A Computer Vision Based Tracking Framework for Medical Training

Bo Sun
Department of Computer Science
Rowan University
Glassboro, NJ, USA
e-mail: sunb@rowan.edu

Lucas Grebe
Department of Computer Science
Rowan University
Glassboro, NJ, USA
e-mail: grebel42@students.rowan.edu

Abstract— **Medical simulation often utilizes standardized patients or mannequins to train and assess medical students. We present a computer vision-based framework where we can accurately augment virtual pathology to a standardized patient/mannequin. The framework offers a unique marker-based tracking approach using the computer vision method, which delivers a fast and precise augmentation for a simulated cranial (head) ultrasound prototype. The technique is very efficient and cost-effective compared with other tracking approaches.**

Keywords-medical simulation; computer vision; human computer interaction framework.

I. INTRODUCTION

According to the Safe Healthcare blog of the US Centers for Disease Control and Prevention (CDC) [1], more people die from medical errors and related issues than from lack of access to medical services. In the US, unintentional injury, including medical errors, kills more than 400,000 people annually [1]. Medical simulation, a computer-based simulation technique, has been widely adopted by medical schools to strengthen clinical skills by increasing indirect patient-related practices. The simulation techniques often use mannequins and Standardized Patients (SPs) to produce a realistic training scenario. However, since most SPs/mannequins are "healthy" individuals, medical pathologies are added via computer simulation. Depending on the diagnosis procedures, rendering of the simulated pathology based on the movement of a diagnostic tool (such as a stethoscope or an ultrasound scanner) requires a fast and precise tracking approach.

Many traditional simulations use electromagnetic trackers, which include a heavy magnetic device, reference cube, and marker, to enable tracking. This tracking system has space and environment requirements and is often connected with long cables. Other simulations involve haptic devices, where the force and movement of medical devices are tracked. These devices are typically heavy-duty and very expensive, depending on the tracking accuracy required by the simulator.

To make the simulation more portable, we present a Computer Vision-based Tracking (CVT) technique in this paper. The system can be carried around easily and offers a cost-effective solution compared to a magnetic tracker and haptic device. The new approach permits a more convenient way for medical training and assessment.

The CVT system consists of a camera and a box-shape marker. The technique uses OpenCV to detect the marker and produce translation and rotation vectors related to the marker. We applied a subtraction factor to increase the tracking space limitation of one camera. Comparing with existing optical based tracking techniques that often use multiple cameras or other sensor devices, our CVT only adopted one camera and relies on a unique box shape marker to ensure the tracking accuracy. The new approach permits a medical trainee to conduct training at any chosen location, such as home, office and lab, using a laptop with a build-in camera. In contrast, traditional optical setting involving multicamera or sensors often requires dedicated lab space to support the training since the set up requires specific distances and angles among the devices.

The framework was applied to a cranial (head) ultrasound prototype using a 3D brain mesh model. The system runs efficiently and accurately to produce an ultrasound slice when examining a brain.

The paper will present the related work in Section 2, detail the method in Section 3, summarize the results in Section 4, and conclude the study in Section 5.

II. RELATED WORKS

Medical simulation techniques could produce two types of outcomes depending on the stage of development. One is a research prototype, likely produced by higher education. This type of work can be retrieved through paper publication. The other is a commercial simulator. Its technique is not available for the reproduction of research; therefore, has little value to the research community.

The tracking techniques generally used by medical simulation are 1) electromagnetic tracker, 2) haptic device, and 3) optical tracking [2]. Therefore, we separate our reviews into these areas in this section. We will focus on tracking techniques often seen in ultrasound simulation initially, then detail optical tracking usage in general medical simulation.

A. *Tracking Techniques in Ultrasound Simulation*

Schallware Ultrasound Simulator [3] adopted a mannequin and several dummy probes connected with a magnetic tracker. The simulator system consists of pre-collected clinic pathological data as training cases in internal medicine, cardiology, obstetrics, and gynecology. Simbionix ultrasound mentor [4] is a high-fidelity simulator. The system uses magnetic tracking techniques along with mock probes to scan mannequins in developing ultrasound training for different genders and ages for various medical specialties.

