Segmentation of the spinous process and its acoustic shadow in vertebral ultrasound images

Florian Berton[a], Farida Cheriet[a,c], Marie-Claude Miron[c],
Catherine Laporte[b,c,*]

[a]*Polytechnique Montréal, 2900 boul. Édouard-Montpetit, Montréal, QC H3T 1J4, Canada*
[b]*École de technologie supérieure, 1100 Rue Notre-Dame O., Montréal, QC H3C 1K3, Canada*
[c]*Sainte-Justine University Hospital Research Centre, 3175 Chemin de la Côte-Sainte-Catherine, Montréal, QC H3T 1C4, Canada*

Abstract

Spinal ultrasound imaging is emerging as a low-cost, radiation-free alternative to conventional X-ray imaging for the clinical follow-up of patients with scoliosis. Currently, deformity measurement relies almost entirely on manual identification of key vertebral landmarks. However, the interpretation of vertebral ultrasound images is challenging, primarily because acoustic waves are entirely reflected by bone. To alleviate this problem, we propose an algorithm to segment these images into three regions: the spinous process, its acoustic shadow and other tissues. This method consists, first, in the extraction of several image features and the selection of the most relevant ones for the discrimination of the three regions. Then, using this set of features and a LDA classifier, each pixel of the image is classified as belonging to one of the three

*Corresponding author

Email addresses: `florian.berton@polymtl.ca` (Florian Berton), `farida.cheriet@polymtl.ca` (Farida Cheriet), `marie-claude_miron@ssss.gouv.qc.ca` (Marie-Claude Miron), `catherine.laporte@etsmtl.ca` (Catherine Laporte)

regions. Finally, the image is segmented by regularizing the pixel-wise classification results to account for some geometrical properties of vertebrae. The feature set was first validated by analyzing the classification results across a learning database. The database contained 107 vertebral ultrasound images acquired with convex and linear probes. Classification rates of 84%, 92% and 91% were achieved for the spinous process, the acoustic shadow and other tissues, respectively. Dice similarity coefficients of 0.72 and 0.88 were obtained respectively for the spinous process and acoustic shadow, confirming that the proposed method accurately segments the spinous process and its acoustic shadow in vertebral ultrasound images. Furthermore, the centroid of the automatically segmented spinous process was located at an average distance of 0.38 mm from that of the manually labeled spinous process, which is on the order of image resolution. This suggests that the proposed method is a promising tool for the measurement of the Spinous Process Angle and, more generally, for assisting ultrasound-based assessment of scoliosis progression.

Keywords: Image processing, Ultrasound, Vertebrae, Image segmentation, Pattern classification, Acoustic shadow, Scoliosis

---

1. Introduction

Ultrasound imaging is abundantly used in medical diagnostic and interventional procedures requiring the visualization of soft tissues (Zhang et al., 2015; Baumann et al., 2009; Unsgaard et al., 2006; Sato et al., 1998; Wu et al., 2003; Zhang et al., 2010). Recently, it has also emerged as a useful, low cost tool for applications that involve imaging bone structures such as guidance for epidural anesthesia (Ashab et al., 2013), bone fracture assessment (Haci-

haliloglu et al., 2009) and scoliosis assessment and monitoring (Young et al., 2015; Ungi et al., 2014). However, the interpretation of bone ultrasound images is very challenging. Ultrasound waves are totally reflected by bone surfaces, creating an acoustic shadow below them and, depending on the orientation of the probe, a bright area at the soft tissue-bone interface.

In this paper, we consider the specific context of spine imaging for scoliosis, where ultrasound-based methods have shown potential as a radiation-free alternative to X-ray imaging. Potential applications include deformity assessment (Ungi et al., 2014; Cheung et al., 2015; Young et al., 2015), spinal brace adjustment (Lou et al., 2015; Li et al., 2012) and dense 3D spine surface reconstruction (Nguyen et al., 2015). The current gold standard measure for assessing scoliosis (as well as its progression or the effectiveness of a treatment) is the Cobb angle, measured on postero-anterior X-ray images. Most of the work cited above has focused on reproducing this measurement, or rather proxies of this measurement (e.g., using the Spinous Process Angle (Li et al., 2012), the Center of Laminae method (Nguyen et al., 2015) or the Transverse Process Angle (Ungi et al., 2014)), since the orientation of the vertebral endplates, which define the Cobb angle, cannot be visualized in ultrasound images. The proposed *methods* of measurement mirror what is typically done in clinical practice with the X-ray images, i.e., they are entirely based on the manual identification of anatomical features, which is time consuming and lacks repeatability. These difficulties are exacerbated by the low perceptual quality of vertebral ultrasound images.

In the particular case of the spinous process, which is the prominent bone feature appearing in transverse vertebral ultrasound images, the inter-

face with soft tissue is very short and its brightness is strongly dependent on the orientation of the probe. In addition, while its acoustic shadow could provide useful cues about the shape of the vertebra and the possible location of the laminae and transverse processes, the boundary of this shadow is difficult to delineate accurately in practice. As a first step towards automated vertebral landmark identification, this paper proposes a method to automatically segment the most prominent features appearing in transverse vertebral ultrasound images: the spinous process and its acoustic shadow.

Several methods were developed for the automatic detection of key anatomical structures in vertebral ultrasound images in the context of ultrasound-guided epidural needle insertion. Their goal is ultimately to determine where to insert the needle for effective injection of the anesthetic. Automatic extraction of the lamina was proposed by Tran and Rohling (2010). Their method uses a ridge detector followed by matching with a binary template of the lamina to capture all the lamina contained in the image; it then segments the ligamentum flavium (also using template matching) on the regions below the lamina. The method detects the lamina and ligamentum flavium with a success rate of 87%. Kerby et al. (2008) proposed a method for the automatic identification of lumbar levels in panoramic ultrasound images. This method extracts a wave-like profile from the image with a median filter and two linear filters operating in the horizontal and vertical directions. Lumbar levels are identified as the local maxima of this profile. Ashab et al. (2013) used this method for guidance by projecting the image of the segmented lumbar levels onto the back of the patient. This system was shown to provide needle insertion accuracy within a clinically acceptable range. The previously

4

cited methods are concerned with vertebral ultrasound images acquired in the paramedian view. In our case, we are interested in the transverse view, which provides more information regarding the shape of each vertebra, and, when combined to a freehand 3D acquisition protocol, is better suited to the use of the Spinous Process Angle and the Center of Laminae methods for assessing scoliosis.

