

SonoAssist: Open source acquisition software for ultrasound imaging studies

David Olivier¹ and Catherine Laporte¹

¹ Department of electrical engineering, École de technologie supérieure, Montreal, Canada

Abstract— We present “SonoAssist”, an open-source acquisition software designed to facilitate the development of point-of-care ultrasound (POCUS) assistance systems by simplifying the data collection process. This software caters to research utilizing ultrasound images along with gaze data or probe movement measurements to tackle tasks like standard scan plane detection, anatomical landmark detection, and ultrasound probe guidance. Through SonoAssist’s simple interface, users can easily collect data from the following sensors: an ultrasound probe, an RGBD camera, an eye tracker, a screen recorder, and IMUs (Inertial Measurement Unit). Furthermore, SonoAssist timestamps data as they are acquired with a single time reference, removing the need for additional synchronization steps. To document the software’s performance, we characterized the synchronization between the ultrasound image and IMU data streams, the eye tracker accuracy, and the acquisition frequencies (ultrasound probe: 22 Hz, eye tracker: 87 Hz, external IMU: 100 Hz, screen recorder: 13Hz).

Keywords— Ultrasound imaging, eye tracking, motion tracking, open source

I. INTRODUCTION

The advent of point-of-care ultrasound (POCUS) created a need for tools that make ultrasound imaging more accessible to medical practitioners outside radiology departments. Indeed, recent research has investigated methods to assist new users of ultrasound by automating interpretation tasks like standard scan plane detection, scan plane classification, landmark detection, and volume reconstruction or providing some form of visual guidance during scanning. Several of these methods utilize information from sources other than the ultrasound images, like the sonographer’s gaze [1,2,3,4,5] and the movements of the ultrasound probe [6, 7]. Sonographer gaze data are useful for the development of saliency prediction models [3,4,5] and attention gated models [6, 8], while probe movement and position data are useful for representation learning [9] and developing volume reconstruction models [6], and probe guidance models [7].

This paper presents an open-source data acquisition tool, SonoAssist¹, that we developed to facilitate the creation of

POCUS assistance systems. The software is easy to use and can acquire synchronized streams from the following devices: an eye tracker, an RGBD camera, IMUs (Inertial Measurement Unit), and an ultrasound scanner. Other open-source projects such as PLUS [10] and IGSTK [11] have previously facilitated the rapid prototyping of ultrasound-guided intervention systems through compatibility with a wide range of tracking and imaging devices as well as a broad variety of functionalities including routines for spatial and temporal calibration and synthetic ultrasound image generation. In contrast with these open-source toolkits, SonoAssist supports the use of eye tracking, is specifically designed for the previously mentioned set of use cases and is ready for use out of the box with no development required.

Recent papers about the utilization of sonographer expertise to automate tasks related to medical ultrasound present their own data acquisition systems tailored to the researchers’ specific needs [1, 7]. These systems are very similar to ours. For instance, the acquisition system developed by Ahmed and Noble [1] acquires ultrasound images from a portable ultrasound scanner and gaze data from an eye tracker to develop a scan plane classifier that utilizes human eye movements. However, despite having been used for research, the design of such systems is rarely thoroughly documented nor is their performance characterized in the literature. This paper aims to bridge this gap by detailing the design of such a system, thoroughly characterizing some of its performance metrics, and making it available for consultation and customization.

The rest of this paper covers the general design of SonoAssist (Section 2), the mechanisms involved in the acquisition and synchronization of data streams (Section 3), and the handling of gaze data (Section 4). We also provide a quantitative performance evaluation of the software’s performance in terms of device acquisition rates and the accuracy and precision of eye tracking.

II. SOFTWARE DESIGN AND USAGE

A. Software I/O

SonoAssist can acquire data from five different sources: an ultrasound probe (L7 linear, Clarius), an RGBD camera (Realsense D435, Intel), an eye tracker (4C, Tobii), an IMU

¹ <https://github.com/OneWizzardBoi/SonoAssist>

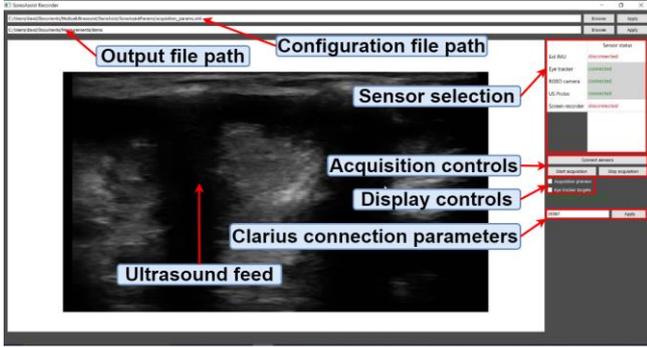


Fig. 1 SonoAssist’s main mode interface.