As seen in these commercial simulators' manuals and brochures, magnetic tracking equipment is used to support the tracking. The probes from both simulators connect with a heavy-duty cable to track the device's movement precisely. In addition, the sensitivity of the magnetic tracking can be quickly impacted by other metal particles nearby, as described in [5].

Other commercial simulators use haptic devices as input interfaces. The device uses electric sensors to monitor the hand motion of trainees along with the force applied to the device. So, haptic devices are commonly seen among simulators. ScanTrainer [6], a transvaginal and transabdominal ultrasound simulation, adopted a haptic device to deliver the training. They used a virtual patient to mimic a physical body. To study the skills, a trainee must watch two monitors (one for simulated ultrasound and one for VR patient). The tracking is completely handled by the haptic device without visual indicators.

The most recent development from Surgical Science uses VR techniques [7] that adopted optical tracking embedded in the VR headset to monitor the movement of mock devices, which permits trainees to complete the entire training in VR.

Neither haptic devices nor VR headsets are feasible devices for developing a low-cost portable system. Therefore, the field needs an optical tracking platform to support the future development of ultrasound simulation.

B. *Optical Tracking*

Markov-Vetter et al. created an ultrasound simulation, CranUS [8], based on optical tracking. The system simulates cranial (head) ultrasound for infants. However, the simulation focused on the training outcomes and had no details about the optical tracking techniques used in the research prototype. According to a figure illustration of the paper, they adopted two large overhead-mounted cameras to visually track the operational space of the scanning. Duan et al. [9] created a 3D tracking system using three cameras and two markers for a virtual surgery simulation. The system obtained 3D coordinates of each marker for every two cameras, then used the average of six coordinates produced by three pairs of camera groups to archive accurate 3D coordination. OptiTrack [10], a commercial optical tracking system, requires multiple high-end overhead cameras. The system has less accuracy compared with a laser tracking system, SteamVR, according to [11]. After all, none of these techniques are portable when multiple cameras are involved. Volutracer O.P.U.S. [12] is a relatively new commercial ultrasound simulator that uses optical tracking. It adopts a web camera, a mock transducer, a marker, and a cardboard to mimic ultrasound scanning. The system is portable and convenient for trainees. However, its tracking technique is not publicly available since it is a commercial simulator. Nevertheless, Optical Tracking Systems (OTS) are often used as surgical tools to register computer-added images for surgical navigation, as seen in Lee et al.'s work [13]. These commercial OTS [10] [14] [15] often use other techniques, such as near-infrared light or senor-based calibration, to enable camera-based tracking. Therefore, they are not portable and unavailable for any open-source editing to fit the needs of other medical simulations.

Overall, optical based tracking framework requires either multiple cameras or involves other sensor devices to improve the tracking accuracy. To our best knowledge, a single camera based optical tracking framework that supports open-source development of medical simulation does not exist.

III. METHOD

Our method used portable devices and open-source CVT. Any simulation prototype can adopt the approach. In this section, we will detail our approach in terms of experiment setup, including adopted devices, a unique marker design, and the implementation of the OpenCV tracking software.

A. *Experiment Setup*

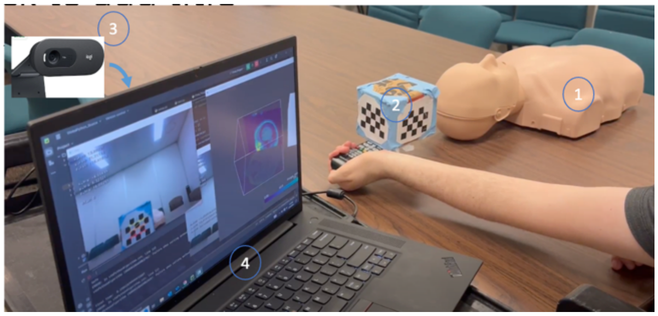


Figure 1. Experiment Setup: 1) Phantom; 2) Cubic Shape Marker Attached to A Mock Transducer; 3) Camera; 4) Laptop.