Recently, a method for automatic detection of the bone and inter-spinous regions in vertebral ultrasound images acquired in the transverse plane was developed by Yu et al. (2013). The method first involves pre-processing the image using a difference of Gaussians filter and local intensity normalization. Then, binary templates describing the transverse processes and the vertebral body are used with a template matching algorithm and a position correlation function in order to identify the key anatomical landmarks of interest. This approach was extended by using machine learning techniques to automatically detect the transverse processes. In this context, Yu and Tan (2014) tested a neural network, a cascading classifier (Yu et al., 2014) and a support vector machine (Yu et al., 2015) as classifiers. Support vector machines obtained the best results with a success rate of 92.3% for the automatic detection of bone and inter-spinous regions. This work focused on the lumbar region of the spine, where the spinous process is larger and aligned with the vertebral body and the transverse processes. This is not the case in the thoracic region of the spine, which we are also interested in.

Alternatively, spinal anatomy could implicitly be detected by establishing correspondence (i.e. registration) between freehand 3D ultrasound data and a generic 3D model (or, more generally, a statistical atlas) of the spine

surface, using methods proposed in the context of image-guided epidural anesthesia. Khallaghi et al. (2010) optimize the similarity between the ultrasound data and synthetic ultrasound images generated from the 3D spine model using a simulation method. Hacihaliloglu et al. (2014) propose instead to match bone surfaces from the model to bone surfaces enhanced using local phase tensor features in the ultrasound image. Dense segmentation of the ultrasound images, as proposed in this paper, and segmentation of acoustic shadow in particular, could certainly be exploited as part of such model-based registration strategies to utilize only meaningful correspondences, as proposed by Coupé et al. (2012) in the context of image-guided neurosurgery.

Outside the context of spine imaging, automatic segmentation of bone surfaces and acoustic shadows in ultrasound images has been more deeply explored. Daanen et al. (2004) proposed a segmentation algorithm for bone surfaces using the main properties of these regions: highly echogenic structures, creation of acoustic shadow, specular reflection, continuous structures and homogeneous contrast within these structures. This algorithm achieved a mean localization error of 1 mm for sacrum images. Foroughi et al. (2007) combined a local edge detector and the sum of the intensities below each pixel to create a "bone probability map", then segmented the bone surfaces by minimizing a cost-function based on the continuity and smoothness of these regions. The latter two methods exploited the continuity of bone regions for segmentation, which is very effective for large bone structures, but not for short surfaces like that of the spinous process. Another method proposed by Hacihaliloglu et al. (2009) measures the presence of a bone surface in ultrasound images using local phase symmetry. This was applied to ultra-

sound images of the radius. In vertebral ultrasound images, fat tissues and ligaments are present in addition to bone, and these structures also lead to high phase symmetry, thereby limiting the specificity of the approach. Hellier et al. (2010) developed a method to detect the boundary between acoustic shadows and the tissues above them using statistical tests along each scan line, accounting for the geometry of the ultrasound transducer. This method was applied to the shadows cast by brain tumors, with a Dice similarity coefficient of 0.95 at a depth of 6 cm. More recently, Karamalis et al. (2012) used random walks incorporating a model of ultrasound image formation to create a confidence map which highlights acoustic shadows. A Dice similarity coefficient of 0.85 was obtained for a database of humerus ultrasound images.

In preliminary work (Berton et al., 2015), we developed a method to automatically segment vertebrae in ultrasound images. The method combines a variety of features (many of them borrowed from the papers cited above) to drive a pixel-wise random forest classifier (Breiman, 2001). A spatial regularization step based on context-specific geometrical constraints was used to automatically segment the spinous process and its acoustic shadow post-classification. We obtained good results with classification rates of 81.97% for bone surfaces and 91.01% for acoustic shadow. In this paper, we extend our preliminary work in the following ways:

- An additional feature based on Gabor filters is introduced that helps characterize the contrast between the acoustic shadow and the visible surrounding tissues.

- The regularization step is improved by constraining the segmentation to identify a unique region corresponding to the spinous process and to

account for the convexity of the acoustic shadow cast by the spinous process.

- A feature selection step is also introduced to select the most relevant features. The selected features are rigorously validated based on their classification performance within a learning database.

- The image databases used for learning and testing now include data from a larger number of individuals as well as images acquired using two different transducers to better demonstrate the robustness of the method.

- The additional data and features allow us to use conventional linear discriminant analysis (LDA) for classification, which simplifies the segmentation process significantly without compromising accuracy or generalization.

- The method is thoroughly validated through application on the larger test set, as well as detailed analysis of its sensitivity and specificity. Dice similarity coefficient and Hausdorff distance measures are also computed to further evaluate segmentation accuracy.

The rest of this paper is organized as follows: in Section 2, the different steps of the segmentation algorithm and the experimental methodology are described in detail. Our results are presented in Section 3, illustrating the promise of the proposed approach. Conclusions and directions for future work are given in Section 4.

2. Materials and methods

The purpose of the method is to classify each pixel of the ultrasound image into one of three categories: "spinous process", "acoustic shadow" and "other tissues" (which correspond to the other anatomical structures in the image such as muscles, fat, ligaments, etc.). The principle of the method is illustrated Fig. 1. First, feature extraction is applied to the images of the learning database, providing one feature vector for each pixel. Next, feature selection is performed and pixels representing the three categories are randomly chosen from the images of the learning database to create the data set used to train a LDA classifier. The automatic segmentation of a new image is achieved in three steps. First, features are once again computed for each pixel of the image. Then, these pixels are classified into one of the three categories using the trained LDA classifier. Finally regularization of the different regions of the segmented image is performed to account for geometric constraints specific to vertebral images.

2.1. Materials for image acquisition and processing

The ultrasound images of the spine used in this study were collected from 7 healthy volunteers between 21 years and 45 years of age at Sainte-Justine Hospital, Montreal, Canada. During the acquisition, the volunteers were lying on a stretcher in prone position, and both the thoracic and lumbar regions of the spine were scanned using a Siemens Acuson S2000 ultrasound system (Siemens Healthcare, Erlangen, Germany). Images were acquired using both a curvilinear ultrasound probe (6C2) with a scanning depth of 6 cm and central frequency set to 6 MHZ and a linear ultrasound probe (14L5)
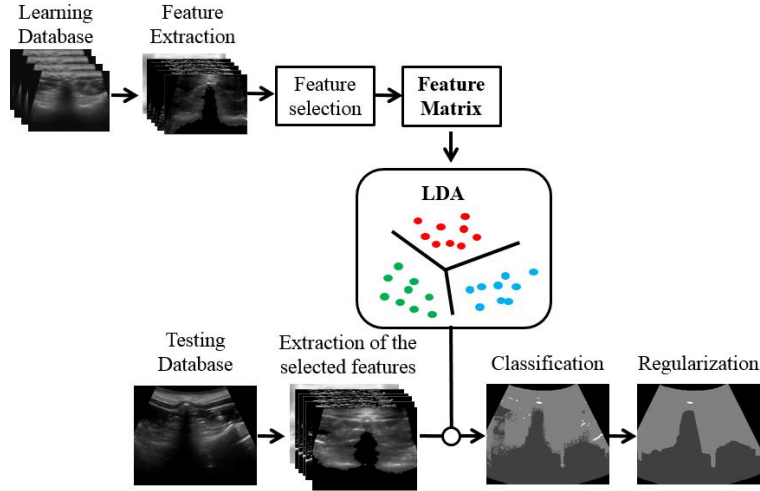
Figure 1: Block diagram of the proposed image segmentation method.

