GPU implementation of coherence-based photoacoustic beamforming for autonomous visual servoing

Eduardo Gonzalez*, Mardava Rajugopal Gubbi† 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—Visual servoing is a promising technique for surgical tool tip tracking and automated visualization of photoacoustic targets during interventional procedures. However, visual servoing in photoacoustic imaging is challenged by the trade-off between laser safety limits and the energy needed for reliable segmentation. Short-lag spatial coherence (SLSC) imaging has the potential to overcome this challenge by improving the signal quality of images acquired with low laser energies. This study introduces the first known GPU-based real-time implementation of SLSC imaging for photoacoustic imaging and applies this real-time algorithm to enhance segmentations for visual servoing. Results with *ex vivo* bovine tissue demonstrate that SLSC imaging recovers signals obtained with low energy (i.e., ≤ 268 μJ) with a signal-to-noise ratio (SNR) of 11.2 ± 2.4, compared to a SNR of 3.5 ± 0.8 with conventional delay-and-sum (DAS) imaging. When energies were lower than the safety limit for skin (i.e., 400 μJ for 900 nm wavelength), real-time SLSC imaging produced lower errors during fiber tracking tasks (i.e., accuracies of 0.67 ± 0.42 mm and 4.91 ± 6.01 mm with SLSC and DAS, respectively). Similarly, for probe centering tests, the tracking error with SLSC and DAS imaging was 0.85 mm $\pm$ 0.44 and 1.05 ± 0.30 mm, respectively. These results are promising for complicated visual servoing tasks in high-noise environments.

I. INTRODUCTION

Visual servoing is a promising technique to maintain visualization of surgical tool tips and other targets during interventional procedures. This technique utilizes a robot-held ultrasound probe to acquire images and an image segmentation technique to locate features of interest within the beamformed image. The robot-held ultrasound probe can either receive ultrasound or photoacoustic signals. Ultrasound-based visual servoing is limited to regions where the acoustic impedance of targets significantly differs from that of the surrounding tissue [1]. Photoacoustic-based visual servoing offers advantages over its ultrasound counterpart, providing enhanced differentiation of tool tips from surrounding tissue [2].

Photoacoustic imaging is implemented by transmitting laser light, which is absorbed by the structures of interest that then experience thermal expansion and generate an acoustic response. The acoustic response is then received by an ultrasound probe. Beamforming techniques are then applied to create a photoacoustic image. In particular, structures of interest for photoacoustic imaging range from blood vessels [3] to neurosurgical tool tips [4] to needle tips [5]. The reliability of photoacoustic-based visual servoing implementations to date depends on the laser energy, which determines the amplitude of the photoacoustic signal of the fiber tip and is limited by laser safety requirements [6].

Short-lag spatial coherence (SLSC) imaging [7] is an advanced beamforming technique that directly displays measurements of the spatial coherence rather than amplitude. SLSC has been demonstrated to improve the contrast of anechoic targets in ultrasound applications by significantly reducing acoustic clutter. In addition, SLSC has been demonstrated to enhance contrast in photoacoustic images without frame averaging, a typical noise reduction technique which has the disadvantage of reduced temporal resolution [8]. SLSC beamforming has also been shown to enhance the quality of photoacoustic signals acquired with low laser energies, which are often necessary to ensure laser safety with miniaturized low-energy light delivery systems, such as pulsed laser diodes [8].

Hyun *et al.* reported a real-time SLSC implementation for harmonic ultrasound imaging, accomplished with two graphical processing units (GPU) on a Verasonics research ultrasound system [9]. Similarly, several ultrasound algorithms have been successfully translated to real-time applications with the aid of GPUs, such as elastography techniques [10], [11], color Doppler methods [12], and advanced beamformers [13], [14]. With the increased prevalence of GPUs in ultrasound equipment designs, we expect that a real-time implementation of SLSC for photoacoustic imaging is feasible with similar resources, thereby enabling the improvement of tool tip sensing in robotic visual servoing systems. In addition, we hypothesize that a more accurate image segmentation with real-time SLSC will enable more accurate tracking of tool tips during surgical guidance scenarios.

This paper presents the first known real-time implementation of SLSC for photoacoustic imaging. To explore the potential benefits of using real-time SLSC over the conventional delay-and-sum (DAS) beamformer in photoacoustic imaging, we conducted two *ex vivo* experiments. First, we measured the SNR of photoacoustic signals in a noisy and aberrated environment, achieved with stacked layers of bovine meat. Second, we assessed performance in two visual servoing tasks: (1) probe centering and (2) fiber tip tracking. In both experiments, we compared the performance of our proposed real-time SLSC beamformer with that of the DAS beamformer.

