.040	0.862 ± 0.063
				S1	L	0.930 ± 0.044	0.940 ± 0.034	0.853 ± 0.069	0.790 ± 0.114
				S2	T	0.958 ± 0.018	0.873 ± 0.083	0.810 ± 0.031	0.767 ± 0.081
				SS	T	0.970 ± 0.013	0.864 ± 0.083	0.840 ± 0.040	0.827 ± 0.069
				SS	L	0.941 ± 0.043	0.936 ± 0.035	0.866 ± 0.072	0.818 ± 0.116
	Narayan [29]	Full	U	S1	T	0.977 ± 0.017	0.819 ± 0.090	0.822 ± 0.048	0.842 ± 0.101
				S1	L	0.933 ± 0.075	0.857 ± 0.107	0.811 ± 0.096	0.805 ± 0.162
				S2	T	0.863 ± 0.122	0.636 ± 0.303	0.487 ± 0.270	0.460 ± 0.339
				SS	T	0.955 ± 0.056	0.740 ± 0.270	0.702 ± 0.246	0.705 ± 0.280
				SS	L	0.966 ± 0.015	0.844 ± 0.095	0.850 ± 0.051	0.867 ± 0.060
	JCR [21]	Full	S	SS	T	0.679 ± 0.102	0.922 ± 0.077	0.490 ± 0.067	0.337 ± 0.060
				SS	L	0.449 ± 0.038	0.930 ± 0.062	0.468 ± 0.064	0.315 ± 0.053
	Chang [12]	Full	S	SS	T	0.891 ± 0.106	0.662 ± 0.258	0.589 ± 0.199	0.598 ± 0.263
				SS	L	0.856 ± 0.127	0.457 ± 0.389	0.435 ± 0.376	0.464 ± 0.431
	Garg [13]	Full	S	SS	T	0.524 ± 0.157	0.955 ± 0.077	0.411 ± 0.074	0.264 ± 0.060
				SS	L	0.421 ± 0.206	0.772 ± 0.370	0.388 ± 0.212	0.266 ± 0.161
Carotid	*Proposed*	Full	U	S1	T	0.998 ± 0.002	0.886 ± 0.077	0.887 ± 0.037	0.897 ± 0.065
				S2	T	0.999 ± 0.001	0.836 ± 0.072	0.881 ± 0.031	0.939 ± 0.054
				SS	T	0.997 ± 0.002	0.908 ± 0.056	0.892 ± 0.031	0.882 ± 0.063
	Star [18]	Semi	U	S1	T	0.935 ± 0.089	0.693 ± 0.338	0.404 ± 0.239	0.531 ± 0.361
				S2	T	0.875 ± 0.104	0.813 ± 0.350	0.303 ± 0.236	0.371 ± 0.334
	Star Kalman [15]	Semi	U	S1	T	0.999 ± 0.001	0.576 ± 0.212	0.700 ± 0.154	0.982 ± 0.044
				S2	T	0.999 ± 0.001	0.566 ± 0.197	0.702 ± 0.166	0.998 ± 0.003
	Star Extended Kalman [16]	Semi	U	S1	T	0.987 ± 0.032	0.932 ± 0.099	0.827 ± 0.167	0.793 ± 0.177
				S2	T	0.982 ± 0.018	0.930 ± 0.143	0.757 ± 0.125	0.692 ± 0.212
	Spoke Ellipse [17]	Semi	U	S1	T	0.982 ± 0.030	0.935 ± 0.090	0.742 ± 0.206	0.667 ± 0.244
				S2	T	0.990 ± 0.008	0.962 ± 0.036	0.839 ± 0.073	0.755 ± 0.120
	Narayan [29]	Full	U	S1	T	0.995 ± 0.005	0.845 ± 0.183	0.809 ± 0.151	0.805 ± 0.181
				S2	T	0.986 ± 0.011	0.783 ± 0.355	0.680 ± 0.311	0.620 ± 0.307
				SS	T	0.992 ± 0.008	0.841 ± 0.286	0.751 ± 0.261	0.695 ± 0.266
	JCR [21]	Full	S	SS	T	0.965 ± 0.040	0.748 ± 0.204	0.503 ± 0.191	0.417 ± 0.219
Muscles	*Proposed*	Full	U	S1	T	0.975 ± 0.023	0.856 ± 0.130	0.838 ± 0.074	0.850 ± 0.120
				S1	L	0.980 ± 0.026	0.689 ± 0.237	0.712 ± 0.203	0.823 ± 0.246
				S2	T	0.984 ± 0.010	0.729 ± 0.151	0.786 ± 0.100	0.878 ± 0.083
				SS	T	0.978 ± 0.021	0.806 ± 0.135	0.817 ± 0.075	0.859 ± 0.118
				SS	L	0.993 ± 0.008	0.819 ± 0.103	0.878 ± 0.049	0.960 ± 0.040
	Narayan [29]	Full	U	S1	T	0.980 ± 0.017	0.737 ± 0.164	0.780 ± 0.116	0.862 ± 0.116
				S1	L	0.959 ± 0.038	0.859 ± 0.120	0.788 ± 0.106	0.780 ± 0.201
				S2	T	0.983 ± 0.016	0.601 ± 0.254	0.676 ± 0.195	0.855 ± 0.117
				SS	T	0.977 ± 0.017	0.676 ± 0.202	0.732 ± 0.148	0.841 ± 0.114
				SS	L	0.978 ± 0.025	0.839 ± 0.114	0.853 ± 0.051	0.894 ± 0.106
	JCR [21]	Full	S	SS	T	0.917 ± 0.035	0.624 ± 0.258	0.556 ± 0.160	0.551 ± 0.145
				SS	L	0.885 ± 0.043	0.940 ± 0.031	0.745 ± 0.108	0.632 ± 0.148
Trachea	*Proposed*	Full	U	S1	T	0.977 ± 0.018	0.880 ± 0.092	0.850 ± 0.053	0.835 ± 0.096
				S2	T	0.956 ± 0.040	0.900 ± 0.078	0.832 ± 0.085	0.804 ± 0.156
				SS	T	0.972 ± 0.026	0.907 ± 0.080	0.856 ± 0.057	0.826 ± 0.110
	Narayan [29]	Full	U	S1	T	0.963 ± 0.031	0.876 ± 0.122	0.780 ± 0.110	0.744 ± 0.177
				S2	T	0.977 ± 0.031	0.291 ± 0.376	0.249 ± 0.298	0.555 ± 0.294
				SS	T	0.955 ± 0.034	0.767 ± 0.350	0.607 ± 0.296	0.607 ± 0.232
	JCR [21]	Full	S	SS	T	0.862 ± 0.036	0.804 ± 0.146	0.526 ± 0.131	0.407 ± 0.132

