Segmenting bone structures in ultrasound images with Locally Weighted SLSC (LW-SLSC) beamforming

Eduardo Gonzalez*and Muyinatu A. Lediju Bell*†‡
*Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD
†Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD
‡Department of Computer Science, Johns Hopkins University, Baltimore, MD

***Abstract*—The process of registering ultrasound (US) images to computed tomography (CT) images relies on accurate segmentation of bony structures in US images. However, segmentation of US images often suffers from the presence of speckle noise, clutter, and acoustic shadowing. We propose to improve the US bone segmentation process with a novel Locally Weighted SLSC (LW-SLSC) beamforming method, which is based on the minimization of the total variation of a spatial coherence weighted sum. Application of this beamformer to an *ex vivo* human vertebra resulted in a 911% contrast-to-noise ratio (CNR) increase in LW-SLSC images (CNR=23.66) when compared to traditional delay-and-sum (DAS) images (CNR=2.34) created from the same channel data. Application to an *ex vivo* caprine vertebra with surrounding tissue intact similarly resulted in a 55.8% CNR increase in the LW-SLSC images (CNR=2.01) compared to DAS images (CNR=1.29) created from the same channel data. Bone boundaries in the caprine vertebra were segmented from the US and CT images, and the LW-SLSC beamformer enabled approximately 5.5 mm thinner boundary lines than the DAS beamformer when compared to segmentation results based on CT images. Similarly, the location error of boundary lines was also reduced with 70% of the total spatial error within $\pm$1 mm in LW-SLSC images compared to 47% in DAS images. These results demonstrate that LW-SLSC imaging provides improved bone segmentation over traditional DAS imaging, which has promising implications for real-time segmentation of bone boundaries during spinal fusion surgeries and other procedures that may benefit from accurate US-based bone segmentation.**

I. INTRODUCTION

In spinal fusion surgeries, x-ray fluoroscopy is frequently used and correlated with preoperative computed tomography (CT) images to reduce the risk of complications related to pedicle screw misplacement [1]. Ultrasound (US) registration to preoperative CT is one alternative to the ionizing radiation of x-rays [2]. However, in order to address the noise and artifacts associated with US images, additional data and filtering are typically required. For example, Wein *et al.* designed a biomechanical model for simulated US backscattered signals from CT volumetric data of soft tissues [3], which was later registered to intraoperative volumetric US data from the liver and kidney [4]. Gill *et al.* [5] simulated US volumes from 3D CT spine data in order to segment the posterior surface of the vertebra and registered the data to *ex vivo* sheep samples. Similarly, Brendel *et al.* [6] simulated US volumes from 3D CT spine data and registered the simulated US data to real US data of an *ex vivo* human spine. Filters are additionally utilized in sophisticated registration procedures based on multicomponent similarity measurements [7]. These measurements require a combination of weighted mutual information term, edge correlation, clamping the compressed skin surface, and shadowing-related artifact removal to assess the alignment between structures in US images and structures in CT images [7].

Based on these examples, it is clear that the CT-US registration approaches for spinal fusion surgery require well-defined bone structures in order to achieve accurate segmentation in real time, and this requirement is currently achieved with redundant information provided by 3D US volumetric data. However, 3D segmentation is often computationally expensive and requires custom hardware for feasible implementation. In addition, depending on the angle of incidence of the US beams to the vertebra, an adapted 3D CT surface must be simulated in real time as well [2]. We hypothesize that improving the bone boundary visualization in ultrasound images with advanced beamforming techniques will benefit surgeons with segmenting and registering 2D US images and reduce the computational burden associated with requiring 3D images to complete a segmentation task.

To explore this hypothesis, Short-Lag Spatial Coherence (SLSC) [8] beamforming and a novel variation of SLSC beamforming that we developed — i.e., Locally Weighted SLSC (LW-SLSC) — were used as alternatives to conventional delay-and-sum (DAS) beamforming to reduce speckle noise and enhance bone boundaries. US and CT acquisitions were conducted on an *ex vivo* human lumbar vertebra and an *ex vivo* caprine thoracic section. Segmentation results in US images created with the LW-SLSC beamformer were quantitatively compared to segmentation results from co-registered CT images.