II. METHODS

A. GPU implementation of SLSC for photoacoustic imaging

A summary of the overall pipeline for real-time SLSC in photoacoustic imaging is presented in Fig. 1. First, raw

channel data were acquired by the ultrasound system, which was triggered by a signal from the laser system. Depending on the ultrasound system memory allocation and the number of available channels, a regrouping process (i.e., "Regroup channels" in Fig. 1) was needed prior to beamforming to obtain a $N \times P$ matrix, where N is the number of elements and P is the number of axial samples. DC removal was computed with 1D convolutions of time-domain kernels executed independently along the axial dimension. Next, the Hilbert transform was computed along the axial dimension using the FFT libraries embedded in CUDA (NVIDIA, Santa Clara, CA, USA). The computation of receive delays was distributed in the ratio of 1 thread per axial sample, and the number of scanlines was equal to the number of receive elements (i.e. 128).

For the SLSC computation, a parallel processing scheme was adapted from the real-time implementation of SLSC for ultrasound imaging demonstrated by Hyun *et al.* [9]. In contrast to the conventional SLSC [7], a set of coherence factors were computed element-wise for each lag m:

$$C_{ij}(z,x,m) = \sum_{i=1}^{N-m} s_i(z,x) s_{i+m}(z,x)^* \tag{1}$$

$$C_{ii}(z,x,m) = \sum_{i=1}^{N-m} |s_i(z,x)|^2 \tag{2}$$

$$C_{jj}(z,x,m) = \sum_{i=1}^{N-m} |s_{i+m}(z,x)|^2 \tag{3}$$

where $s_i(z,x)$ is a complex signal at channel i, scanline x, and axial sample z, $\{*\}$ denotes the complex conjugate, and N is the total number of channels. The coherence factors (C_{ij}, C_{ii}, and C_{jj}) were stored in the device memory. They were then compounded across a spatial kernel k_z and a cumulative lag M as follows:

$$\text{SLSC}(z,x) = \sum_{m=1}^{M} \frac{\sum_{\hat{z} \in k_z} C_{ij}(\hat{z},x,m)}{\sqrt{\sum_{\hat{z} \in k_z} C_{ii}(\hat{z},x,m) \sum_{\hat{z} \in k_z} C_{jj}(\hat{z},x,m)}} \tag{4}$$

Finally, the image was normalized, truncated, and log-compressed. The maximum term for normalization was computed using logarithmic reduction strategies [15]. The SLSC images were computed on a GeForce GTX 1080 GPU (NVIDIA Corporation, Santa Clara, CA, USA), with $M = 25$ (39% of the aperture) and $k_z = 7$ axial samples.
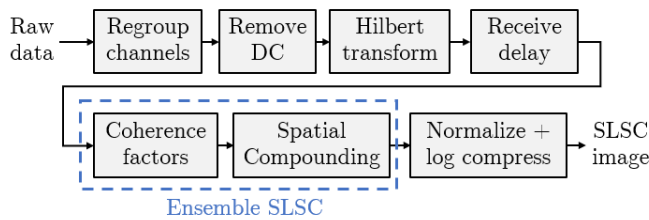
B. Experimental setup

Real-time DAS and SLSC for photoacoustic imaging was compared in two experimental setups. The first setup focused on assessing the similarities between the GPU and CPU implementations of the real-time SLSC algorithm, while also comparing the SNR of the SLSC algorithm with that of the DAS algorithm in a noisy environment. A 1 mm-diameter optical fiber was inserted into stacked layers of bovine meat at a depth of 25 mm. Considering that the main source to photoacoustic interaction comes from the tip of the fiber, we did not attach it to any surgical tools in this study. An L3-8 linear probe was connected to an Alpinion E-Cube 12R system (Alpinion Medical Systems, Seoul, Korea), and positioned at the top of the bovine meat, as shown in Fig. 2(a). The fiber transmitted 900 nm-wavelength light from a Phocus Mobile laser (Opotek Inc., Carlsbad, CA, USA), with energies ranging from 118 to 645 μJ. For this wavelength and geometry, the energy equivalent to the maximum permissible exposure for skin is 400 μJ. The SNR was computed with the following formula:

$$\text{SNR} = \mu_{\text{signal}} / \sigma_{\text{background}} \tag{5}$$

The second setup focused on the effect of SLSC and DAS on visual servoing. The same optical fiber was inserted into the muscle section of the *ex vivo* bovine tissue at a depth of 4 cm, and the probe was attached to a Sawyer robot (Rethink Robotics, Boston, MA) as shown in Fig. 2 (b). The visual servoing algorithm consisted of a series of morphological operations and binary thresholding to detect the photoacoustic signal in order to maintain this signal at the center of the photoacoustic image, which corresponds to the center of the robot-held ultrasound probe.