[*] Results are averaged over both experts.
[†] Methods whose names are italicized represent multi-organ segmentation algorithms. Unsupervised algorithms are listed first followed by supervised algorithms for every organ. The results of each algorithm are sorted according to the dataset used from S1 to SS.

TABLE VI
INTEROBSERVER VARIATION *.

Organ	GT	SE	SP	DSC	PPV
Thyroid	E12	0.993 ±0.006	0.832 ±0.088	0.890 ±0.050	0.961 ±0.029
	E21	0.965 ±0.024	0.961 ±0.029	0.889 ±0.051	0.833 ±0.088
Carotid	E12	0.999 ±0.001	0.855 ±0.072	0.901 ±0.030	0.959 ±0.042
	E21	0.996 ±0.002	0.959 ±0.042	0.901 ±0.049	0.856 ±0.072
Muscles	E12	0.997 ±0.002	0.620 ±0.146	0.749 ±0.117	0.978 ±0.021
	E21	0.926 ±0.026	0.978 ±0.021	0.749 ±0.117	0.620 ±0.146
Trachea	E12	0.990 ±0.009	0.786 ±0.181	0.833 ±0.173	0.895 ±0.186
	E21	0.974 ±0.019	0.895 ±0.186	0.833 ±0.173	0.786 ±0.181

* E12: Manual segmentation by expert 1 used as GT to quantify the segmentation by expert 2; E21: Manual segmentation by expert 2 used as GT to quantify the segmentation by expert 1.

TABLE VII
PERFORMANCE ANALYSIS OF THE CAROTID
SEGMENTATION ALGORITHM ON 971 IMAGES OF
A PUBLIC DATABASE [58].

	Proposed method	Riha et al. [58] [a]
SE (%)	98.14	97

[a] Values are as reported in [58] for the same database.

Carotid (F(1,50) = 4.034, $p = 0.289$), Muscles (F(1,78) = 3.965, $p = 0.964$) and Trachea ((F(1,50) = 4.034, $p = 0.067$)), at 0.05 significance. Hence, it can be inferred that the results of the segmentation algorithm do not vary significantly from the inter-observer variation.

