SAM Zero-Shot detection of vertebral surfaces from freehand 3D ultrasound

Sen Li[1], Farida Cheriet[2] and Catherine Laporte[1]

[1] École de technologie supérieure, 1100 Rue Notre-Dame O., Montréal, Canada
[2] Polytechnique Montréal, 2900 boul. Édouard-Montpetit, Montréal, Canada

Abstract— **Freehand 3D ultrasound imaging has emerged as a promising modality for spine imaging due to its non-invasive nature and cost-effectiveness. Accurate extraction of the 3D bone surface from freehand ultrasound image sequences is crucial for comprehensive interpretation of spinal structures, geometric spine deformity analysis, and for guiding ultrasound-assisted spine interventions. However, the vertical bone surface adjacent to the spinous process is usually invisible in ultrasound scans. In this preliminary study, a novel method for tissue-bone separation utilizing landmark-guided Segment Anything Model (SAM) is proposed to achieve accurate 3D bone surface reconstruction from low-quality ultrasound image sequences. The resulting bone surface, represented as a 3D mesh, effectively reproduces a smooth outline of the vertebra, encompassing the left lamina, spinous process, and right lamina. To evaluate the performance of the proposed method, we collected three freehand ultrasound sequences from three participants. The acoustic shadow masks beneath the extracted bone surface were evaluated against manually labeled masks, achieving a mean Intersection over Union (IoU) above 0.92.**

Keywords— **SAM Zero-Shot, Freehand 3D Ultrasound, Vertebral Surface Detection, 3D Point Cloud**

I. INTRODUCTION

In recent years, freehand 3D ultrasound has emerged as a promising modality for 3D spine imaging, offering advantages such as non-invasiveness, portability, real-time visualization, and cost-effectiveness. The extraction of vertebral landmarks from 3D ultrasound sequences provides comprehensive information on the 3D spine shape and greatly facilitates anatomical structure-oriented investigations across diverse applications using ultrasound imaging, including epidural anesthesia [1], intra-operative image guidance [2, 3], assessment of the severity of spine deformities [4], and more specifically scoliosis assessment [5, 6].

The 3D reconstruction of the vertebral surface holds significant clinical significance in terms of diagnosing and prognosticating spine deformities as well as providing guidance during intra-operative imaging procedures. Ungi et al. [7] proposed to use U-Net to segment the bone surface in paramedian sagittal ultrasound images for the 3D reconstruction of left/right transverse processes (TP). The spine shape represented by paired TPs in 3D provided assistance in visualization and scoliosis measurement. Tang et al. [8] introduced a U-Net-based approach for extracting the bone surface to reconstruct a phantom lumbar spine, where their method produces non-contiguous masks encompassing all discernible bone surface echo responses, including the anterior complex. Jiang et al. [9] proposed a manual contour-tracing based method for 3D spinal surface reconstruction, where the contour line effectively connects the left lamina, spinous process, and the right lamina, allowing for a potentially accurate representation of individual vertebra outlines in the reconstructed 3D spinal surface. However, this approach is labor-intensive.

Automated extraction of the spinous process surface presents a significant challenge in spinal ultrasound imaging due to the inherently poor contrasts and adverse bone geometry in these images. Yet, a precise 3D spine model is essential for comprehending spinal structure, with the spinous process providing key 3D positional information and facilitating vertebral rotation analysis, particularly in the context of ultrasound-guided spine interventions. Accurate online 3D spinal surface reconstruction from ultrasound sequences could facilitate interpretation of vertebral structure in ultrasound images, streamline sonographer training, and ultimately optimize the quality of acquired ultrasound sequences.

The present work aims to develop an automated method to extract a smooth vertebral bone surface, including the spinous process, from freehand 3D ultrasound data. In this preliminary study, we investigate the application of the Segment Anything Model (SAM) — specifically SAM Zero-Shot, which has shown potential for other medical imaging applications [10, 11] — for segmenting vertebral surfaces in ultrasound images obtained from a spine phantom and human subjects. We compare the performance of SAM's automated mask generation (denoted as "SAM Auto") and prompt-based segmentation methods. Additionally, we propose a novel vertebral bone surface extraction method, **L**andmark prompted **S**AM - **T**rio **S**urface **E**xtraction (LS-TSE), specifically designed for transverse ultrasound images. The LS-TSE method combines the advantages of AI-generated laminae landmarks for precise targeting of the left lamina, right lamina, and spinous process, pre-trained SAM for controlled segmentation of ultrasound response from left/right laminae surface using prompts, and asymmetrical peak model for spinous process bone surface approximation with a trio fitting design based on the extracted curves of the left and right lamina bone surfaces.