with a scanning depth of 4.5-6 cm and central frequency set to 9 MHz. A database of 175 images was thus obtained; their sizes varied from (400×260) to (680×553) pixels. The algorithms were implemented with MATLAB R2014a and run on a personal computer (Intel Core i7-4720HQ, 8GB memory and NVIDIA GTX965M).

2.2. Feature extraction

Several feature extraction methods previously used by others for describing bones surfaces, including image gradient, Foroughi et al. (2007)'s bone probability map and phase symmetry (Hacihaliloglu et al., 2009), were implemented. In addition, to characterize the acoustic shadow, the rupture points described by Hellier et al. (2010) were also detected. Furthermore, to discriminate the acoustic shadow and the spinous process from other tissues present in the ultrasound images, texture descriptors based on Local Binary Patterns (LBP) (Guo et al., 2010) and Gabor filters (Clausi and Jernigan,
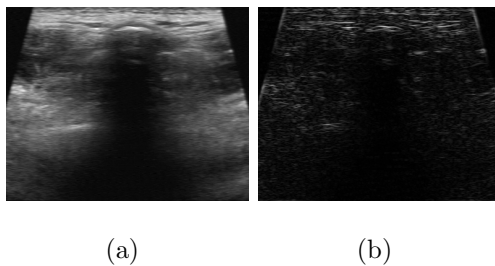
10

(a)   (b)

Figure 2: Intensity (left) and gradient (right) of a vertebral ultrasound image.

2000) were also extracted. Each of these features is described in more detail below.

2.2.1. Intensity and gradient images

The intensity image and its gradient obtained with the Sobel gradient operator (Fig. 2) are used as features. Some information about the acoustic shadow and the bone surfaces is naturally provided by the image itself. The gradient image highlights the boundaries between distinct structures.

2.2.2. Bone probability

Foroughi et al. (2007) proposed a *bone probability map* based on two key characteristics of bone surfaces in ultrasound images: the presence of a bright ridge due to specular reflection, and the presence of acoustic shadow below this bright ridge. To obtain the bone probability map, a ridge map that highlights potential specular reflections is first built by filtering the image with a Laplacian of Gaussian filter, and this map is multiplied by a map of the *quantity of shadow* (SH), obtained by computing, for each pixel, the sum of the intensities of pixels lying below it:

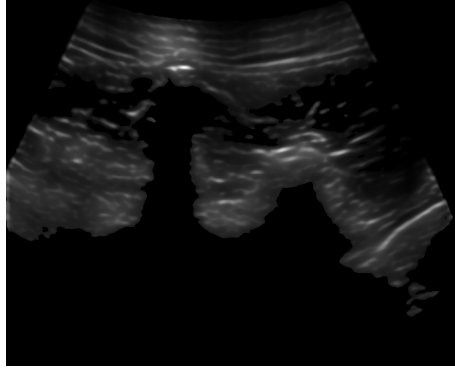$$SH(x,y) = \frac{\sum_{j=y}^{H} G(j-y)I(x,j)}{\sum_{j=y}^{H} G(j-y)} \quad , \tag{1}$$

11

Figure 3: Bone probability map of Foroughi et al. (2007). The map highlights the spinous process, but also displays bright areas that correspond to other structures.

where x and y are pixel coordinates, $G(.)$ and $I(.)$ represent a Gaussian windowing function and the image intensity, respectively, and H is the number of rows in the image. The standard deviation of the Gaussian window was empirically set to 2/3 of the image height. While the resulting bone probability map does highlight the spinous process, it may also highlight other tissues like ligaments or fat near a bone surface (Fig. 3). Such tissues also generate strong reflections and the acoustic shadow created by the nearby bone structure can be incorrectly associated with them.

2.2.3. Phase symmetry

Since bone surfaces are smooth and highly echogenic, Hacihaliloglu et al. (2009) use a ridge detector based on local phase symmetry to detect them. Measurements of local phase symmetry $PS(x, y)$ at each pixel position (x, y), (Fig. 4) are obtained by combining the image's responses to Log Gabor filters at 2 different scales and 6 orientations:

$$PS(x, y) = \frac{\sum_r \sum_m \lfloor [|e_{rm}(x, y)| - |o_{rm}(x, y)|] - T_r \rfloor}{\sum_r \sum_m \sqrt{e_{rm}^2(x, y) + o_{rm}^2(x, y)} + \epsilon} \quad . \tag{2}$$

12

In Eq. 2, T_r is a noise threshold, and $e_{rm}(x, y)$ and $o_{rm}(x, y)$ correspond to the responses of quadrature 2D Log Gabor filters (Hacihaliloglu et al., 2009) with scale r and orientation m in the frequency domain:

$$e_{rm}(x, y) = \Re(F^{-1}(2DLGF(I(x, y)))), \tag{3}$$

$$o_{rm}(x, y) = \Im(F^{-1}(2DLGF(I(x, y)))), \tag{4}$$

where $F()$ and $F^{-1}()$ correspond respectively to the forward and inverse Fourier transforms, $\Re()$ and $\Im()$ respectively denote real and imaginary parts of a complex number and $I(x, y)$ is the ultrasound image. $2DLG$ denotes the two-dimensional frequency domain Log Gabor filter defined as (Field, 1987)

$$2DLG(\omega, \phi) = \exp\left(-\left(\frac{\log((\omega/\omega_0)^2}{2\log((K/\omega_0)^2} + \frac{\phi - \phi_0}{2\sigma_\phi}\right)\right), \tag{5}$$

where K is the bandwidth of the filter in the radial direction, ω_0 is the center frequency, ϕ_0 is the orientation of the filter, i.e., the ridge orientation to which it is most sensitive, and σ_ϕ is the orientation spacing between the different filters. In this work, all free parameters were set exactly as recommended by Hacihaliloglu et al. (2009).