II. METHODS

A. Short Lag Spatial Coherence (SLSC)

SLSC beamforming displays the similarity of received signals in the aperture as a function of element separation m. A received time-delayed sample in the aperture is represented as $s_i(n)$, where i is the channel index and n the depth index in a zero-mean radio frequency signal s_i. First, the coherence function $R(m)$ is calculated using a small axial kernel as defined by:

$$\hat{R}(m) = \frac{1}{N-m} \sum_{i=1}^{N-m} \frac{\sum_{n=n_1}^{n_2} s_i(n) s_{i+m}(n)}{\sqrt{\sum_{n=n_1}^{n_2} s_i^2(n) \sum_{n=n_1}^{n_2} s_{i+m}^2(n)}}, \tag{1}$$

where N is the the number of elements in the aperture, and n_1 and n_2 are the limits of the axial kernel. Then, an SLSC image of lag M is generated as the integral of the spatial coherence function over the first M lags:

$$\text{SLSC}(M) = \int_1^M \hat{R}(m)dm \approx \sum_{m=1}^{M} \hat{R}(m). \quad (2)$$

SLSC images were computed with lag M = 8 and with an axial kernel size equivalent to 1.2λ, where λ is the wavelength associated with the center frequency of the transducer.

B. Locally Weighted Short Lag Spatial Coherence (LW-SLSC)

We hypothesize that additional enhancement of bone boundaries can be achieved by implementing a regularized version of the SLSC beamforming. Instead of averaging the cumulative sum up to a lag value M (out of a preselected total of N_L lags, where M> N_L), LW-SLSC beamforming computes the weighted coefficients for N_L lags by minimizing the total variation (TV) of the weighted sum within a moving kernel of size k_z x k_x x N_L. In order to preserve the high resolution information available at higher lags (i.e. M>15), this adaptive solution is regularized using the L2-norm with a gradient operator, defined as:

$$\hat{w_i} = \text{argmin}\{\text{TV}(w_i K_i) + \alpha \|\nabla w_i\|^2\}, \quad (3)$$

where TV is the total variation function applied to the cost function, K_i is the kernel i of the correlation matrix $\hat{R}$, and w_i (1 x N_L) is the optimized weight vector for the calculated summed lags of K_i. The main advantage of the LW-SLSC relies on the adaptive selection of lower lags in kernels surrounding isoechoic regions and higher lags otherwise, reducing noise commonly present in high-lag SLSC images. Factors that contribute to the quality of the LW-SLSC are mainly the kernel size (k_z x k_x), overlap, and the regularization term α.

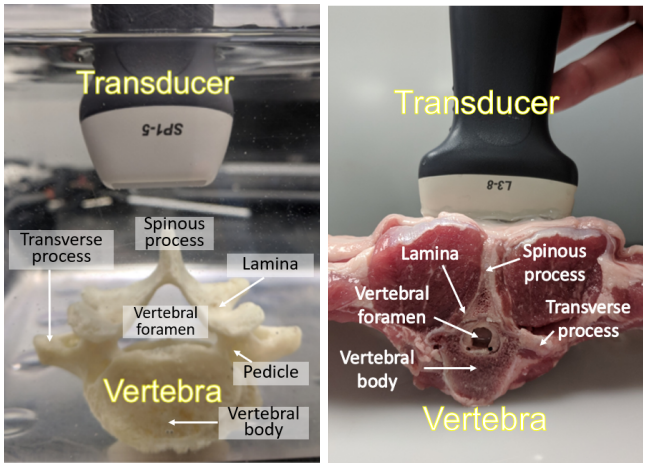
LW-SLSC images were computed with a 1.20 mm (lateral) x 1.92 mm (axial) kernel, 50% overlap and N_L = 50. For the regularization coefficient selection ($\alpha = 0.1$), a conventional L-curve method was applied over a set of vertebra results to find the optimized value [9].