The proposed method in this study used four pieces of equipment to develop a running prototype, as seen in Figure 1:

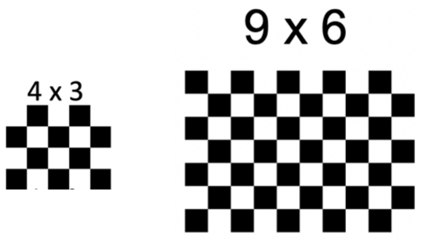
1) An upper body mannequin as a phantom
2) Four chessboard markers glued to a cubic shape cardboard and attached to the top of a mock transducer
3) A Logitech web camera (hidden behind of the laptop) with a resolution of 720p/30fps
4) A Lenovo laptop with AMD Ryzen 7 5700u CPU, AMD Radeon Graphics, and 40 GB memory to run the OpenCV method for computer vision-based tracking

B. Tracking Method

Our tracking method used OpenCV [16], an open-source computer vision library. The original OpenCV program was designed to use several cameras from different angles to track a single marker. The setting can easily capture the translation and rotation of a marker based on streaming tracking of all cameras. However, in order to provide a portable framework, we only adopted one camera and designed a cubic-shape marker box in our setting.

1) Marker

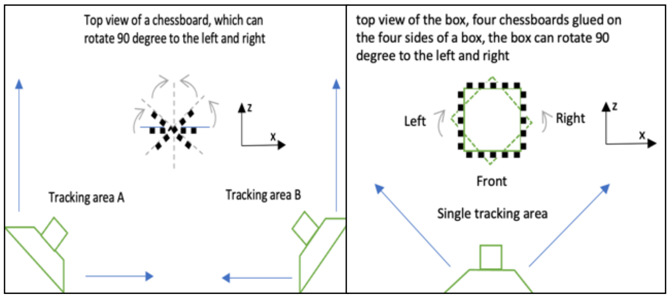
The cubic-shaped box consists of four markers glued on the four sides (left, right, front, and back) of a box, as seen in Figure 1 (2). This unique marker design can accommodate the development of a portable system. Each marker is measured as 4 by 3 in terms of the number of Internal Corner Points (ICP), points where the black square corners connecting with each other [17]. This size is cropped from the original marker measured 9 by 4 ICP in order to fit a hand-held medical device. Figure 2 shows the comparison of the two-marker sizes.



(a) (b)

Figure 2. Markers a) size 4x3; b) size 9x6

OpenCV is designed for multiple tracking of several cameras. Therefore, many sample programs consist of multiple streaming tracking. Figure 3 shows the top view (xz-plane) of two experiment settings. The camera lens is directly facing the xy-plane. The y coordinate is perpendicular to the xz plane. As seen in Figure 3 (a), two cameras placed diagonally (45-degree angle) facing against a marker can easily monitor the rotation of the marker along the y coordinate. The marker can rotate 90 degrees to the left or right, while cameras take turns to track the rotation when the marker is no longer seen by one camera in a 180-degree tracking angle. However, two camera-based tracking would be inconvenient for a portable system design. After all, we came up with the design of the box shape marker, which allows us to track the rotation of a marker easily when there is only one camera tracking the marker at a limited tracking angle. For example, as seen in Figure 3(b), when the front marker is rotating to the left along the y coordinate at approximately 45 degrees, it starts to lose the tracking of the camera since its chessboard is no longer seen by the camera lens. However, at this point, the right-side marker on the box becomes available for camera detection, which is rotating toward the camera with a less than 45-degree angle. The system would then add a 90-degree angle change to the rotation vector as a subtraction factor. This solution permits single-camera tracking while producing a fast and accurate rotation vector for the study. Similarly, when tracking the markers glued on the top and bottom of the box, the system can easily monitor the rotation along the x coordinate.