The multi-organ segmentation algorithm presented in this paper can be used as a general framework to perform: (i) Computer Aided Diagnosis; (ii) volumetric analysis; and (iii) US guided interventions. As an example, we demonstrate the use of this algorithm to assist in US guided interventions by detecting and tracking the carotid artery in 2D US image sequences. The carotid segmentation algorithm of Sec. III-B is tested on 971 images of a publicly available dataset for carotid detection and tracking. The acquisition details and the link to the database can be found in [58]. The results of comparison are provided in Table VII. From the table it can be observed that the proposed method performs better than the performance reported in [58] for the same database. Extension of the proposed method to applications related to computer aided diagnosis and volumetric analysis will be taken up as one of the future works which includes multi-organ segmentation in images with abnormality.

VI. CONCLUSION

In this research work, the echogenicity information in speckle related pixels is made use of in a constructive fashion to categorize tissues into different iso-echoic regions that have been termed as Similarly Reflective Regions. The SRR's are used to detect hyper-echoic anatomical landmarks which are used to detect and segment other organs using local phase based methods. The proposed method has an overall accuracy of over 92% and performs better than existing methods to segment both individual and multiple organs. In conclusion, speckle related pixels need not be considered as noise and the echogenicity information contained in these pixels can be successfully used to detect and segment organs in ultrasound images.

REFERENCES

[1] G. E. Mailloux, M. Bertrand, R. Stampfler, and S. Ethier, "Texture analysis of ultrasound b-mode images by segmentation," *Ultrasonic imaging*, vol. 6, no. 3, pp. 262–277, 1984.

[2] J. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Transactions on Medical Imaging*, vol. 25, no. 8, pp. 987–1010, Aug. 2006.

[3] J. U. Quistgaard, "Signal acquisition and processing in medical diagnostic ultrasound," *IEEE Signal Processing Magazine*, vol. 14, no. 1, pp. 67–74, 1997.

[4] S. W. Smith, H. G. Pavy Jr, and O. T. Von Ramm, "High-speed ultrasound volumetric imaging system. i. transducer design and beam steering," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 38, no. 2, pp. 100–108, 1991.

[5] O. T. Von Ramm, S. W. Smith, and H. G. Pavy Jr, "High-speed ultrasound volumetric imaging system. ii. parallel processing and image display," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 38, no. 2, pp. 109–115, 1991.

[6] R. L. Powis and W. J. Powis, "A thinker's guide to ultrasonic imaging," 1984.

[7] H. Kobatake, "Future cad in multi-dimensional medical images:–project on multi-organ, multi-disease cad system–," *Computerized Medical Imaging and Graphics*, vol. 31, no. 4, pp. 258–266, 2007.

[8] E. Keramidas, D. Iakovidis, D. Maroulis, and S. Karkanis, "Efficient and effective ultrasound image analysis scheme for thyroid nodule detection," *Lecture Notes in Computer Science*, vol. 4633, pp. 1052–1060, 2007.

[9] D. Maroulis, M. Savelonas, D. Iakovidis, S. Karkanis, and N. Dimitropoulos, "Variable background active contour model for computer-aided delineation of nodules in thyroid ultrasound images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 5, pp. 537–543, 2007.

[10] E. A. Mylona, M. A. Savelonas, and D. Maroulis, "Self-parameterized active contours based on regional edge structure for medical image segmentation," *SpringerPlus*, vol. 3, no. 1, p. 424, 2014.

[11] M. Savelonas, D. K. Iakovidis, I. Legakis, D. Maroulis *et al.*, "Active contours guided by echogenicity and texture for delineation of thyroid nodules in ultrasound images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 4, pp. 519–527, 2009.

[12] C.-Y. Chang, Y.-F. Lei, C.-H. Tseng, and S.-R. Shih, "Thyroid segmentation and volume estimation in ultrasound images," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 6, pp. 1348 –1357, Jun. 2010.

[13] H. Garg and A. Jindal, "Segmentation of thyroid gland in ultrasound image using neural network," *Proceedings of 4th International Conference on Computing, Communications and Networking Technologies*, pp. 1–5, 2013.

[14] E. N. Kollorz, D. A. Hahn, R. Linke, T. W. Goecke, J. Hornegger, and T. Kuwert, "Quantification of thyroid volume using 3-d ultrasound imaging," *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 457–466, 2008.

[15] P. Abolmaesumi, S. E. Salcudean, W.-H. Zhu, M. R. Sirouspour, and S. P. DiMaio, "Image-guided control of a robot for medical ultrasound," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 1, pp. 11–23, 2002.

