

Ultrasound Augmentation: Rapid 3-D Scanning for Tracking and On-Body Display

Maeliss Jallais, Hastings Greer, Sam Gerber, Matt McCormick, Deepak Chittajallu, Neal Siekierski, and Stephen Aylward

Kitware Inc, Carrboro NC 27510, USA

Abstract. By using a laser projector and high speed camera, we can add three capabilities to an ultrasound system: tracking the probe, tracking the patient, and projecting information onto the probe and patient. We can use these capabilities to guide an untrained operator to take high quality, well framed ultrasound images for computer-augmented, point-of-care ultrasound applications.

1 Motivation

We have developed algorithms that analyze ultrasound signals to detect internal bleeding [1] and increased intracranial pressure (increased optic nerve sheath diameter) [2]. These algorithms are part of a computer-augmented, point-of-care ultrasound (CAPCUS) system that is intended to aid in the triage of patients with abdominal or head trauma at the scene of an accident, helping emergency medical service (EMS) personnel decide when to order expedited transport and initiate life-saving measures in the field.

One of the critical, remaining challenges with deploying our CAPCUS system is informing the EMS personnel on where to place and how to manipulate an ultrasound probe at various anatomic locations in order to, for example, thoroughly examine the regions of the abdomen where blood tends to pool. This positioning task requires extensive anatomic and ultrasound training, beyond what most clinicians and/or EMS personnel typically receive.

This paper presents an innovative device that combines patient and probe tracking with augmented reality to create an "augmented ultrasound" system. It is one of a variety of methods and user interfaces that we are investigating to guide EMS personnel in probe placement, to inform them of data quality and to convey diagnoses.

2 Augmented Ultrasound

Rather than displaying data and instructions on a screen, our augmented ultrasound system proposes to concisely convey guidance and diagnoses by projecting instructions onto the surface of the patient. There are several advantages. Unlike a screen, a projection does not require the operator to look away from the

patient. In addition, when combined with a tracker, the projection can provide directions in absolute terms. For example, a CAPCUS system that uses our algorithm for on body display would be able to direct the operator to "Place the scanner on the bullseye," instead of relying on more abstract instructions such as "Place the scanner on the right side of the patient's chest, half-way between the nipple and the shoulder." Furthermore, instructions can be updated based on the tracked movement of the ultrasound probe and the images/anatomy captured in the ultrasound data, ensuring coverage of an anatomic area or flashing a warning icon when the probe is not making sufficient contact with the skin.

The augmented ultrasound system must be able to track the ultrasound probe and the surface of the patient's body and then use the generated model of the body and the position of the probe to accurately project graphics/instructions onto the patient's body. Furthermore, the augmented ultrasound system must work in sunlight and in rugged environments with minimal set-up time so that it can be applied at the scene of an accident.

In the next section, we discuss our system for forming a three-dimensional (3D) model of a scene using a laser projector and a high-speed camera. Then, we describe how we track objects in that scene.

3 3D Scene Modeling

To form a 3D model of a scene, we replicated and extended the system described in a paper by researchers at Carnegie Mellon University that performed depth from structured light reconstruction using a portable projector, which uses laser projection technology, and a high-speed camera. (See Fig. 1, adapted from [3].) The laser projector draws only one line of a projected image at a time, but it does so fast enough that the human eye only sees the complete image. The high-speed camera is capable of precise timing and fast shutter speed. These properties allow the camera to take a picture as the projector draws a specific line of the image. Through a one-time calibration of the camera and the projector, and by using structured-light reconstruction algorithms, each point on a projected line that is seen by the camera can be efficiently and accurately triangulated to a point in 3D space (i.e., each illuminated point seen in a camera image is at the intersection of the plane/line of light emitted by the projector and the corresponding pixel/line viewed by the camera). The camera/projector calibration process is adapted from the method developed at Brown University [?], which is based on OpenCV and uses the camera application programming interface (API).

Our system is composed of a Grasshopper3 camera and a SONY mobile projector MP CL1. With a resolution of 2048 x 1536, we can use the camera at a frame rate of 300 fps. The update rate per image of the laser projector is 60Hz in scan line order. Given its resolution of 1920 x 720, it projects 43,200 lines/s.

The augmented ultrasound system works under varied lighting conditions for three reasons. First, the shutter speed simplifies the detection of which pixels are illuminated by the projector. It suppresses ambient light, even direct sunlight, simply by remaining open for no more than a tiny fraction of a second.

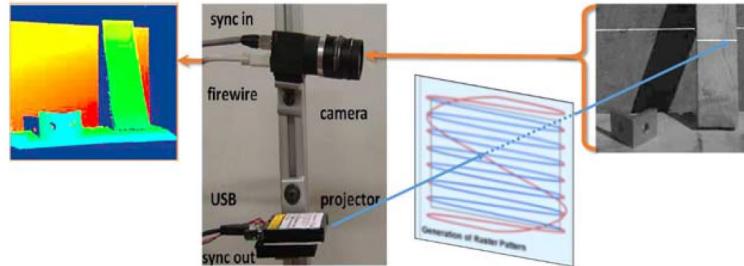


Fig. 1: The augmented reality system uses a pico projector (based on laser projection technology) to display an image in a scene. A high-speed camera sees the individual lines being drawn by the projector to create the image. Via appropriate camera calibration, custom circuitry to detect the vertical reset of the laser and depth from structured lighting algorithms, a 3D model of the scene can be formed at very high resolution in about one second. Figure adapted from [3].