C. Data Acquisition and Imaging Parameters

Raw ultrasound channel data were acquired from two vertebra samples using an Alpinion ECUBE-12R system. The first sample was a human lumbar vertebra, which was submerged in a water tank and imaged with a SP1-5 phased array probe, as shown in Fig. 1(a). The phased array had 64 elements, 0.3 mm pitch, and 3.8 MHz center frequency. The image depth was 65 mm, and the focus was located at a depth of 50 mm.

The second sample was a thoracic horizontal section of a caprine vertebra with surrounding tissue intact. This sample was imaged with a L3-8 linear array probe, as shown in Fig. 1(b). The linear array had 128 elements, 0.3 mm pitch, and 4.0 MHz center frequency. The image depth was 40 mm, and the focus was located at a depth of 30 mm.

CT acquisitions were performed using a SIEMENS ARCADIS Orbic 3D C-Arm with 190 raw projections, generating a 6cm^3 volume of 0.23 mm^3 voxel resolution. For quantitative evaluation of the reconstructed images, contrast-to-noise ratio (CNR) was computed from rectangular regions in the vertebral foramen (F) and lamina (L) with the following equation:



(a) (b)

Fig. 1. US acquisition setup. (a) Human vertebra with a phased array probe. (b) Caprine thoracic horizontal section with a linear array probe.

$$\text{CNR} = \frac{|\mu_F - \mu_L|}{\sqrt{\sigma_F^2 + \sigma_L^2}}, \quad (4)$$

where μ and σ^2 are the mean and variance, respectively, and the subscripts represent the ROIs associated with the two regions described above.

III. RESULTS & DISCUSSION

A. Ex Vivo Human Vertebra

Examples of CT, DAS, SLSC and LW-SLSC images of the human vertebra are shown in Fig. 2. DAS images were

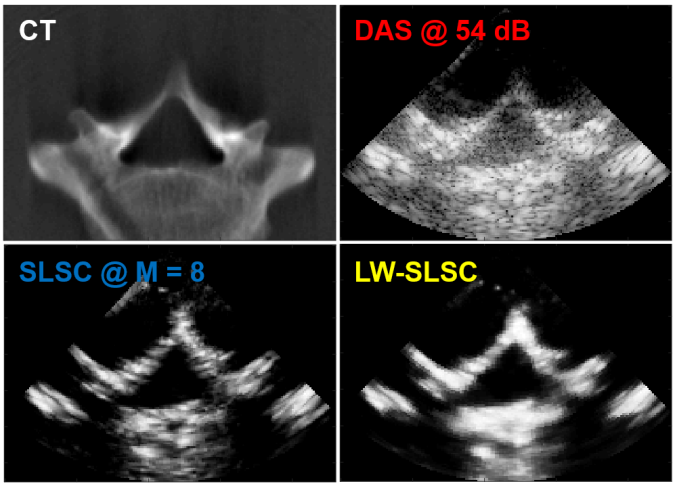


Fig. 2. Examples of reconstructed CT, DAS, SLSC and LW-SLSC images (65 mm axial x 90 mm lateral) of the human vertebra (non-registered). Observe that both SLSC and LW-SLSC have similar lateral resolution and the LW-SLSC image has greater CNR.

created with a rectangular window apodization and a dynamic range of -54 dB. Qualitatively, both SLSC and LW-SLSC beamforming remove clutter artifacts that are apparent in in water region of the DAS image, resulting in the enhanced contrast of bone structures. The CNR of the SLSC image was 11.48, outperforming that of the DAS image, which was 2.34. The lateral and axial resolution is also preserved in the SLSC and LW-SLSC images when considering the similarity between bone structures in these images and the CT and US-based images. Similarly, an additional contrast enhancement was observed in the LW-SLSC image, which was 23.66.