(MetaMotionC, MbiEntLab), and a screen recorder. In terms of data flow, dedicated acquisition threads pull data from the sensors to write them to the appropriate output files. For instance, sensors providing a stream of images are associated with two file types. The first one (.avi or .bag) contains the acquired images and the second one (.csv) indexes them for synchronization purposes. The remaining sensors are associated with (.csv) output files which store and index the measurements.

B. User interface and intended use

Simplicity of use is the principal requirement that guided the design of SonoAssist. Consequently, the user interface is light and the more complex configuration parameters are tucked away in a configuration file. In terms of use, the software has two display modes. The first mode is the preview mode and it allows the user to verify the different input streams when setting up an acquisition. The second mode is the main mode (Fig. 1) and it displays the ultrasound image feed in real-time. During an acquisition, the sonographer uses the main mode to analyze the ultrasound images from the probe in real-time while the eye tracker positioned under the laptop screen is recording his or her gaze and the RGBD camera is recording the movements of the ultrasound probe.

III. TIMING AND SYNCHRONIZATION

A. Ultrasound image display rate

In our benchmarking experiments, we used an HP Pavilion 15-cs2 laptop equipped with an Intel i7-8565U CPU.

The first row of Table 1 provides the ultrasound display rates for two different ultrasound display sizes: (540 x 960) px for config 1 and (720 x 1280) px for config 2. We measured an average display rate of 22.72 Hz for the smaller display and 22.43 Hz for the larger one. To obtain these

measurements, we simply counted the number of frames displayed during benchmarking acquisition sequences and divided the count by the duration of the sequences. In total, ten acquisition sequences, ranging from 50 to 65 seconds in length, were analyzed for each configuration.

In addition to the average value, the variation of the display rate was also characterized to ensure that it remains stable throughout the acquisition sequences. We measured an average time interval of 43 milliseconds (23.19 Hz) separating the display of consecutive ultrasound images with a standard deviation of 6.7 milliseconds. Note that this measurement covers both display sizes (config 1 and config 2) and was obtained by analyzing 20 acquisition sequences ranging from 50 to 65 seconds in length. By analyzing the measured frequency distribution, we can conclude that 95.5% of the frames are displayed at a rate ranging between 16.9 Hz and 29.4 Hz. On a practical level, these results validate SonoAssist’s acquisition design because the display rate values are relatively close to the previously discussed average display rates (22.72 Hz and 22.43 Hz). This means that, over the length of an acquisition sequence, the displayed images are evenly spread out in time. Note that the rate variations are mainly due to the fact that the software does not run on dedicated hardware and that the display rate is tied to the ultrasound scanner’s acquisition rate, which can vary.

Handling data from sources other than the ultrasound scanner (i.e. the screen recorder, the eye tracker, and the external IMU) requires fewer steps because the corresponding data are not displayed. Indeed, after the reception of the data, or their generation, in the case of the screen recorder, they are timestamped and directly saved to the disk. Consequently, measuring the acquisition rate associated with the previously mentioned sensors boils down to analyzing the timestamps to calculate the average number of acquisitions saved to the disk every second. However, future versions of the software could include additional data visualization functionalities which would call for a change in our benchmarking approach. Table 1 provides the average acquisition rate for each sensor. We can observe that the screen recorder’s average acquisition rate is 13.64 Hz and that across the 3 different sensor configurations, the average acquisition rates for the external IMU and the eye tracker are respectively about 101 Hz and 86 Hz. As expected, the average acquisition rates for the external IMU and the eye tracker are close to the rates specified by the manufacturer which are respectively 100 Hz and 90 Hz.

B. Ultrasound probe data synchronization

When characterizing SonoAssist, we did not assume that the image and IMU data pairs from the ultrasound probe are

synchronized. In practice, one of the data streams could consistently be ahead of the other due to the ultrasound scanner’s inner workings. Therefore, to quantify the time

Table 1 Average acquisition frequencies (Hz) for sensors across different sets of configurations. The averaging was done over 10 acquisitions of 50 seconds for all configurations. Acquisition configuration #1: external IMU, eye tracker, RGBD camera and Clarius probe with a (540 x 960) (px) display. Acquisition configuration #2: external IMU, eye tracker, RGBD camera and Clarius probe with a (720 x 1280) (px) display. Acquisition configuration #3: external IMU, eye tracker, RGBD camera, and screen recorder.