Table 1: Hardware used, with prices

Device	Manufacturer	Price
Grasshopper 3 GS3-U3-32S4C-C	Point Grey	\$975.00
MP-CL1 Projector	Sony	\$349.00
Photodiode, BPW34S	Vishay Semiconductor Opto Division	\$6.66

This benefit arises because the laser projector illuminates each point very brightly, but only for a fraction of a millisecond. Alternatively, ambient sources like the sun illuminate continuously at a lower intensity. Second, we use background subtraction to further reduce the effect of ambient light. In particular, from each image that is taken, we subtract an image from when the laser is in a different location. Third, we apply a custom method that quickly detects the brightest point on a vertical scan line in a grayscale image.

The most difficult challenge with our low-cost, compact ultrasound augmentation system was that the horizontal sync signal produced by the projector/HDMI protocol was not phase-locked to the actual vertical movement of the laser. For this reason, we opted to trigger the camera off the light emitted by the projector, using a photocell and a pair of Operational Amplifiers. This horizontal sync is critically important because it is needed to determine which line is being projected at any point in time. The amplifier circuit converts the noisy photocell into a digital trigger signal. See Fig. 2.

Example reconstructions generated by the system are shown in Fig. 3.

4 Object Tracking

Having formed a 3D model of a scene, the next step in ultrasound augmentation is to track an ultrasound probe. To that end, we mounted a multi-cubed, color-

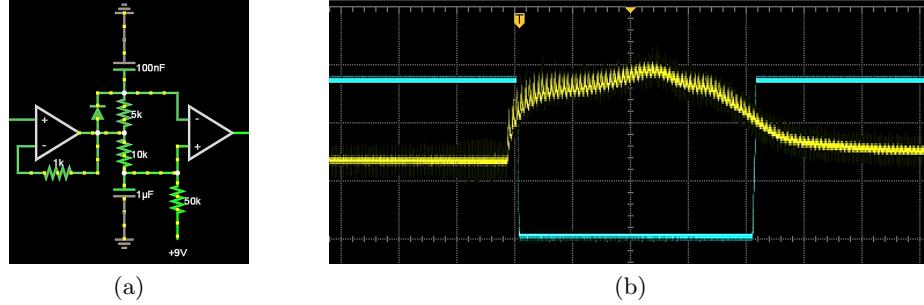


Fig. 2: (a) Circuit diagram of the photo detector system used to determine when the laser returns to the upper left and begins drawing a new image (approximately every 1/60th of a second). (b) Oscilloscope-measured photocell output (yellow) and circuit output (blue) signals. The circuit output's falling edge is used as a camera trigger.

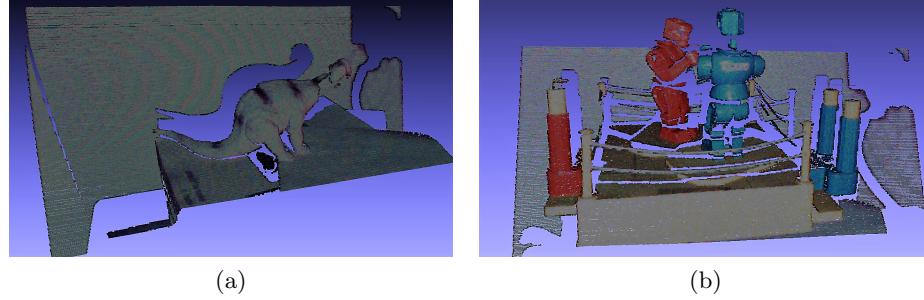


Fig. 3: Examples of 3D scenes captured by our system.

tagged marker (designed by InnerOptic) on the probe. The colors of the cube are red, blue and green. They are optimized for the response curves of the color detectors in the camera. By determining the intersection of three adjacent faces on the cubes, we can determine the position and the orientation of the ultrasound probe.

To simplify and speed the tracking algorithm, instead of estimating the specific location of each cube face, we estimate the planes that contain each of the faces of the cube and then solve for the intersection of those planes. The plane detection method begins with color pixel detection, using statistical models of the appearance of the cube faces under a range of lighting conditions. Depending on the image from the projector, the ambient light and the adjacent objects in the environment, extraneous pixels may be incorrectly identified as cube pixels based on color alone. To eliminate extraneous pixels, we compute the centers of gravity for each target color using robust statistics. Knowing the expected

position of each face relative to the others, we can redefine the estimates of the centers of gravity by eliminating colored pixels that are inconsistent with the centers of gravity or with the expected relative positions of each face. To further reduce the influence of extraneous pixels, we also estimate the equations of the planes using the random sample consensus (RANSAC) algorithm. It randomly picks three points, computes the plane defined by them and then scores that plane based on how many other cube points are included in that plane. Ultimately, the planes with the best scores are chosen. The intersection point of the three chosen planes is then used to define the position and the orientation of the probe in space. A sample result of three detected planes intersecting the scene is shown in Fig. 4.