DAS beamforming is commonly used for US-guided surgery of the spine [2], but it is suboptimal when used to delineate bony structures due to acoustic clutter and reflections from bone. For the human spine submerged in water, reducing the dynamic range partially reduces these undesirable artifacts, but also eliminates part of the vertebral body and spinous process. Similarly, SLSC images created with M>15 remove reflection artifacts without affecting the vertebral body and spinous process, but causes discontinuity in the bone structure that can possibly be mistaken for a bone fracture or other abnormalities. On the other hand, SLSC beamforming with M<4 overcomes this discontinuous appearance of otherwise continuous bony structures, but reduces the lateral resolution and accuracy needed to distinguish the pedicle from the transverse process. The texture of the bony structure is notably improved in LW-SLSC images, balancing the trade off between a high-contrast bone boundary and sufficient spatial resolution.

For some US acquisition angles in the human vertebra, vertebral components such as the apex of the spinous process and ventral limit of the vertebral foramen that were not present in the current elevation plane were observed in SLSC and LW-SLSC images. A possible explanation is that projections of backscattered radiofrequency signals from other elevation planes contribute to the computation of the coherence matrix $\hat{R}$. While this could benefit surgeons with more landmarks for segmentation, it also limits accuracy in the elevation dimension.

B. Ex Vivo Caprine Vertebra

Examples of CT, DAS, SLSC and LW-SLSC images of the caprine thoracic section are shown in Fig. 3. In comparison to the human vertebra submerged in water, SLSC and LW-SLSC improve the boundary between soft tissue and the spinous, lamina and transverse process. This bone boundary was segmented in the CT image and the segmentation is overlaid as a white outline on the CT image in Fig. 3. CNR was enhanced in the SLSC and LW-SLSC images, which were 1.30 and 2.01, respectively, when compared to that of the DAS image, which was 1.29.

The improvement in visualization of the bone boundary is expected to provide an improvement in the accuracy of bone segmentation from ultrasound images. Fig. 4 shows the registration of bone boundaries segmented from LW-SLSC and DAS images with the bone boundary segmented from the CT image. Rigid monomodal registration was conducted using a regular step gradient descent optimizer. These results demonstrate bone boundaries segmented from the US images are well registered to the bone boundary segmented from the CT image. However, the shape and thickness of the segmented boundary in the LW-SLSC image is more similar to that in the CT image when compared to the bone boundary segmented from the DAS image, given the same threshold segmentation value for each ultrasound image (i.e., 30% of the maximum normalized intensity).

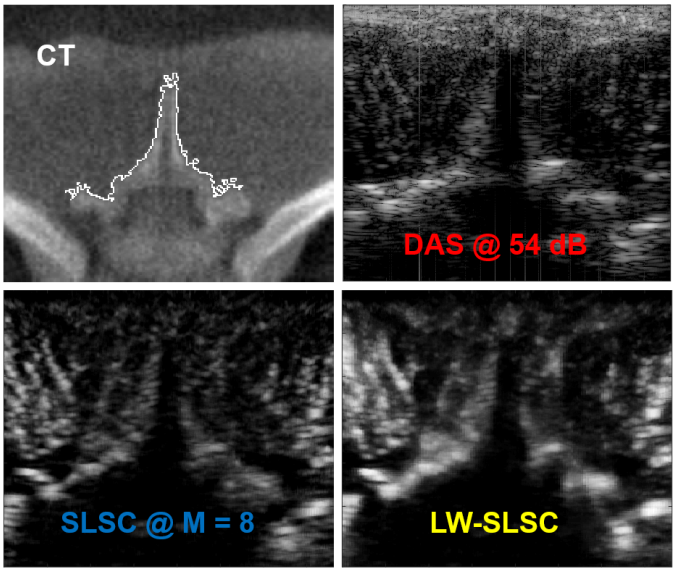


Fig. 3. Examples of reconstructed CT (46.8 mm axial x 45.4 mm lateral), DAS, SLSC and LW-SLSC images (40 mm axial x 38.4 mm lateral) of the caprine sample (non-registered).