Fig. 4 shows the phase symmetry of a vertebral ultrasound image. The phase symmetry image was enhanced and normalized to demonstrate how this feature enhances bone structures but also other interfaces present in the image such as layers of fat and muscle fibers.

2.2.4. Rupture points

This feature was developed as part of the acoustic shadow detection method proposed by Hellier et al. (2010). For each scan line of the ultrasound image, a statistical test is applied based on two criteria to detect the
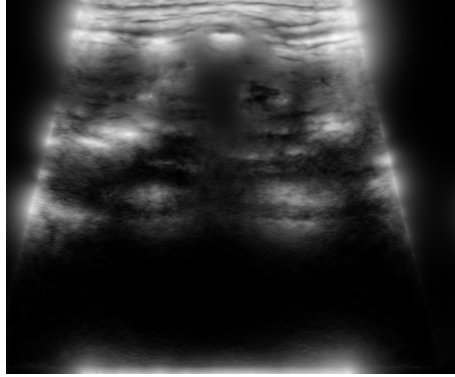
13

Figure 4: Image feature of the phase symmetry.

boundary of a potential acoustic shadow. The first criterion is a rupture in the signal. For each point of the scan line S, the local symmetric entropy coefficient ρ is computed as

$$\rho(x) = \sum_{i=1}^{n} \left(S(x-i) \log \frac{S(x-i)}{S(x+i)} + S(x+i) \log \frac{S(x+i)}{S(x-i)} \right), \qquad (6)$$

where x is depth of the point tested along the scan line and n is a neighborhood size, which we empirically set to 8. *Rupture points* are then determined to be the maxima of ρ for each scan line. Acoustic shadows are regions where the noise variance is low; thus, the second criterion is based on the similarity between the statistics of the signal below the rupture points and noise model estimated above the rupture points. The image is binarized, associating shadow with regions below the shallowest rupture points failing the statistical test. As demonstrated in Fig. 5, this does provide a coarse indication of the presence of acoustic shadow, but sometimes detects acoustic shadow where there is none.
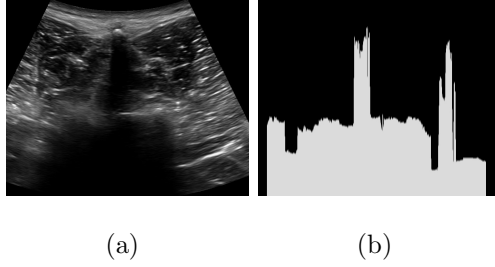
14

(a)   (b)

Figure 5: Vertebral ultrasound image (left) and binarized image based on rupture point detection (right).

2.2.5. Local Binary Patterns

LBPs (Guo et al., 2010) are often used as image texture descriptors. In this representation, each pixel is assigned two binary patterns derived respectively from the sign s and magnitude m of the dissimilarities observed between this pixel and its neighbors:

$$LBP_{\text{Sign}_{P,R}} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, \quad s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & < 0 \end{cases} \tag{7}$$
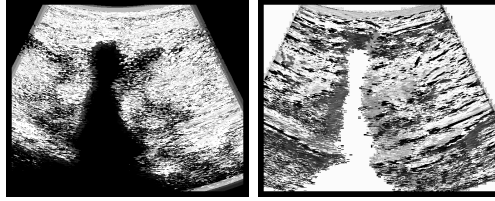
and

$$LBP_{\text{Mag}_{P,R}} = \sum_{p=0}^{P-1} m(g_p - g_c)2^p, \quad m(x) = |x| \ . \tag{8}$$

In Eq. 7 and Eq. 8, g_c and g_p respectively correspond to the intensities of the central pixel and the neighboring pixels, P is the neighborhood size and R is a scale parameter. In this work, we empirically set $P = 8$. The scale parameter R was optimized in terms of its Fisher score (Duda et al., 2000) with respect to classification, based on independent treatment of LBP sign and magnitude features:

$$D(F_R) = \frac{\sum_{c=1}^{C} n_c (\mu_c^{F_R} - \mu^{F_R})^2}{\sum_{c=1}^{C} n_c n_c (\sigma_c^{F_R})^2}, \tag{9}$$

15

Table 1: Fisher score for each features

Scale	LBP sign	LBP magnitude
2	0.533	0.045
4	0.707	0.0474
8	0.925	0.0553
16	0.998	0.0748
20	0.971	0.0726
24	0.933	0.0725
30	0.860	0.0631

(a) (b)

Figure 6: Sign (left) and magnitude (right) LBP features.

where F_R is the sign or magnitude LBP feature at scale R and n_c, μ^{F_R} and $\sigma_c^{F_R}$ correspond respectively to the number, mean and standard deviation of F_R for the class c (which can correspond to spinous process, acoustic shadow or other tissues). Fisher scores $D(F_R)$ for different values of R are shown Table 1, and according to these data, the scale of the LBP was set to 16.

LBPs are texture descriptors. When applied to ultrasound images, they especially highlight the speckle, so that there will be a contrast between speckle-free regions (e.g., most of the acoustic shadow) and regions with speckle, as shown in Fig. 6.

2.2.6. Gabor filter response

The combined response of the input image I to a set of Gabor filters at 8 orientations evenly sampled on the interval $(0, 2\pi)$,

$$GTD(x, y) = \sum_{i=1^8} g_i(x, y) * I(x, y), \qquad (10)$$

was used as an additional texture descriptor for the image. In Eq. 10, '$*$' is the convolution operator and g_i denotes the ith Gabor filter of orientation θ_i defined in the spatial domain as

$$g_i(x, y) = \exp\left(-\frac{x_i'^2 + \gamma^2 y_i'^2}{2\sigma^2}\right) \exp\left(j(2\pi\frac{x_i'}{\lambda} + \psi)\right) \quad \psi \in \{0, \pi\}, \qquad (11)$$

where $x_i' = x\cos(\theta_i) + y\sin(\theta_i)$ and $y_i' = -x\sin(\theta_i) + y\cos(\theta_i)$ and $j = \sqrt{-1}$. We empirically set the wavelength of the sinusoidal factor $\lambda = 1$, the filter scale $\sigma = 2.7$ and the spatial aspect ratio $\gamma = 0.4$. The results were not found to be especially sensitive to these choices of parameters.

The resulting feature enhances bright linear structures. These correspond to bones structures, but also to the layers of fat and muscle fibers in the ultrasound image (Fig. 7). Thus, the contrast between the acoustic shadow and the visible surrounding tissues is enhanced.