II. MATERIALS AND METHODS

To evaluate the potential of SAM Zero-Shot, both with and without prompts, our experimental setup involves the acquisition of freehand 3D ultrasound data from both spine phantoms and human subjects, as detailed in section A. The proposed LS-TSE method designed to handle low-quality ultrasound data is elaborated on in section B. The efficacy of LS-TSE in generating acoustic shadow (area below its extracted vertebral bone surface) masks is assessed against manually labeled masks on frames where laminae are visible, using the Intersection over Union (IoU) metric. A B-Spline surface is a mathematical representation employed in computer graphics for modeling smooth surfaces through the manipulation of control points. The accuracy of vertebral (B-Spline) surface extraction using LS-TSE is evaluated using the Mean Surface Distance (MSD) metric. The upper boundary of the manually labeled acoustic shadow mask was considered as the true bone surface, and 65 uniformly distributed control points along these curves were employed to generate a B-Spline 3D mesh.

A. Freehand 3D Ultrasound Data Acquisition

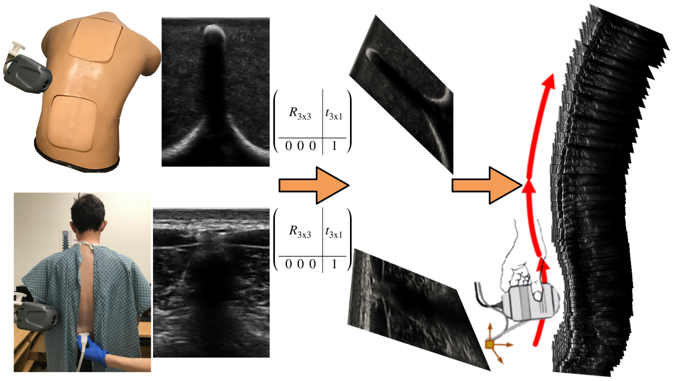

Fig. 1: Freehand 3D ultrasound imaging of the spine

An Ultrasonix Sonix Tablet equipped with a linear probe and an integrated magnetic position sensor (L14/5 GPS) was employed to acquire ultrasound sequences from a spine phantom (Blue Phantom, Kirkland, WA) and three healthy adults in the standing position, yielding high-quality and low-quality images respectively. The in-vivo image quality can be influenced by factors, e.g., body mass index, spine shape, probe operation, etc. The sonographer manipulated the ultrasound probe in an inferior-to-superior direction along each subject's spine, maintaining the probe orientation orthogonal to the skin surface. As shown in Fig. 1, freehand 3D ultrasound captures ultrasound data sequences with each frame comprising a distinct 2D image and its corresponding 4 by 4 transformation matrix obtained from the integrated magnetic position sensor. The matrices map the images from the transducer to the world coordinate system, enabling the rendering of a 3D spinal ultrasound sequence in world coordinates. The collected data were imported from the Sonix Tablet to a desktop computer utilizing the PLUS toolkit [12], integrated within the 3D Slicer software. In total, three ultrasound sequences were acquired from three volunteered participants (one sequence per participant). In these images, 508 pairs of lamina landmarks were identified and localized through manual labeling, while 531 pairs were identified using the deep learning approach [13].

B. LS-TSE vertebral surface extraction method

The **T**rio **S**urface **E**xtraction (TSE) determines the transverse vertebral surface by joining the left lamina, the spinous process, and the right lamina. The detection of the left/right lamina surfaces is detailed in section B1, and that of the spinous process surface in section B2.