[16] J. Guerrero, S. E. Salcudean, J. Mcewen, B. Masri, S. Nicolaou *et al.*, "Real-time vessel segmentation and tracking for ultrasound imaging applications," *IEEE Transactions on Medical Imaging*, vol. 26, no. 8, pp. 1079–1090, 2007.

[17] D. C. Wang, R. Klatzky, B. Wu, G. Weller, A. R. Sampson, and G. D. Stetten, "Fully automated common carotid artery and internal jugular vein identification and tracking using b-mode ultrasound," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 6, pp. 1691–1699, Jun. 2009.

[18] N. Friedland and D. Adam, "Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing," *IEEE Transactions on Medical Imaging*, vol. 8, no. 4, pp. 344–353, 1989.

[19] M. Wels, Y. Zheng, G. Carneiro, M. Huber, J. Hornegger, and D. Comaniciu, "Fast and robust 3-d mri brain structure segmentation," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 575–583, 2009.

[20] B. Georgescu, X. S. Zhou, D. Comaniciu, and A. Gupta, "Database-guided segmentation of anatomical structures with complex appearance," *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 429–436, 2005.

[21] B. Glocker, O. Pauly, E. Konukoglu, and A. Criminisi, "Joint classification-regression forests for spatially structured multi-object segmentation," *European Conference on Computer Vision (ECCV)*, pp. 870–881, 2012.

[22] T. Kohlberger, M. Sofka, J. Zhang, N. Birkbeck, J. Wetzl, J. Kaftan, J. Declerck, and S. K. Zhou, "Automatic multi-organ segmentation using learning-based segmentation and level set optimization," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 338–345, 2011.

[23] T. Okada, M. Linguraru, M. Hori, Y. Suzuki, R. Summers, N. Tomiyama, and Y. Sato, "Multi-organ segmentation in abdominal ct images," *Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3986–3989, Aug. 2012.

[24] R. Wolz, C. Chu, K. Misawa, M. Fujiwara, K. Mori, and D. Rueckert, "Automated abdominal multi-organ segmentation with subject-specific atlas generation," *IEEE Transactions on Medical Imaging*, vol. 32, no. 9, pp. 1723–1730, Sept. 2013.

[25] G. E. Mailloux, M. Bertrand, R. Stampfler, and S. Ethier, "Texture Analysis Of Ultraosund B-Mode Images by Segmentation," *Ultrasonic Imaging*, vol. 6, no. 3, pp. 262–277, 1984.

[26] W. Lu and J. Tan, "Segmentation of ultrasound fetal images," *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE)*, vol. 4203, pp. 81–90, Dec. 2000.

[27] P. R. Bai, Q. Y. Liu, L. Li, S. H. Teng, J. Li, and M. Y. Cao, "A novel region-based level set method initialized with mean shift clustering for automated medical image segmentation," *Computers in biology and medicine*, vol. 43, no. 11, pp. 1827–1832, 2013.

[28] M. G. Linguraru, J. A. Pura, V. Pamulapati, and R. M. Summers, "Statistical 4d graphs for multi-organ abdominal segmentation from multiphase ct," *Medical image analysis*, vol. 16, no. 4, pp. 904–914, 2012.

[29] N. S. Narayan, P. Marziliano, J. Kanagalingam, and C. G. Hobbs, "Speckle in ultrasound images: Friend or foe?" in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*. IEEE, 2014, pp. 5816–5820.

[30] Y. Yu and S. T. Acton, "Speckle reducing anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 11, no. 11, pp. 1260–1270, 2002.

[31] Y. Yue, M. M. Croitoru, A. Bidani, J. B. Zwischenberger, and J. W. Clark Jr, "Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images," *IEEE Transactions on Medical Imaging*, vol. 25, no. 3, pp. 297–311, 2006.

[32] Y. Deng, Y. Wang, and Y. Shen, "Speckle reduction of ultrasound images based on rayleigh-trimmed anisotropic diffusion filter," *Pattern Recognition Letters*, vol. 32, no. 13, pp. 1516–1525, 2011.

[33] T. Eltoft, "Modeling the amplitude statistics of ultrasonic images," *IEEE Transactions on Medical Imaging*, vol. 25, no. 2, pp. 229–240, 2006.

[34] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound b-scans." *IEEE Transactions on Sonics and Ultrasonics*, vol. 30, no. 3, pp. 156–163, 1983.

[35] R. W. Prager, A. H. Gee, G. M. Treece, and L. H. Berman, "Decompression and speckle detection for ultrasound images using the homodyned k - distribution," *Pattern Recognition Letters*, vol. 24, no. 4, pp. 705–713, 2003.