(a) (b)

Figure 3. (a) Multiple camera and one marker tracking (b) One camera and box shape marker-based tracking

Figure 4 presents the real case scenarios for our experiment. The camera is facing the front marker initially. When a subject was rotating the marker along the y coordinate (perpendicular to the mannequin's face) to the left side of the marker, the right side marker became available for tracking. At this time, we will update the subtraction factor to indicate the rotation angle of -90 degrees.

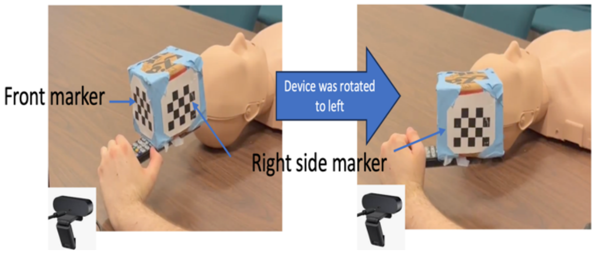


Figure 4. The Setting of One Camera Tracking Experiment

2) Tracking Framework

Figure 5 presents the system diagram of the implemented tracking software using OpenCV. The framework system started with camera-based calibration to learn the orientation of the phantom. The marker was then converted to a grayscale image for chessboard recognition using the method findChessBoardCorners(). We adopted three flags in this method call: 1) threshold to convert the image to black and white only, 2) normalize the gamma of the image in order to find the chessboard easier, and 3) add a fast check that speeds up the chessboard detection process. If a chessboard is detected, the findChessBoardCorners() method searches 6 ICPs during the stream tracking. Once the 6 ICPs were found, they went through the cornerSubPix() method to return a more accurate ICP measurement using the image gradient described in Forstner's method [18]. Afterward, the calibrateCamera() method runs the solvePnP algorithm, which will estimate the marker's pose using the 6 processed ICPs and their corresponding image projection points. The algorithm runs Levenberg-Marquardt optimization [19] iteratively, which helps to calculate such a pose by minimizing the reprojection error. This approach will finalize the pose by generating the translation and rotation vectors. Meanwhile, our script checks the rotation vector along the y coordinate to update the subtraction factor. The framework produces a translation vector, a rotation vector, and a subtraction factor. These three outputs can be used by any simulation system that needs a tracking feature.