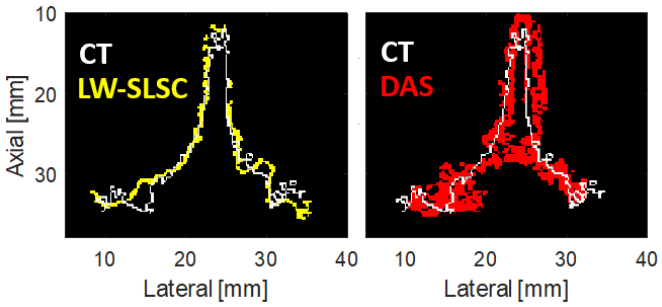


Fig. 4. Registered US-CT bone boundaries after applying threshold segmentation to images of the *ex vivo* caprine vertebra. The US images were beamformed using DAS (left) and LW-SLSC (right).

The integrated thickness of the segmented bone boundary is shown in Fig. 5, which was calculated by summing pixels of the binary masks along the lateral and axial dimension. The average integrated thickness for CT, LW-SLSC, and DAS results were 2.07 mm, 2.19 mm and 7.91 mm, respectively. Fig. 6(a) shows the integrated thickness difference when the CT results are compared to DAS and LW-SLSC results. These results are consistent with the thicker segmented bone boundary derived from the DAS image and the thinner segmented

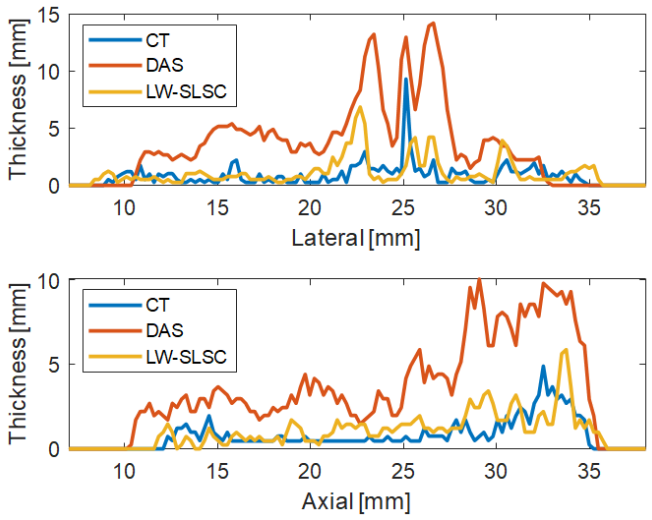


Fig. 5. Integrated thickness of the segmented bone boundary in DAS, LW-SLSC and CT image along the lateral (top) and axial (bottom) dimensions.

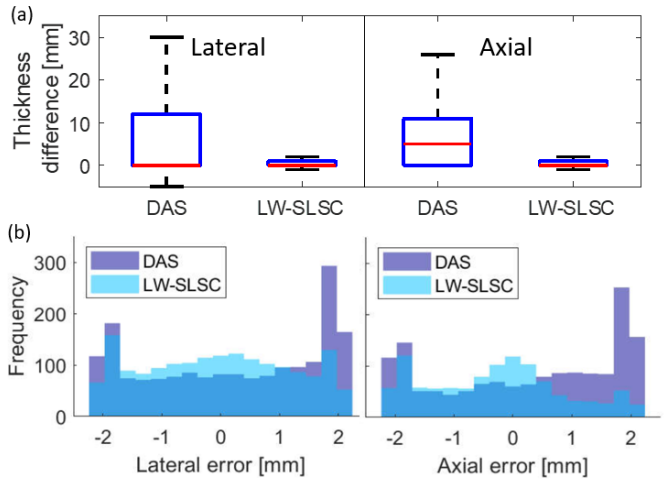


Fig. 6. (a) Integrated thickness difference between segmented bone boundaries in ultrasound and CT images. (b) Histograms of distance errors in the lateral and axial dimensions.

bone boundary derived from the LW-SLSC image

An alternative metric to assess segmentation accuracy is to measure the error between the segmented boundaries in the lateral and axial dimensions within a 4.7 mm x 4.7 mm moving kernel. The histogram of Fig. 6(b) shows this result, demonstrating lower errors with LW-SLSC images than DAS images. The percentage of total pixels within $\pm$ 1 mm error was 70% with the LW-SLSC image and 47% with the DAS image.