B1 Lamina surface detection

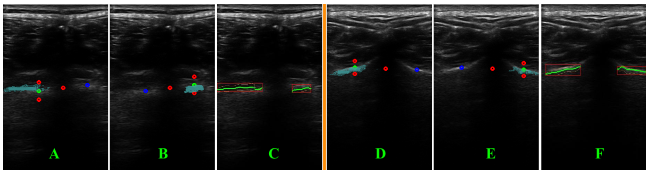

Fig. 2: Laminae bone surface extraction using Lamina landmark prompted SAM Zero-Shot. Images A/B (lumbar) and D/E (thoracic) present the prompt points of SAM Zero-Shot segmentation for left/right laminae areas. Blue circles are laminae landmarks, and red/green circles are background/foreground prompt points. SAM masks were overlapped in semitransparent color cerulean. Images C (lumbar) and F (thoracic) show the extracted bone surface (green curve), in which the red rectangles are bounding boxes of segmented laminae masks.

The lamina landmarks extracted from the ultrasound images are used as point prompts to SAM Zero-Shot towards the extraction of left and right lamina segmentation masks, mask-L and mask-R, respectively illustrated in Fig. 2 A/D and Fig. 2 B/E. To generate each mask, one foreground prompt point (the corresponding lamina landmark) and three background prompt points are employed, including the midpoint between two laminae landmarks and two points located slightly above and below each lamina landmark, at 0.9 and 1.1 times their depths with respect to the skin surface, respectively. The left/right lamina bone surface curve is obtained by tracing the curve that is vertically centered within its respective lamina mask, followed by a vertical translation of this curve to align it with the lamina landmark, thereby ensuring its accuracy.

B2 Spinous process surface estimation

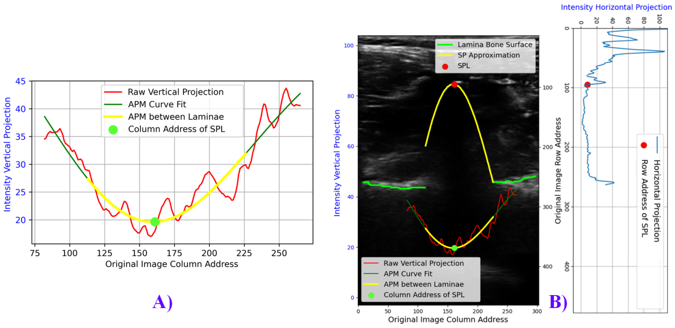

Fig. 3: Spinous process surface estimation.

To model the shape of the spinous process, as approximated by that of its acoustic shadow, we used an Asymmetrical Peak Model (AMP), which can be described as a skewed Gaussian curve. This is a flexible, yet simple model, which can account for complex spine anatomy and variations in the orientation of the ultrasound beam which can cause asymmetry in the apparent shape of the spinous process. The APM is mathematically described as:

$$f(x;A,\mu,\sigma,\alpha) = A \cdot exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \cdot \left(1+\alpha \cdot \frac{x-\mu}{\sigma}\right) \tag{1}$$

where A is the amplitude of the peak, μ is the center of the peak, σ is the standard deviation controlling the width of the peak, and α is the asymmetry parameter, which determines how much the peak is shifted asymmetrically.

The parameters of the APM are determined by fitting the model to the vertical image intensity projection curve using the Levenberg-Marquardt algorithm, as shown in Fig. 3 A. Inspired by Brignol et al. [14], we localized the row address of the spinous process landmark from the horizontal intensity projection to determine the height h of the flipped and stretched APM curve, as shown in panel B. Therefore, the spinous process outline is represented by a vertically translated flipped and stretched APM curve, whose potential shorter "foot" is filled by a vertical line to connect to its corresponding lamina bone surface (as shown in Fig. 5).

III. RESULTS

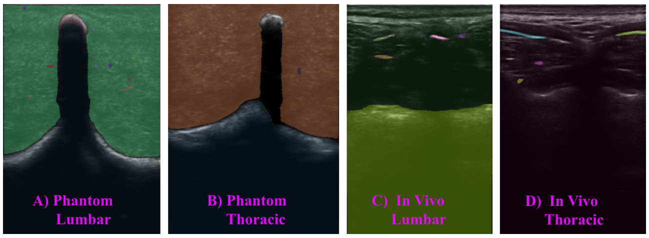

Fig. 4: SAM Zero-Shot with no prompts in Spinal Ultrasound Object Segmentation. Random colors were used for masks segmented by "SAM Auto" method. Ultrasound images in A∼B are scanned from the phantom, and C∼D are scanned from the in-vivo human subject.