2.3. Feature selection

To determine the most relevant subset of features necessary for classification, we use the sequential floating forward selection (SFFS) method proposed by Pudil et al. (1994). Its principle is to start with an empty set of features, iteratively adding the best features and removing the worst features according to a utility criterion. In our case, we seek to maximize the stability of the classifier. To this end, classification rates are measured each time
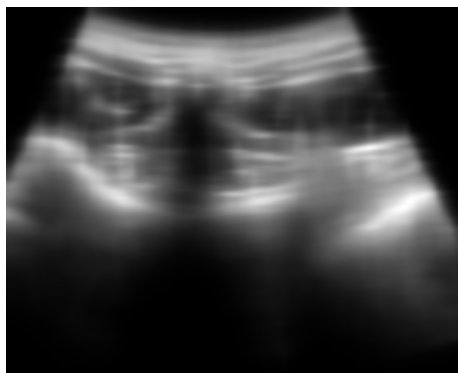
17

Figure 7: Gabor filter response feature.

a feature is added or removed from the set using threefold cross-validation. The algorithm terminates when all the features were added to the set or there is no significant change in classification rate. The resulting set of features is taken to be optimal.

2.4. Classification

We used LDA McLachlan (2004), a generalization of Fisher's linear discriminant, as a classification method because it is simple and fast. Its principle is to create linear sub-spaces, obtained from linear combinations of the features, which maximize the ratios of between-class variance to within-class variance of the learning database. Based on the projections of the training data onto these sub-spaces, linear decision boundaries are computed that optimally separate the data belonging to each class from the data belonging to each of the other classes. New data requiring classification are subsequently projected onto the learned sub-spaces and assigned to a class based on the learned decision boundaries.

2.5. Regularization

Pixel classification provides an initial segmentation of the image into regions corresponding to the "spinous process", "acoustic shadow" and "other tissues" classes. Regularization of the resulting regions is necessary to reflect the following context-specific geometrical constraints (Fig 8):

- **Neighboring pixels are more likely to belong to the same region than to belong to different regions**. To account for this, opening morphological operations with small structuring elements are first applied to the segmented image.

- **The acoustic shadow of a vertebra is a very large region**. In our images, the average acoustic shadow occupies on the order of 200,000 pixels. Thus, connected components initially labeled as "acoustic shadow" are relabeled as "other tissues" if they are smaller than 0.3% of this average size. This operation removes artifacts that could interfere with the next regularization operations.

- **Pixels that are part of the spinous process region lie just above the top of the acoustic shadow**. Therefore, connected components classified as "spinous process" which are not in the neighborhood of the top of the acoustic shadow are relabeled as "other tissues". For each of the remaining connected components, the mean of the bone probability feature (see Section 2.2.2) is calculated. The connected component with the largest bone probability value is the only one to keep its "spinous process" label and the others are relabeled as "other tissues".
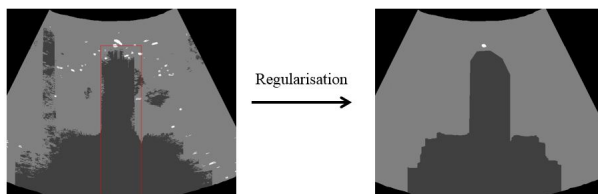
Figure 8: Sample result for the regularization step. The boundary of the region of interest for the acoustic shadow is in red.

- **A vertebra casts only one acoustic shadow and this shadow has a smooth, convex shape**. Therefore, a region of interest around the shallowest "acoustic shadow" pixel is defined and all connected components labeled as " acoustic shadow" within the region of interest are merged. The region labeled as "acoustic shadow" is then expanded to fill the boundaries of its convex hull. Finally, the (now convex) boundary of the acoustic shadow is smoothed using a median filter.

2.6. Experimental methodology

2.6.1. Image database

The gold standard used for training and testing the classifiers was a manual segmentation of all images validated by a radiologist. Furthermore, the visual quality of all images in the database were also assessed by the radiologist and the images of exceedingly poor quality were removed from the database. The final database was thus comprised of 107 images (out of the original 175) of reasonable quality. We used 2/3 of the images from each transducer for the learning database and the rest of images formed the testing database.

20

Since the spinous process occupies a much smaller area in the images than the other two regions, using all the pixels from the learning database to train the classifier would lead to a highly unbalanced training set and poor classifier performance. Therefore, while the training set contains all the pixels belonging to the spinous process in the learning database, it does not include all pixels from the other two classes. For these, a total of seven times more pixels than for the "spinous process" class are used, thereby preserving a known, constant and reasonable ratio between the number of pixels for the spinous process and the number of pixels for the other two classes. These pixels are chosen randomly.

2.6.2. Assessment criteria

Four measures are used to assess accuracy of the proposed method:

1. **Classification rate**. The classification rate measures the sensitivity of the segmentation algorithm with respect to each class.

2. **Dice similarity coefficient** (DSC). The DSC measures the spatial overlap between two point sets A and B:

$$DSC = \frac{2|A \cap B|}{|A| + |B|}.$$ (12)

   Here, A and B respectively correspond to the automatically and manually segmented spinous processes (or acoustic shadows). $DSC = 1$ corresponds to a complete overlap.

3. **Hausdorff distance**. The Hausdorff distance, defined as

$$HD = \max(h(A, B), h(B, A)), \quad h(A, B) = \min_{a \in A} \max_{b \in B} ||a - b||,$$ (13)

measures how far two sets of points A and B are from each other. In this paper, we evaluate the Hausdorff distance between the boundaries of the automatically and manually segmented acoustic shadows. Only the boundary in a region of interest around the spinous process (see Fig. 8) is considered because the acoustic shadow created in the vicinity of the spinous process is the region which interests us most.

4. **Euclidean distance between centroids**. We apply this measure to the centroids of the automatically and manually segmented spinous processes to evaluate our algorithm's localization accuracy.

3. Results and discussion

3.1. Feature relevance

This section concerns the relevance of the image features considered. We performed feature selection on the learning database using the SFFS method summarized in Section 2.3 with threefold cross-validation. Fig. 9 shows the classification rate, corresponding to the true positive rate, for the three classes, as the number of included features in SFFS increases. These curves do not directly show specificity when taken individually, but specificity is indirectly reflected in the growth of the classification rate for the other classes. Except for the intensity of the image, all the features were kept in the final feature set.