Sensors	Config #1	Config #2	Config #3
US Probe with integrated IMU	22.72 Hz	22.43 Hz	None
Screen recorder	None	None	13.64 Hz
Eye tracker	87.01 Hz	86.90 Hz	85.82 Hz
External IMU	101.75 Hz	101.76 Hz	101.2 Hz

offset between the data pair members, we performed an experiment inspired by the temporal calibration procedure described in the Stradx freehand 3D ultrasound software paper [12]. The experiment consists of simultaneously applying a step input to the image and IMU data streams by holding the probe still against a phantom for a few seconds and then jerking it away. This jerk motion causes a large variation in both data streams, and the time difference between these variations corresponds to the offset between the data pair members.

To quantify the variations in the image stream, we used the same method as in Stradx [12], which takes in two consecutive images and produces a difference measure (Δim) that is robust to speckle noise and sensitive to changes across their whole surface. To measure changes in the IMU data stream, we calculated the variation in the linear acceleration vector of consecutive acquisitions. This calculation (Eq. 1) takes in 2 consecutive IMU acquisition (A_1) and (A_2) and calculates the variation (ΔA):

$$\Delta A = \left| \frac{\|A_2\| - \|A_1\|}{\|A_1\|} \right|, \quad (1)$$

where $A_1 = [a_{1x}, a_{1y}, a_{1z}]$ and $A_2 = [a_{2x}, a_{2y}, a_{2z}]$.

To identify the start of the jerk motion, we set a variation threshold of 2% for (ΔA) and 100% for (Δim) when compared to the average value. The results of the experiment (Fig. 2), show that, for the selected thresholds, the IMU data and the ultrasound image streams are offset by a single acquisition period (43 milliseconds). However, because the identified motion start points are so close, strategic threshold values could be selected to reduce the offset to 0 seconds. Therefore, we consider the ultrasound image and IMU data streams to be synchronized.

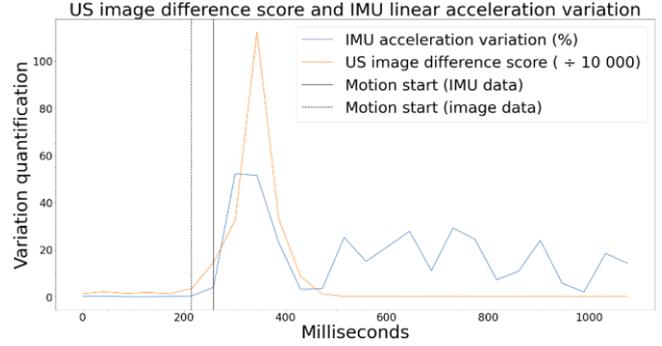


Fig. 2 Results of the synchronization test performed on the paired IMU data and ultrasound images from the ultrasound probe.

IV. GAZE DATA PROCESSING

A. Measuring the default gaze tracking accuracy

SonoAssist utilizes the Tobii 4C, an inexpensive consumer-grade eye tracker made for desktop gaming and mouse control to collect gaze data. Since these applications do not require the collected gaze data to be saved for later analysis, the manufacturer does not provide precise accuracy benchmarks for the device. Despite this, we still decided to use the Tobii 4C because of its accessible price point and because it was used in previous studies [3, 5].

To generate accuracy benchmarks, we asked three volunteers to conduct custom acquisition sequences designed to enable the comparison of gaze point positions with ground truth positions. Indeed, the acquisitions utilize visual targets (red dots), positioned on the corners of the main mode’s ultrasound display for which the exact positions are known. During these acquisitions, the user places his or her head approximately 0.75 meters away from the display and sequentially stares at each target for 5 seconds. Furthermore, the user must visit each target three times while moving his or her gaze as fast as possible in between them. Globally, each volunteer performed this sequence three times, so that nine acquisitions were conducted in total.

When analyzing data from these experiments, gaze points associated with a high movement speed were filtered out by a 0.5°/second gaze speed filter, leaving only the points associated with fixations on the visual targets. After the filtering, the average distance, in pixels, between each gaze point and its nearest visual target was calculated. This measurement represents the error between the position of collected gaze positions and their intended target. The top half of Table 2 presents this average error for each of the nine benchmarking acquisitions. Note that for the global

error, the (X) and (Y) percentages correspond to the portion of the screen's width and height occupied by the error.

B. Gaze tracking error correction

Because the average tracking error for users remains relatively constant across their acquisition sequences, the tracking accuracy can be improved in post-processing, for each user, by shifting the position of all gaze points by the opposite of their tracking error. The bottom half of table 2 presents the average tracking error for each of the nine benchmarking acquisitions after the error correction procedure. When comparing the gaze accuracy measurements before and after the correction, we observe that the procedure produces a significant accuracy improvement. Indeed, the global average error is reduced from 52.8 (px) to 23.9 (px) which corresponds to a 54.7% improvement in accuracy.