The thinner segmented bone boundary and lower errors achieved with LW-SLSC is expected to decrease uncertainty in the real-time identification of bone boundaries in ultrasound images, which has implications for improving targeting accuracy during pedicle screw insertion. In addition, the improved edges in the bone boundary could enable real-time segmentation during spinal fusion procedures without requiring 3D images. This work is also promising for combining ultrasound image segmentation with novel photoacoustic-guided spinal fusion surgery approaches [10].

IV. CONCLUSION

To the authors' knowledge, this work is the first to demonstrate that coherence-based beamforming can be used to highlight bone boundaries for US-based segmentation tasks. In particular, our novel LW-SLSC beamforming method segments bone structures in ultrasound images with a striking similarity to segmentation results achieved by CT images. This is a significant improvement over the bone segmentation results that were achieved with traditional DAS beamforming. These results have implications for enabling real-time, ultrasound-based segmentation of bone boundaries during spinal fusion surgeries and other procedures that may benefit from accurate bone boundary segmentation. Future studies include additional validation with human cadaver experiments as well as evaluating the feasibility of LW-SLSC in clinical applications.

ACKNOWLEDGMENT

The authors acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. In addition, the authors thank Gerhard Kleinzig and Sebastian Vogt from Siemens Healthineers for making a Siemens ARCADIS Orbic 3D available.

REFERENCES

[1] R. M. Kretzer, C. Chaput, D. M. Sciubba, I. M. Garonzik, G. I. Jallo, P. C. McAfee, B. W. Cunningham, and P. J. Tortolani, "A computed tomography–based morphometric study of thoracic pedicle anatomy in a random united states trauma population," *Journal of Neurosurgery: Spine*, vol. 14, no. 2, pp. 235–243, 2011.

[2] S. Winter, B. Brendel, I. Pechlivanis, K. Schmieder, and C. Igel, "Registration of ct and intraoperative 3-d ultrasound images of the spine using evolutionary and gradient-based methods," *IEEE Transactions on Evolutionary Computation*, vol. 12, no. 3, pp. 284–296, 2008.

[3] W. Wein, A. Khamene, D.-A. Clevert, O. Kutter, and N. Navab, "Simulation and fully automatic multimodal registration of medical ultrasound," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2007, pp. 136–143.

[4] W. Wein, S. Brunke, A. Khamene, M. R. Callstrom, and N. Navab, "Automatic ct-ultrasound registration for diagnostic imaging and image-guided intervention," *Medical image analysis*, vol. 12, no. 5, pp. 577–585, 2008.

[5] S. Gill, P. Abolmaesumi, G. Fichtinger, J. Boisvert, D. Pichora, D. Borshneck, and P. Mousavi, "Biomechanically constrained groupwise ultrasound to ct registration of the lumbar spine," *Medical image analysis*, vol. 16, no. 3, pp. 662–674, 2012.

[6] B. Brendel, S. W. A. Rick, M. Stockheim, and H. Ermert, "Registration of 3d ct and ultrasound datasets of the spine using bone structures," *Computer Aided Surgery*, vol. 7, no. 3, pp. 146–155, 2002.

[7] W. Wein, B. Roper, and N. Navab, "Integrating diagnostic b-mode ultrasonography into ct-based radiation treatment planning," *IEEE transactions on medical imaging*, vol. 26, no. 6, pp. 866–879, 2007.

[8] M. A. Lediju, G. E. Trahey, B. C. Byram, and J. J. Dahl, "Short-lag spatial coherence of backscattered echoes: Imaging characteristics," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 58, no. 7, 2011.

[9] P. C. Hansen and D. P. OLeary, "The use of the l-curve in the regularization of discrete ill-posed problems," *SIAM Journal on Scientific Computing*, vol. 14, no. 6, pp. 1487–1503, 1993.

[10] J. Shubert and M. A. L. Bell, "Photoacoustic imaging of a human vertebra: implications for guiding spinal fusion surgeries," *Physics in Medicine and Biology*, 2018.