Regarding the acoustic shadow, Fig. 9 shows that after including the rupture point, LBP sign and Gabor filter features, the classification rate does not improve, suggesting that these three features adequately predict the presence of acoustic shadow. This is not surprising since the rupture point
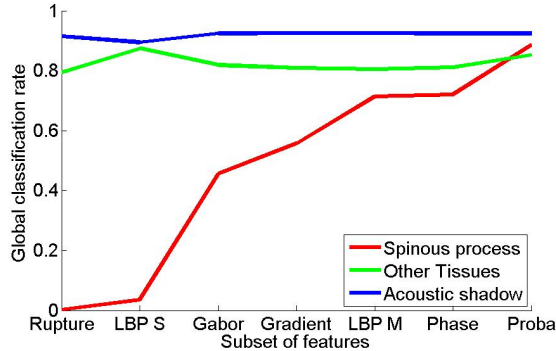
Figure 9: Classification rate for the three regions during the SFFS procedure.

feature is specifically designed to detect acoustic shadows and the Gabor filters and the LBP sign highlight most of the anatomical structures which surround them.

For the spinous process, the classification rate starts at a very low value and increases steadily with the inclusion of new features. Each of these features characterizes one property of bone surfaces, but other anatomical structures also share this property. Consequently, the combination of all these features is essential to the accurate segmentation of the spinous process.

The confusion matrix (normalized over the total number of pixels in each image and averaged over all the images in the test dataset, see Table 2) shows mis-classification tendencies for each class. It also illustrates the proportion of the images occupied by each region. "Other tissues" are very significant since they compose 57% of an image on average. Thus, the better the classification rate for "other tissues", the more the specific the method is with respect to the spinous process and its acoustic shadow. Generally, when mis-classified, pixels belonging to the spinous process tend to be mis-classified as "other tissues", and similarly for the pixels belonging to the acoustic shadow.

23

Table 2: Normalized confusion matrix of the method, averaged over all the images

		Classifier output		
		Spinous process	Acoustic shadow	Other tissues
Actual value	Spinous process	0.08%	0%	0.01%
	Acoustic shadow	0.03%	38.95%	3.86%
	Other tissues	0.94%	6.94 %	49.19 %

There is little confusion between the acoustic shadow and the spinous process, indicating that these two regions are well separated by the classifier. "Other tissues" are frequently mis-classified as part of the acoustic shadow. This is one reason why a regularization step is required to improve the segmentation.

3.2. Segmentation results

In this section, the segmentation of vertebral ultrasound images is assessed. We used the set of features obtained with the SFFS, classified each image of the testing database and applied regularization as described in Section 2. Table 3 shows our results averaged over all the images. The results obtained after regularization are very promising, with classification rates around 85% for the spinous process, 92% for the acoustic shadow and 91% for the "other tissues". Typical segmentation examples are shown in Fig. 10.

Segmentation quality measures are reported in Tables 3 and 4. The proposed method generally performed an accurate segmentation of the acoustic

Table 3: Classification rates before and after regularization. The results before regularization were computed from the normalized confusion matrix of Table 2. The results after regularization are averages over the images of the test database.

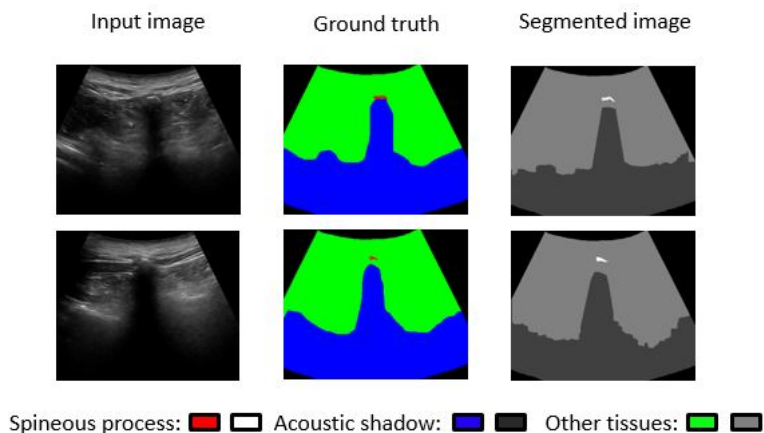|	Classification rate	
	Before regularization	After regularization
Spinous process	88.88%	84.72% ± 19.59%
Acoustic shadow	90.92%	92.49% ± 5.09%
Other tissues	86.19%	90.75% ± 7.15%

Figure 10: Typical segmentation results. Left: input ultrasound image, Middle: Manual segmentation, Right: Automatic segmentation.

25

Table 4: Segmentation quality measures averaged over all the images of the testing database.

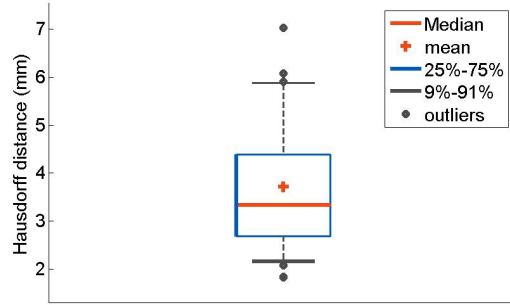| Segmentation quality measure	
DSC (acoustic shadow)	0.88 ± 0.06
DSC (spinous process)	0.72 ± 0.17
Hausdorff distance (mm)	3.90 ± 1.4
Euclidean distance between spinous process centroids (mm)	0.38 ± 0.4

Figure 11: Boxplot of Hausdorff distances between the true and automatically segmented acoustic shadows.

shadow. We obtained a high classification rate, a DSC of 0.88 with a low standard deviation, and a Hausdorff distance of 3.9 mm for the acoustic shadow below the spinous process. However, the standard deviation for the Hausdorff distance is very high. A boxplot representing the distribution of the Hausdorff distances among the test database is shown in Fig. 11. This distribution is skewed towards low Hausdorff distances, implying that the high standard deviation is due to a few inaccurate segmentation results, like those shown in Fig. 12, occurring when there is a region containing little
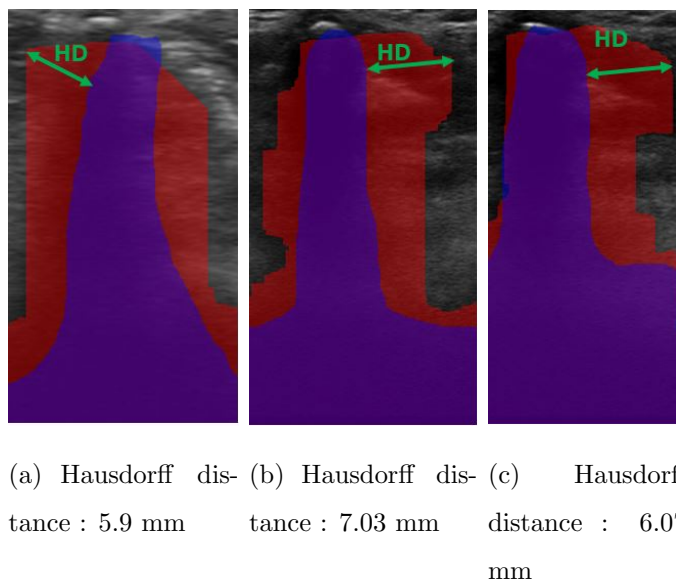
(a) Hausdorff distance : 5.9 mm