The visual servoing system was first validated by the probe centering experiment described in [2]. The probe was placed on the surface of the bovine tissue and the length of the optical fiber was aligned with the imaging plane. The tip of the optical fiber was laterally displaced from the center of the image. The visual servoing system was activated and the robot was

Fig. 1. Pipeline for generating a SLSC photoacoustic image in real time.
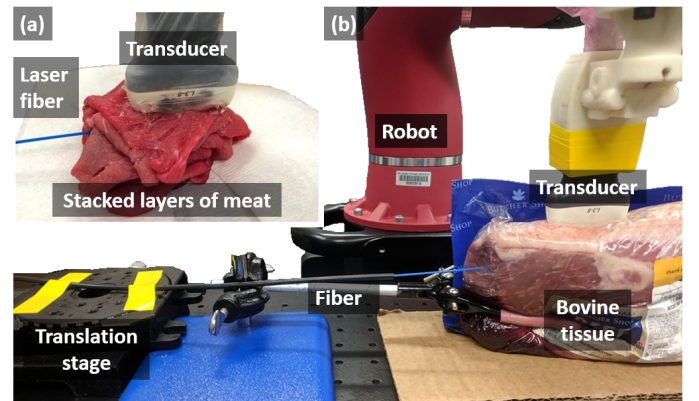

Fig. 2. Photoacoustic acquisition setups. (a) Optical fiber inserted into stacked layers of *ex vivo* bovine tissue. (b) Optical fiber attached to translation stage and inserted into *ex vivo* bovine tissue. As the optical fiber was translated, an ultrasound transducer connected to a robot arm performed visual servoing.
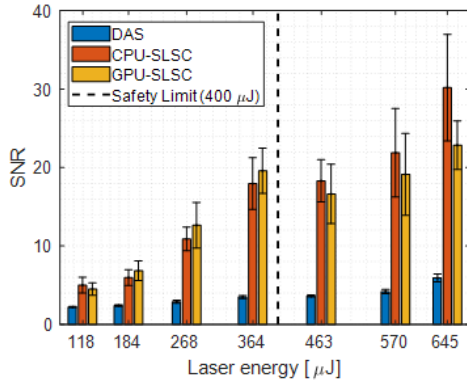
Fig. 3. SNR results from the optical fiber inserted into stacked layers of bovine tissue as a function of the laser energy.

Fig. 4. Results of target tracking at 248 μJ. The circles represent time stamps where the visual servoing algorithm fails to segment the photoacoustic signal.

Fig. 5. Tracking error of the fiber tracking experiment.

commanded to maintain the lateral center of the probe on the fiber tip.

Once the probe was centered, the ability of the visual servoing system to follow the fiber tip was assessed over a distance of 10 mm using a translation stage. We refer to this experiment as the fiber tracking experiment. The two visual servoing experiments were repeated with laser energies of 169, 248, and 322 μJ, which are below the safety limit of 400 μJ for a wavelength of 900 nm.

III. RESULTS & DISCUSSION

A. SNR and computation measurements in tissue layers

Fig. 3 shows the quantitative results of the SNR measurements. SLSC consistently outperformed DAS and visualized low-energy signals ($\leq$ 268 μJ) with a mean SNR of 11.2 $\pm$ 2.4 when compared to that of DAS (i.e., 3.5 $\pm$ 0.8). In addition, the SNR difference between the CPU-based and GPU-based SLSC implementations was 1.14 $\pm$ 3.99. This minimal SNR difference is likely due to the single precision libraries and interpolation in the GPU texture memory.

The computation time was significantly reduced with the GPU-based SLSC (i.e., 24.3 ms per frame), performing more than 402 times faster than the CPU-based SLSC, which had a frame rate of 9.8 Hz. Considering that the pulse repetition frequency of the laser was 10 Hz, we conclude from these SNR and computation time comparisons that the GPU-based SLSC implementation is more beneficial for real-time imaging.