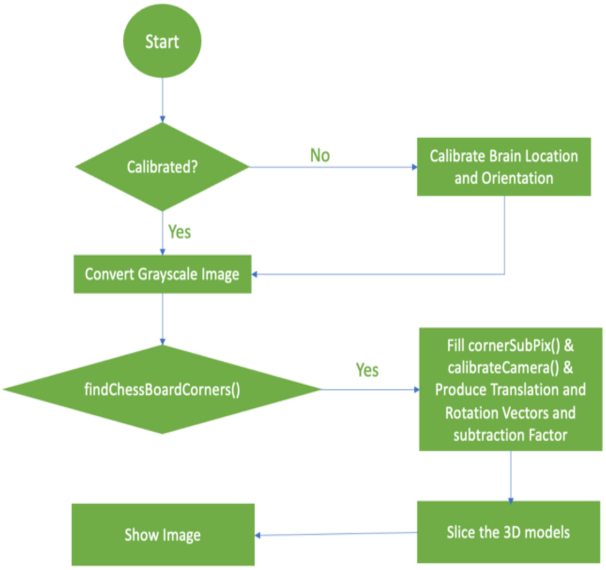


Figure 5. OpenCV Program Diagram

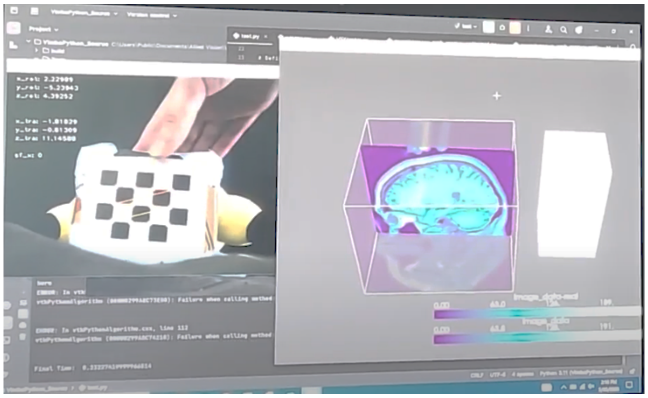
IV. A CASE STUDY

The framework was applied to a simulated cranial (head) ultrasound to verify its usage. We adopted a 3D brain mesh model from PyVista [20] in VTK format. To properly slice the 3D model based on the tracked chessboard that is orthogonal to the slice, the ManySlicesAlongPoints() method [21] from PyVista was used. We used the translation vector [X_t, Y_t, Z_t] and rotation vector [X_r, Y_r, Z_r] obtained from the OpenCV tracking program to calculate a required path [X, Y] as shown below:

$$X = \left(\tan\left((X_r - 90)\frac{\pi}{180}\right) \times Y\right) + X_t;$$

$$Y = \text{points: } 0 - 216;$$

Then, the SliceAlongManyPoint() method produced a line/path based on the calculated [X, Y] and sliced the 3D brain model that is perpendicular to the path. As seen in Figure 4, the face of a mock transducer, facing against the top head of the mannequin, is orthogonal to the marker planes on all four sides. Therefore, an orthogonal slice is the proper setting for our experimental setup.



```
x_rot: 2.22989
y_rot: -5.23943
z_rot: 4.39252

x_tra: -1.81829
y_tra: -0.81309
z_tra: 11.14588

sf_x: 0
```

Figure 6. A prototype for a simulated cranial ultrasound

In Figure 6, the simulated ultrasound prototype presents three windows side by side. The top left window showed the real-time tracking of the box marker by highlighting the tracked marker; the bottom left window displayed the running script of the program by outputting both translation and rotation vectors; the top right window produced a slice of the 3D brain model for the simulated cranial ultrasound based on the location and orientation of the mock transducer/box marker. A white cubic on the right side of this output window indicates the box marker's corresponding position toward the calibrated brain in virtual space. Note that our prototype has not yet presented its form in ultrasound texture due to time constraints; however, our previous ultrasound image simulation techniques [22] can be easily adopted by the platform to fulfill the case study of the simulation. Nevertheless, the presented outcome does not impact our research goal for a tracking framework based on CVT to support medical simulation.

V. RESULTS

The proposed tracking framework can produce real-time tracking based on our simulated cranial ultrasound experiments. The adopted devices work well to support the newly developed tracking technique. The simulated ultrasound is presented instantly at a 30 FPS rate according to the movement of the mock transducer. The rotation was captured accurately at 90 degrees to the left and right along the y coordinate using the subtraction factor.

The dual windows design (tracked marker and simulation production), as seen in Figure 6, enables us to explore the tracking space fully. The showing of the mock transducer in the top right window presents the true location of the device in the physical world. The tracked marker reflected the rotation angle. The script is running effectively to support the movement of the transducer. The most important evaluation is that the generated cranial ultrasound slice reflects the accurate orientation and translation of the brain model in the phantom

The simulation framework can be easily adopted by any simulation research. We plan to release the software as a free resource for academic usages.