(b) Hausdorff distance : 7.03 mm

(c) Hausdorff distance : 6.07 mm

Figure 12: Examples of inaccurate automatic acoustic shadow segmentation with Hausdorff distance measurements (green arrows labeled "HD"). Red: automatic segmentation result. Blue: ground truth segmentation. Purple: overlap of the previous two.

speckle adjacent to the acoustic shadow. The acoustic shadow overextends horizontally over a small region. When the convex hull is computed, the size of this overextension is reduced but it is not entirely removed. To address this problem, the use of a shape template for the acoustic shadow, designed from the input of several experts (Tran and Rohling, 2010) or a statistical atlas (Khallaghi et al., 2010; Hacihaliloglu et al., 2014), could be a solution. Improving the reliability of acoustic shadow segmentation in such a fashion could be useful for constraining the location of key anatomic landmarks such as the laminae, which in turn would be useful for obtaining automated assessments of scoliosis using the Center of Laminae method (Nguyen et al., 2015).

Notice from Table 3 that the classification rate is higher for the "other

tissues" and lower for the spinous process after regularization than before regularization (i.e., than the classification rates corresponding to the results reported in Table 2). The regularization step selects a single connected component to represent the spinous process. This improves the the specificity of the segmentation (hence the improvement in the classification rate for "other tissues"). However, this also slightly reduces sensitivity. As a matter of fact, in our database, there was one case where no spinous process was found at all after regularization.

For the spinous process, we obtained $DSC = 0.72$, indicating under-segmentation, as illustrated in Fig. 13. The visible part of the spinous process is a very small region, so a few mis-classified pixels have a strong impact on the classification rate and the DSC. However, from the point of view of localization accuracy, we also measured a distance of 0.38 ± 0.4 mm between the centroids of the spinous process in the manually and automatically segmented images, demonstrating an accuracy and precision on the order of X-ray image resolution. This implies that measurements of the Spinous Process Angle, recommended by Li et al. (2012) as a proxy to the Cobb angle, would be more accurate using our automatic method on ultrasound images than with manual landmark identification on X-ray images, currently used in clinical practice.

4. Conclusions

In this paper, we proposed an automatic method for the segmentation of vertebral ultrasound images in the transverse view. This method automatically and simultaneously segments the spinous process and its acoustic
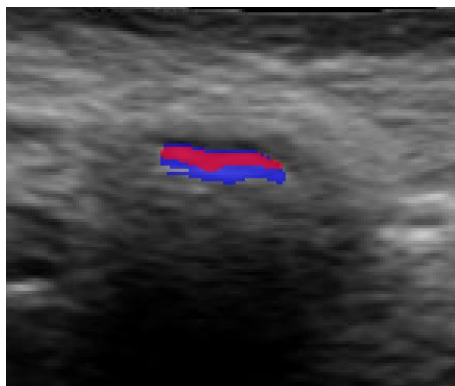
Figure 13: Example of an under-segmented spinous process. Red: manual segmentation, blue: automatic segmentation.

shadow in the image. It is based on a set of features derived and adapted from the literature, and designed to capture the characteristics of bone and shadow in ultrasound images, and their contrast with respect to the background textures. A LDA classifier is used to classify each pixel in three classes : "spinous process", "acoustic shadow" and "other tissues". Then, a regularization step is applied to the individually classified pixels to account for some of the geometric properties of vertebral images. The feature set was validated using established feature selection techniques. The segmentation method was then tested on a large database, and we obtained classification rates around 84% for the spinous process, 92% for the acoustic shadow and 91% for "other tissues". Furthermore, we obtained Dice similarity coefficients of 0.88 for the acoustic shadow and 0.72 for the spinous process, and the centroid of the spinous process was at an average distance of 0.38 mm from the manually segmented spinous process. The high accuracy of spinous process localization using our method is one very encouraging result from a

clinical standpoint, which we hope to leverage in future work involving measurements of the Spinous Process Angle on 3D ultrasound data. To improve accuracy further, our automatic segmentation method can easily be extended to utilize new image features. For instance, Hacihaliloglu et al.'s phase symmetry features (Hacihaliloglu et al., 2009) might advantageously be replaced by their more recently developed local phase tensor features (Hacihaliloglu et al., 2014). Future work will also involve optimizing the image acquisition parameters and protocol in order to capture other anatomical structures like the transverse processes, the laminae and the vertebral bodies, which will be integrated as new classes in the segmentation algorithm. The method will also be validated on images from scoliotic patients so that its use may be investigated to assist dense 3D reconstructions of the scoliotic spine as well as intra-operative surgical guidance.

Acknowledgments

References

Ashab, H. A.-D., Lessoway, V. A., Khallaghi, S., Cheng, A., Rohling, R., Abolmaesumi, P., 2013. An augmented reality system for epidural anesthesia (AREA) prepuncture identification of vertebrae. IEEE Transactions on Biomedical Engineering 60 (9), 2636–2644.

Baumann, M., Mozer, P., Daanen, V., Troccaz, J., 2009. Transrectal ultrasound prostate biopsy tracking with efficient and accurate deformation estimation. In: IEEE International Symposium on Biomedical Imaging: From Nano to Macro. pp. 1211–1214.

Berton, F., Azzabi, W., Cheriet, F., Laporte, C., 2015. Automatic segmentation of vertebrae in ultrasound images. In: International Conference on Image Analysis and Recognition. pp. 344–351.

Breiman, L., 2001. Random forests. Machine Learning 45 (1), 5–32.

Cheung, C., Zhou, G., Law, S., Mak, T., Lai, K., Zheng, Y., 2015. Ultrasound volume projection imaging for assessment of scoliosis. IEEE Transactions on Medical Imaging 34 (8), 1760 –1768.

Clausi, D. A., Jernigan, M. E., 2000. Designing Gabor filters for optimal texture separability. Pattern Recognition 33 (11), 1835–1849.