B. Fiber tracking in bovine muscle

Fig. 4 shows the trajectory of the robot arm during a fiber tracking test when using DAS and SLSC photoacoustic beamforming for segmentation. While both beamformers initially follow the fiber displacement, eventually the noise present in the DAS image contributes to failed segmentations in the visual servoing algorithm. The failed segmentations are identified by the speed of the robot arm, resulting in a displacement of >8 mm over a duration of 1 ms. These failed segmentations are marked in Fig. 4 with circles.
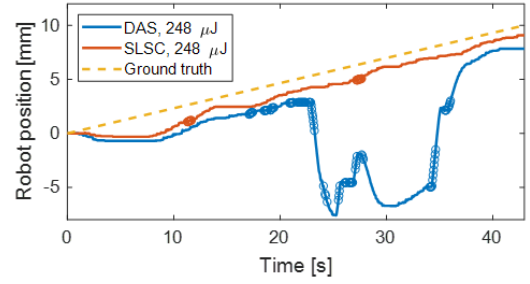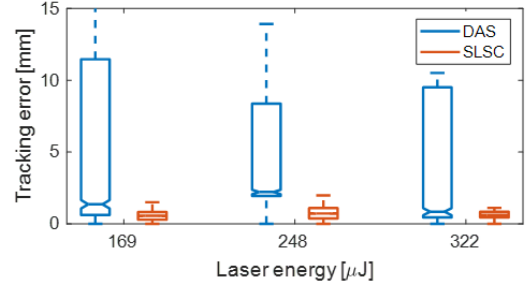
When consecutive instances of failed segmentation are occur over a 1 s time period, the robot performs a search around the current region in the lateral and elevation ultrasound probe directions. This searching algorithm is responsible for the mean $\pm$ one standard deviation of 10.35 $\pm$ 3.47 mm displacement from ground truth with DAS images, as seen within the 21 s to 38 s time interval in Fig. 4. When excluding this interval, both beamformers have deviations from the ground truth locations (DAS: 1.92 $\pm$ 0.55 mm, SLSC: 0.77 $\pm$ 0.51 mm), which is likely caused by the manual operation of the translation stage resulting in a non-constant velocity or deflections of the fiber tip during insertion into the tissue.

Fig. 5 shows the tracking errors between the ground truth locations and the measured robot positions for the fiber tracking tests. Generally, tracking errors are larger with DAS compared to SLSC for each laser energy shown (i.e., 169 μJ, 248 μJ, and 322 μJ). The overall mean $\pm$ one standard deviation of tracking errors were 4.91 $\pm$ 6.01 mm and 0.67 $\pm$ 0.42 mm with DAS and SLSC images, respectively, for the three laser energies.

C. Probe centering in bovine muscle

Fig. 6 shows the trajectory of the robot arm during a probe centering test when using DAS and SLSC photoacoustic beamforming for segmentation. At 248 μJ, the visual servoing system using SLSC has an underdamped response given a step impulse, while the DAS counterpart has an overdamped response. In addition, the steady state of the DAS result produced an increased tracking error (i.e., 1.01 mm) when compared to that of SLSC (i.e., 0.52 mm). We suspect that
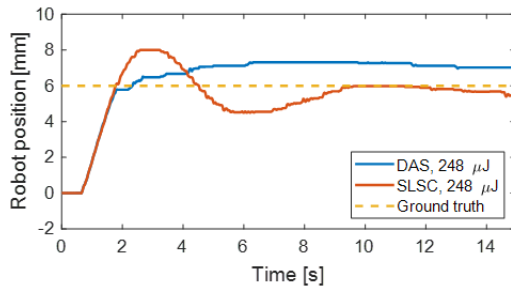
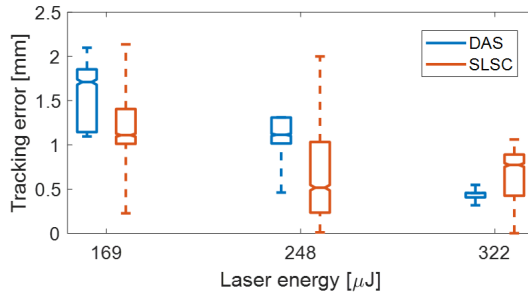Fig. 6. Results of probe centering at 248 μJ energy.

Fig. 7. Tracking error of the probe centering experiment.

the source of the increased tracking error with DAS images is the area of the signal appearance in comparison to that of SLSC. The larger segmented areas in DAS images increase the effective area of segmentation. This increase causes deviations in the position of the centroid, which results in a less accurate position estimate for the detected signal. While the side lobes that were partially responsible for the larger areas in DAS were similarly present in SLSC, they were reduced by increasing the cumulative lag M in Eq. 4. In this study, $M = 25$ was sufficient to provide a small segmented area without compromising the speed of the real-time imaging system.