ISBN: 978-1-68558-197-8

VI. CONCLUSION

The proposed framework developed an innovative approach to track a box shape marker in order to build a portable tracking system using one web camera. The system can be used by any simulation-related prototype, providing conveniences to medical students who can practice training in any setting without labs. The case study presented an ultrasound simulation setting, which has approved the effectiveness of the framework that offers real-time stream tracking. Although the simulation prototype is not ready for ultrasound training, the presented prototype provides us with a way to evaluate the framework regarding tracking accuracy.

The framework can be adopted by any simulation research to monitor the orientation and rotation of a movable medical device. It is particularly useful to augment a virtual object based on the movement of a physical object in any domain setting.

ACKNOWLEDGMENT

We would like to thank US Department of Education for GAANN award #P200A190033 to support this research work.

REFERENCES

[1] A. Nahum. *Patient Safety Representatives Unite to Implement Global Patient Safety Action Plan in the United States*. Centers for Disease Control and Prevention, City, 2021.

[2] T. Blum, A. Rieger, N. Navab, H. Friess and M. Martignoni. "A Review of Computer-Based Simulators for Ultrasound Training". *Simulation in Healthcare*, vol 8, no. 2, pp. 98-108, 2013.

[3] Schallware *Schallware Ulrasound Simulator*. Schallware, City, 2024.

[4] SurgicalScience *Ultrasound Mentor Platforms*. Surgical Science, City, 2024.

[5] Menache, A. *1 - Motion Capture Primer*. Morgan Kaufmann, City, 2011.

[6] Intelligent Ultrasound *ScanTrainer: Simulation in Transvaginal and Transabdomnial Ultrasound*. Intelligent Ultrasound, City, 2024.

[7] SurgicalScience *VR add on for ultrasound mentor*. Surgical Science, City, 2024.

[8] D. Markov-Vetter, J. Muehl, D. Schmalstieg, E. Sorantin and M. Riccabona. "3D augmented reality simulator for neonatal cranial sonography". *Comput Assist Radiol Surg*, vol. 1230, pp. 483-487, 2001.

[9] Z. Duan, Z. Yuan, X. Liao, W. Si and J. Zhao. "3D tracking and positioning of surgical instruments in virtual surgery simulation". *Journal of multimedia*, vol. 6, no. 6 2011.

[10] NaturalPoint Inc. *OptiTrack*. NaturalPoint Inc., City, 2024.

[11] T. Ameler, K. Blohme, L. Brandt, R. Brüngel, A. Hensel, L. Huber, F. Kuper, J. Swoboda, M. Warnecke and M. Warzecha. "*A comparative evaluation of steamvr tracking and the optitrack system for medical device tracking"*. IEEE, City, 2019.

[12] Medge Platforms Inc. *Volutracer O.P.U.S.: Optical Positiong Ultrasound Simulator*. Medge Platforms, Inc, City, 2023.

[13] S. Lee, H. Lee, H. Choi, S. Jeon and J. Hong. "Effective calibration of an endoscope to an optical tracking system for medical augmented reality". *Cogent Engineering*, vol. 4, no. 1, pp. 1359955, 2017.

[14] Northern Digital Inc. *Optical Navigation*. NDI, City, 2022.

[15] Advanced Real Time Tracking GmbH & Co. KG *CAPTA Tracking technolgy*. Advanced Realtime Tracking, City, 2024.

[16] OpenCV *Open Source Computer Vision: OpenCV modules*. OpenCV, City, 2024.

[17] OpenCV *findChessboardCorners(): OpenCV Document*. OpenCV, City, 2024.

[18] OpenCV *cornerSubPix(): OpenCV Document*. City, 2024.

[19] OpenCV *Perspective-N-Point Pose (PnP) Computation: OpenCV Doc*. City, 2024.

[20] PyVista *PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK)*. The PyVista Developers, City, 2023.

[21] B. Sullivan. *Many Slices Along Points - PVGeo Documentation*. GenDocs, City, 2020.

[22] B. Sun. and F. McKenzie. "Real-time sonography simulation for medical training". *International journal of education and information technologies*, vol 5, no. 3, pp. 328-335, 2011.