Coupé, P., Hellier, P., Morandi, X., Barillot, C., 2012. 3D rigid registration of intraoperative ultrasound and preoperative MR brain images based on hyperechogenic structures. Journal of Biomedical Imaging 2012 (531319).

Daanen, V., Tonetti, J., Troccaz, J., 2004. A fully automated method for the delineation of osseous interface in ultrasound images. In: International Conference Medical Image Computing and Computer Assisted Intervention. pp. 549–557.

Duda, R. O., Hart, P. E., Stork, D. G., 2000. Pattern Classification. Vol. Second Edition. John Wiley & Sons Inc.

Field, D. J., 1987. Relations between the statistics of natural images and the response properties of cortical cells. Journal of the Acoustic Society of America A 4 (12), 2379–2394.

Foroughi, P., Boctor, E., Swartz, M., Taylor, R., Fichtinger, G., 2007. Ultrasound bone segmentation using dynamic programming. In: IEEE Ultrasonics Symposium. pp. 2523–2526.

Guo, Z., Zhang, L., Zhang, D., 2010. A complete modeling of local binary pattern operator for texture classification. IEEE Transactions on Image Processing 19 (6), 1657–1663.

Hacihaliloglu, I., Abugharbieh, R., Hodgson, A., Rohling, R., 2009. Bone surface localization in ultrasound using image phase-based features. Ultrasound in Medicine & Biology 35 (9), 1475–1487.

Hacihaliloglu, I., Rasoulian, A., Rohling, R. N., Abolmaesumi, P., 2014. Local phase tensor features for 3-D ultrasound to statistical shape + pose spine model registration. IEEE Transactions on Medical Imaging 33 (11), 2167–2179.

Hellier, P., Coupé, P., Morandi, X., Collins, D. L., 2010. An automatic geometrical and statistical method to detect acoustic shadows in intraoperative ultrasound brain images. Medical Image Analysis 14, 195–204.

Karamalis, A., Wein, W., Klein, T., Navab, N., 2012. Ultrasound confidence maps using random walks. Medical Image Analysis 16, 1101–1112.

Kerby, B., Rholing, R., Nair, V., Abolmaesumi, P., 2008. Automatic identification of lumbar level with ultrasound. In: Annual Meeting of the IEEE Engineering in Medicine and Biology Society. pp. 2980–2983.

Khallaghi, S., Mousavi, P., Gong, R. H., Gill, S., Boisvert, J., Fichtinger, G., Pichora, D., Borschneck, D., Abolmaesumi, P., 2010. Registration of a statistical shape model of the lumbar spine to 3D ultrasound images. In: International Conference on Medical Image Computing and Computer Assisted Intervention. Vol. II. pp. 68–75.

Li, M., Cheng, J., Ying, M., Ng, B., Zheng, Y. P., Lam, T., Wong, W. Y., Wong, M., 2012. Could clinical ultrasound improve the fitting of spinal orthosis for the patients with AIS? European Spine Journal 21 (10), 1926–1935.

Lou, E. H., Chan, A. C., Donauer, A., Tilburn, M., Hill, D. L., 2015. Ultrasound-assisted brace casting for adolescent idiopathic scoliosis. Scoliosis 10 (Suppl. 1), O38.

McLachlan, G., 2004. Discriminant Analysis and Statistical Pattern Recognition. Vol. 544. John Wiley & Sons.

Nguyen, D. V., Vo, Q. N., Le, L. H., Lou, E. H. M., 2015. Validation of 3D surface reconstruction of vertebrae and spinal column using 3D ultrasound data–a pilot study. Medical Engineering & Physics 37 (2), 239–244.

Pudil, P., Novovičová, J., Kittler, J., 1994. Floating search methods in feature selection. Pattern Recognition Letters 15 (11), 1119–1125.

Sato, Y., Nakamoto, M., Tamaki, Y., Sasama, T., Sakita, I., Nakajima, Y., Monden, M., Tamura, S., 1998. Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization. IEEE Transactions on Medical Imaging 17 (5), 681–693.

Tran, D., Rohling, R. N., 2010. Automatic detection of lumbar anatomy in ultrasound images of human subjects. IEEE Transactions on Biomedical Engineering 57 (9), 2248–2256.

Ungi, T., King, F., Kempston, M., Keri, Z., Lasso, A., Mousavi, P., Rudan, J., Borschneck, D. P., Fichtinger, G., 2014. Spinal curvature measurement by tracked ultrasound snapshots. Ultrasound in Medicine & Biology 40 (2), 447–454.

Unsgaard, G., Rygh, O., Selbekk, T., Müller, T. B., Kolstad, F., Lindseth, F., Nagelhus Hernes, T., 2006. Intra-operative 3D ultrasound in neurosurgery. Acta Neurochirurgica 148 (3), 235–253.

Wu, R., Ling, K. V., Shao, W., Ng, W. S., 2003. Registration of organ surface with intra-operative 3D ultrasound image using genetic algorithm. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 383–390.

Young, M., Hill, D. L., Zheng, R., Lou, E., 2015. Reliability and accuracy of ultrasound measurements with and without the aid of previous radiographs in adolescent idiopathic scoliosis (AIS). European Spine Journal 24 (7), 1427–1433.

Yu, S., Tan, K. K., 2014. Classification of lumbar ultrasound images with machine learning. In: Simulated Evolution and Learning. pp. 287–298.

Yu, S., Tan, K. K., Shen, C., Sia, A. T. H., 2013. Ultrasound guided automatic localization of needle insertion site for epidural anesthesia. In: IEEE International Conference on Mechatronics and Automation. pp. 985–990.

Yu, S., Tan, K. K., Sng, B. L., Li, S., Sia, A. T. H., 2014. Automatic identification of needle insertion site in epidural anesthesia with a cascading classifier. Ultrasound in Medicine & Biology 40 (9), 1980–1990.

Yu, S., Tan, K. K., Sng, B. L., Li, S., Sia, A. T. H., 2015. Lumbar ultrasound image feature extraction and classification with support vector machine. Ultrasound in Medicine & Biology 41 (10), 2677–2689.

Zhang, S., Jiang, S., Yang, Z., Liu, R., 2015. 2D ultrasound and 3D MR image registration of the prostate for brachytherapy surgical navigation. Medicine 94 (40), e1643.

Zhang, X., Günther, M., Bongers, A., 2010. Real-time organ tracking in ultrasound imaging using active contours and conditional density propagation. In: International Workshop on Medical Imaging and Augmented Reality. pp. 286–294.