[36] R. Martí, J. Martí, J. Freixenet, J. Vilanova, and J. Barceló, "Robust speckle detection in ultrasound images: evaluation aspects," *Proceedings of the 8th International Conference on Quality Control by Artificial Vision*, pp. 63 560P–63 560P, 2007.

[37] J. Seabra and J. Sanches, "Modeling log-compressed ultrasound images for radio frequency signal recovery," *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008.

[38] W. R. Hedrick, D. L. Hykes, and D. E. Starchman, *Ultrasound physics and instrumentation*. St. Louis, Mo. : Elsevier Mosby, 2005.

[39] J. M. Thijssen, "Ultrasonic speckle formation, analysis and processing applied to tissue characterization," *Pattern Recognition Letters*, vol. 24, no. 4, pp. 659–675, 2003.

[40] B. B. Tempkin, *Ultrasound scanning : principles and protocols*. St. Louis, Mo. : Elsevier Saunders, 2009.

[41] N. S. Narayan, P. Marziliano, and C. G. L. Hobbs, "Automatic removal of manually induced artefacts in ultrasound images of thyroid gland," *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3399–3402, Jul. 2013.

[42] R. A. Sofferman and A. T. Ahuja, *Ultrasound of the Thyroid and Parathyroid Glands*. New York, NY : Springer, 2012.

[43] G. B. Thomas and R. L. Finney, *Calculus and analytic geometry*. Reading, Mass. : Addison-Wesley, 1996.

[44] W. Gander, G. Golub, and R. Strebel, "Least-squares fitting of circles and ellipses," *BIT Numerical Mathematics*, vol. 34, no. 4, pp. 558–578, 1994.

[45] B. Rahmatullah, A. T. Papageorghiou, and J. A. Noble, "Integration of local and global features for anatomical object detection in ultrasound," *Medical Image Computing and Computer-Assisted Intervention(MICCAI)*, pp. 402–409, 2012.

[46] T. M. Rackham, S. Rueda, C. L. Knight, and J. A. Noble, "Ultrasound image segmentation using feature asymmetry and shape guided live wire," *SPIE Medical Imaging*, pp. 86 690P–86 690P, 2013.

[47] D. Boukerroui, J. A. Noble, and M. Brady, "On the choice of band-pass quadrature filters," *Journal of Mathematical Imaging and Vision*, vol. 21, no. 1-2, pp. 53–80, 2004.

[48] M. Mellor and M. Brady, "Phase mutual information as a similarity measure for registration," *Medical image analysis*, vol. 9, no. 4, pp. 330–343, 2005.

[49] D. Jayachandra and A. Makur, "Directional variance: A measure to find the directionality in a given image segment," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 1551–1554, Jun. 2010.

[50] P. A. Yushkevich, J. Piven, H. Cody Hazlett, R. Gimpel Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[51] R. Unnikrishnan and M. Hebert, "Measures of similarity," *Proceedings of the 7th IEEE Workshops on Application of Computer Vision*, vol. 1, pp. 394–394, 2005.

[52] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," *Proceedings of the 8th IEEE International Conference on Computer Vision*, vol. 2, pp. 416–423, 2001.

[53] M. Meil, "Comparing clusterings: an axiomatic view," *Proceedings of the 22nd international conference on Machine learning*, pp. 577–584, 2005.

[54] J. Freixenet, X. Muñoz, D. Raba, J. Martí, and X. Cufí, "Yet another survey on image segmentation: Region and boundary information integration," *Proceedings of the European Conference on Computer Vision*, pp. 408–422, 2002.

[55] A. Y. Yang, J. Wright, Y. Ma, and S. S. Sastry, "Unsupervised segmentation of natural images via lossy data compression," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 212–225, 2008.

[56] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "Toward objective evaluation of image segmentation algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 929–944, 2007.

[57] S. Golemati, J. Stoitsis, E. G. Sifakis, T. Balkizas, and K. S. Nikita, "Using the hough transform to segment ultrasound images of longitudinal and transverse sections of the carotid artery," *Ultrasound in medicine & biology*, vol. 33, no. 12, pp. 1918–1932, 2007.

[58] K. Říha, J. Mašek, R. Burget, R. Beneš, and E. Závodná, "Novel method for localization of common carotid artery transverse section in ultrasound images using modified viola-jones detector," *Ultrasound in medicine & biology*, vol. 39, no. 10, pp. 1887–1902, 2013. [Online]. Available: splab.cz/en/research/zpracovani-medicinskych-signalu/databaze/artery