Fig. 7 shows the tracking errors between ground truth locations and the robot positions during the probe centering experiments. Tracking errors from this probe centering experiment were reported over a time period of 12 to 15 s after the first crossing of the robot position with the ground truth (e.g., 1.8 s in Fig. 6). For 169 μJ and 248 μJ laser energy, SLSC imaging produced lower errors compared to DAS. At a laser energy of 322 μJ, although DAS has lower errors than SLSC, Fig. 5 demonstrates that the fiber tracking error with the DAS implementation was significantly higher than that of the SLSC implementation. Considering that the success of visual servoing is affected by both the tool tip tracking errors and the probe centering errors, the SLSC implementation remains the better choice for clinical applications at this higher energy.

IV. CONCLUSION

This paper presents the first known implementation of real-time SLSC beamforming for photoacoustic data, which was enabled by GPUs and parallel processing techniques. *Ex vivo* results with bovine tissue demonstrate that SLSC is beneficial over DAS for photoacoustic-based visual servoing. Benefits include higher SNRs in the beamformed photoacoustic images at lower energies and smaller, more reliable segmentations of photoacoustic targets. The improved segmentation is responsible for more accurate steady state positioning and improved tracking of a moving signal during visual servoing.

ACKNOWLEDGMENTS

REFERENCES

[1] J. Kettenbach, G. Kronreif, M. Figl, M. Fürst, W. Birkfellner, R. Hanel, and H. Bergmann, "Robot-assisted biopsy using ultrasound guidance: initial results from in vitro tests," *European radiology*, vol. 15, no. 4, pp. 765–771, 2005.

[2] M. A. L. Bell and J. Shubert, "Photoacoustic-based visual servoing of a needle tip," *Scientific reports*, vol. 8, no. 1, p. 15519, 2018.

[3] R. G. Kolkman, W. Steenbergen, and T. G. van Leeuwen, "In vivo photoacoustic imaging of blood vessels with a pulsed laser diode," *Lasers in Medical Science*, vol. 21, no. 3, pp. 134–139, 2006.

[4] B. Eddins and M. A. L. Bell, "Design of a multifiber light delivery system for photoacoustic-guided surgery," *Journal of Biomedical Optics*, vol. 22, no. 4, p. 041011, 2017.

[5] J. Su, A. Karpiouk, B. Wang, and S. Emelianov, "Photoacoustic imaging of clinical metal needles in tissue," *Journal of Biomedical Optics*, vol. 15, no. 2, pp. 021 309–021 309, 2010.

[6] M. Graham, F. Assis, D. Allman, A. Wiacek, E. Gonzalez, M. Gubbi, J. Dong, H. Hou, S. Beck, J. Chrispin, and M. A. L. Bell, "In vivo demonstration of photoacoustic image guidance and robotic visual servoing for cardiac catheter-based interventions," *IEEE Transactions on Medical Imaging*, 2019. [Online]. Available: http://dx.doi.org/10.1109/TMI.2019.2939568

[7] 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.

[8] M. A. L. Bell, X. Guo, H. J. Kang, and E. Boctor, "Improved contrast in laser-diode-based photoacoustic images with short-lag spatial coherence beamforming," in *2014 IEEE International Ultrasonics Symposium*. IEEE, 2014, pp. 37–40.

[9] D. Hyun, A. L. C. Crowley, M. LeFevre, J. Cleve, J. Rosenberg, and J. J. Dahl, "Improved visualization in difficult-to-image stress echocardiography patients using real-time harmonic spatial coherence imaging," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 66, no. 3, pp. 433–441, 2018.

[10] E. A. Gonzalez, S. E. Romero, and B. Castaneda, "Real-time crawling wave sonoelastography for human muscle characterization: Initial results," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 66, no. 3, pp. 563–571, 2018.

[11] A. Baghani, H. Eskandari, W. Wang, D. Da Costa, M. N. Lathiff, R. Sahebjavaher, S. Salcudean, and R. Rohling, "Real-time quantitative elasticity imaging of deep tissue using free-hand conventional ultrasound," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2012, pp. 617–624.

[12] A. J. Chee, B. Y. Yiu, and C. Alfred, "A gpu-parallelized eigen-based clutter filter framework for ultrasound color flow imaging," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 64, no. 1, pp. 150–163, 2016.

[13] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, and M. Fink, "Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 56, no. 3, pp. 489–506, 2009.

[14] E. A. R. Nielsen, "Real-time wavelet filtering on the gpu," Master's thesis, Institutt for datateknikk og informasjonsvitenskap, 2007.

[15] S. Cook, *CUDA programming: a developer's guide to parallel computing with GPUs*. Newnes, 2012.