Table 2 Average gaze tracking error (in pixels) measured in the benchmarking acquisitions, before and after performing gaze error correction. Note that for the global error, the (X) and (Y) percentages correspond to the portion of the screen's width and height occupied by the error.

Before gaze drift correction			
	Subject 1	Subject 2	Subject 3
Acquisition #1	67.8 px	69.3 px	29.2 px
Acquisition #2	67.1 px	67.3 px	28.2 px
Acquisition #3	59.7 px	58.0 px	28.2 px
Global average error: 52.8 px (X: 2.7 % and Y: 4.9%)			
After gaze drift correction			
	Subject 1	Subject 2	Subject 3
Acquisition #1	31.0 px	22.2 px	22.0 px
Acquisition #2	26.8 px	24.1 px	18.7 px
Acquisition #3	31.2 px	21.8 px	17.4 px
Global average error: 23.9 px (X: 1.2 %, Y: 2.2 %)			

V. CONCLUSIONS

This paper presented SonoAssist, software that facilitates the data collection process associated with the development of POCUS assistance systems. Practically, the software's US image display (22 Hz) is responsive enough to be used for freehand imaging and the acquisition rates of the other relevant sensors are appreciably higher (eye tracker: 87 Hz, external IMU: 100 Hz). This means that all acquired ultrasound images can be paired with useful data from these sources. Furthermore, the gaze tracking accuracy (+/- 23.9 pixels) is suitable for the generation of saliency maps because the error is just under 1° of visual angle (33 horizontal pixels and 92 vertical pixels for an HD screen). As such, SonoAssist promises to be a useful tool to speed up the development of research prototypes for the study of sonog-

rapher behavior and the development of POCUS guidance systems.

ACKNOWLEDGMENT

This work was supported by MITACS and NSERC.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. M. Ahmed and J. Noble, "Fetal Ultrasound Image Classification Using a Bag-of-words Model Trained on Sonographers' Eye Movements," *Procedia Computer Science*, vol. 90, pp. 157–162, 2016, doi: 10.1016/j.procs.2016.07.021.
2. Y. Cai, H. Sharma, P. Chatelain, and J. Noble, "SonoEyeNet: Standardized fetal ultrasound plane detection informed by eye tracking," *Proceedings / IEEE International Symposium on Biomedical Imaging: from nano to macro. IEEE International Symposium on Biomedical Imaging*, pp. 1475–1478, 2018, doi: 10.1109/ISBI.2018.8363851.
3. R. Droste et al., "Ultrasound Image Representation Learning by Modeling Sonographer Visual Attention." 2019.
4. R. Droste, Y. Cai, H. Sharma, P. Chatelain, A. T. Papageorghiou, and J. Noble, "Towards Capturing Sonographic Experience: Cognition-Inspired Ultrasound Video Saliency Prediction." 2019.
5. R. Droste, P. Chatelain, L. Drukker, H. Sharma, A. Papageorghiou, and J. Noble, "Discovering Salient Anatomical Landmarks by Predicting Human Gaze." 2020.
6. H. Guo, S. Xu, B. Wood, and P. Yan, "Sensorless Freehand 3D Ultrasound Reconstruction via Deep Contextual Learning." 2020.
7. R. Droste, L. Drukker, A. Papageorghiou, and J. Noble, "Automatic Probe Movement Guidance for Freehand Obstetric Ultrasound." 2020.
8. Y. Gao and J. Noble, "Learning and Understanding Deep Spatio-Temporal Representations from Free-Hand Fetal Ultrasound Sweeps," 2019, pp. 299–308.
9. J. Jiao, R. Droste, L. Drukker, A. Papageorghiou, and J. Noble, "Self-supervised Representation Learning for Ultrasound Video." 2020.
10. A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi, and G. Fichtinger, "PLUS: open-source toolkit for ultrasound-guided intervention systems.," *IEEE Trans Biomed Eng.*, vol. 61, no. 10, pp. 2527–2537, Oct. 2014, doi: 10.1109/TBME.2014.2322864.
11. A. Enquobahrie et al., "The Image-Guided Surgery Toolkit IGSTK: An open source C++ software toolkit," *Journal of digital imaging: the official journal of the Society for Computer Applications in Radiology*, vol. 20 Suppl 1, pp. 21–33, 2007, doi: 10.1007/s10278-007-9054-3.
12. R. W. Prager, A. Gee, and L. Berman, "Stradx: real-time acquisition and visualization of freehand three-dimensional ultrasound," *Medical Image Analysis*, vol. 3, no. 2, pp. 129–140, Jun. 1999, doi: 10.1016/S1361-8415(99)80003-6.