Fig. 4 presents the segmentation results of SAM Zero-Shot on spinal ultrasound images. The "SAM Auto" method perfectly delineated the acoustic shadow in phantom ultrasound as shown in panels A and B. However, it identified the acoustic shadow area beneath the laminae as an object in the human lumbar scan (panel C), and recognized no vertebral structure but only edges in the human thoracic scan (panel D).

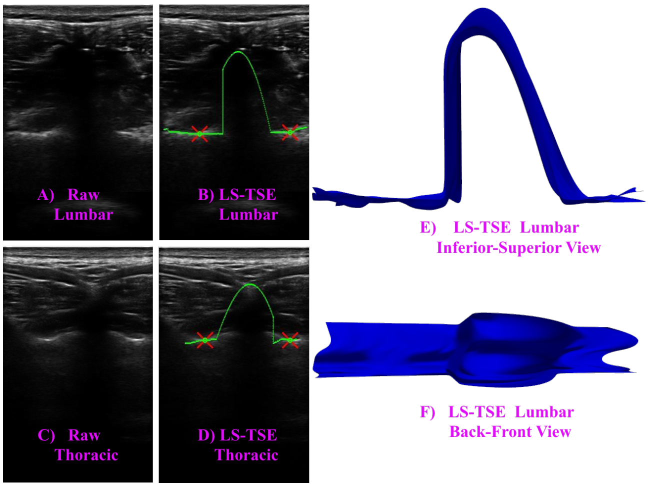

Fig. 5: Vertebral surface extraction from human transverse spinal ultrasound using LS-TSE. In panels A∼D, red crosses denote the lamina landmarks and green curves represent the bone surface. Images E/F present the B-Spline 3D mesh of an extracted lumbar vertebral surface.

Table 1: LS-TSE acoustic mask IoU against manual labeling. PTCP is short for "participant". Method M-L / DL-L refer to lamina **L**andmarks generated by **M**anual labeling and **D**eep **L**earning method [13], respectively.

Method	PTCP #1	PTCP #2	PTCP #3
M-L	0.9326	0.9439	0.9330
DL-L	0.9358	0.9379	0.9243

Fig. 5 A/B and C/D demonstrate bone surface extraction

from individual lumbar and thoracic vertebra images, respectively. Panel B illustrates the ability of LS-TSE to extract the lumbar vertebral surface. Nevertheless, panel D shows a result where the detected spinous process landmark does not belong to the same vertebra level as the lamina bone surface. Instead, the spinous process landmark extracted in panel D actually belongs to the superior vertebral level. Thoracic spinal anatomy complexifies the task due to the physical overlap between consecutive vertebrae, and a surface model derived from these landmarks would not correctly represent the underlying anatomy. Consequently, in this preliminary study, our subsequent evaluation of the LS-TSE method focuses on the lumbar region.

Table 1 presents the assessment of LS-TSE in segmenting acoustic shadow masks under the lumbar vertebral bone surface. The IoU results suggest a high accuracy regardless of the method used to generate the lamina landmark prompts.

Figs. 5 E and F show the 3D visualization of an extracted lumbar vertebral surface obtained from participant #1. The three lumbar vertebrae B-Spline meshes extracted by LS-TSE from participant #1 achieve a MSD error of 0.61 ± 0.12 mm, which would be adequate for image-guided intervention.

IV. CONCLUSIONS

In this preliminary study, we explored the application of SAM Zero-Shot on the detection of vertebral surfaces from freehand 3D ultrasound data. The SAM Zero-Shot segmentation without prompts, while effective for phantom images, exhibits limited performance on human spine images. In contrast, our proposed LS-TSE extracts accurate surfaces for lumbar vertebrae in vivo. Due to the physical overlap between consecutive thoracic vertebrae, it remains a challenge to localize the spinous process landmark that belongs to the same vertebra level as the lamina bone surface.

Vertebral surface detection offers greater value and reliability compared to limited anatomical landmark detection for X-ray to ultrasound registration in preoperative and intraoperative surgical procedures, and our LS-TSE method can enhance the accuracy of this registration process. Fully rendered 3D spine models derived from the LS-TSE generated 3D point clouds can be compared over time to track changes in spinal structure, monitor the effectiveness of treatment, and assess the progression of degenerative conditions. The technique can additionally foster advancements in medical pedagogy and augmented reality assistance guidance surgery. Future work will address spinous process landmark localization for thoracic vertebrae, and validation with data from patients with spinal deformities.

REFERENCES

1. S. Ghosh, C. Madjdpour, and K. Chin, “Ultrasound-guided lumbar central neuraxial block,” *BJA Education*, vol. 16, no. 7, pp. 213–220, 2016.
2. H.-E. Gueziri, S. Drouin, C. X. B. Yan, and D. L. Collins, “Toward real-time rigid registration of intra-operative ultrasound with preoperative CT images for lumbar spinal fusion surgery,” *Int J Comput Assist Radiol Surg*, vol. 14, p. 1933–1943, 2019.
3. H.-E. Gueziri, C. Santaguida, and D. L. Collins, “The state-of-the-art in ultrasound-guided spine interventions,” *Medical Image Analysis*, vol. 65, p. 101769, 2020.
4. W. Chen, E. H. M. Lou, and L. H. Le, “Using ultrasound imaging to identify landmarks in vertebra models to assess spinal deformity,” in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, pp. 8495–8498.
5. W. Chen, E. H. M. Lou, P. Q. Zhang, L. H. Le, and D. Hill, “Reliability of assessing the coronal curvature of children with scoliosis by using US images,” *Journal of Children’s Orthopaedics*, vol. 7, p. 521–529, 2013.
6. C. J. Cheung, G. Zhou, and S. L. et al., “Ultrasound volume projection imaging for assessment of scoliosis,” *IEEE Transactions on Medical Imaging*, vol. 34, no. 8, pp. 1760–1768, Aug 2015.
7. T. Ungi, H. Greer, K. R. Sunderland, V. Wu, Z. M. C. Baum, C. Schlenger, M. Oetgen, K. Cleary, S. R. Aylward, and G. Fichtinger, “Automatic spine ultrasound segmentation for scoliosis visualization and measurement,” *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 11, pp. 3234–3241, 2020.
8. S. Tang, X. Yang, P. Shajudeen, C. Sears, F. Taraballi, B. Weiner, E. Tasciotti, D. Dollahon, H. Park, and R. Righetti, “A CNN-based method to reconstruct 3-D spine surfaces from US images in vivo,” *Medical Image Analysis*, vol. 74, p. 102221, 2021.
9. W. Jiang and J. Zhang, “3d surface reconstruction of spinal ultrasound images based on 2d contour line,” in *Proceedings of the 2016 3rd International Conference on Materials Engineering, Manufacturing Technology and Control*. Atlantis Press, 2016/04, pp. 110–115.
10. C. Mattjie, L. V. de Moura, R. C. Ravazio, L. S. Kupssinskü, O. Parraga, M. M. Delucis, and R. C. Barros, “Zero-shot performance of the segment anything model (sam) in 2d medical imaging: A comprehensive evaluation and practical guidelines,” 2023.
11. P. Shi, J. Qiu, S. M. D. Abaxi, H. Wei, F. P.-W. Lo, and W. Yuan, “Generalist vision foundation models for medical imaging: A case study of segment anything model on zero-shot medical segmentation,” *Diagnostics*, vol. 13, no. 11, p. 1947, jun 2023.
12. A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi, and G. Fichtinger, “Plus: open-source toolkit for ultrasound-guided intervention systems,” *IEEE Transactions on Biomedical Engineering*, no. 10, pp. 2527–2537, Oct 2014.
13. S. Li, L. Gauthier, F. Cheriet, C. Fortin, P. Paquette, and C. Laporte, “Lamina landmark detection in ultrasound images: a preliminary study,” in *Medical Imaging 2023: Ultrasonic Imaging and Tomography*, vol. 12470. International Society for Optics and Photonics, 2023, p. 1247011.
14. A. Brignol, H. Gueziri, F. Cheriet, D. Collins, and C. Laporte, “Automatic extraction of vertebral landmarks from ultrasound images: a pilot study,” *Computers in Biology and Medicine*, vol. 122, 07 2020.