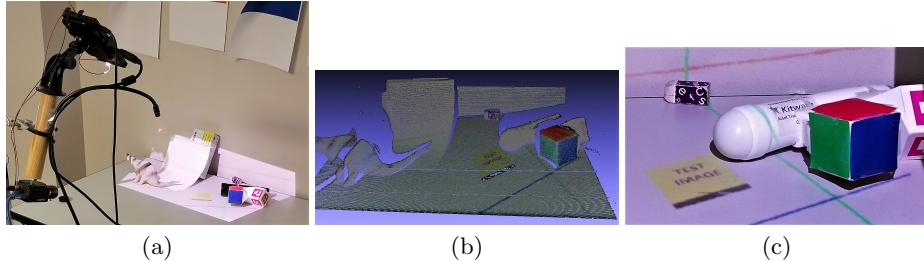


Fig. 4: (a) The setup of the system, with the projector at the top left and the camera and the bottom left. (b) The collected point cloud is shown in white. The intersection of the estimated red, green and blue planes with that point cloud are shown in color. This indicates that the faces of the cube are well represented by the estimated planes. (c) A close-up of the scene shows the projected colored planes and a projected yellow instruction card (labeled "test image") that automatically follows the tip of the ultrasound probe.

Our camera is capable of up to 300 frames per second, and our projector runs at 60 sweeps per second. However, running the camera at full frame rate does not result in an even coverage of the scene by the captured images. In addition, some frames are not useful because they would be captured as the laser is returning to the top of the screen. Finally, it is only possible to trigger the camera at evenly spaced intervals, a variable delay after our optical detection circuit fires. In our system, the camera fires three times every time the optical detection circuit is activated by the laser, leading to a speed of 180 lines per second, which is equivalent to 4 reconstructions per second.

Our system can trade frame rate for resolution of the point cloud, so we tested the precision of tracking for a variety of point cloud resolutions. The results of this test are shown in figure 5. We found that precision did not improve past 45 lines per frame, achieving a standard deviation of 1.2 mm per axis. We also investigated the shape of this error, finding it to be roughly isotropic, as shown if figure 6.

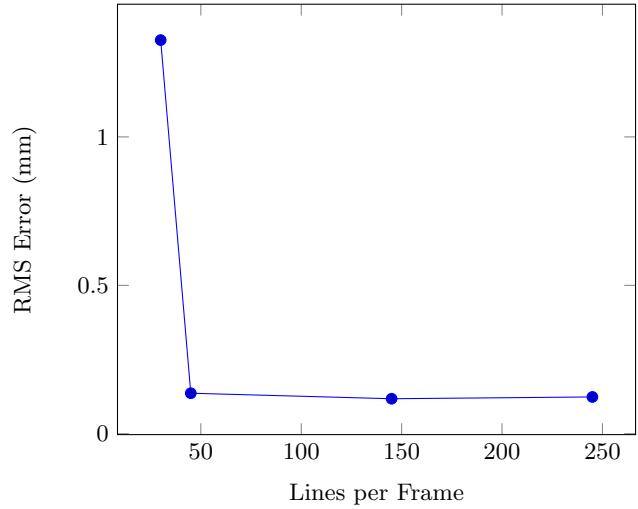


Fig. 5: The behavior of error as the number of lines per reconstruction is increased at a distance of 60 cm

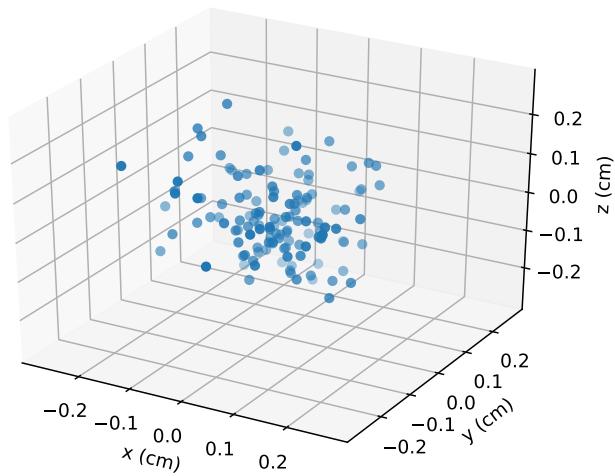


Fig. 6: Representative error in x, y, and z at a distance of 60 cm with 45 lines per reconstruction

Table 2: Accuracy at 60cm and 400Lux on 300 acquisitions

Lines on cube	Std Deviation		
	x	y	z
4	3.001cm	2.248cm	6.846cm
5	0.884cm	1.134cm	2.937cm
6	0.301cm	0.875cm	3.075cm
7	0.229cm	0.570cm	1.000cm
8	0.105cm	0.455cm	1.234cm
10	0.074cm	0.087cm	0.084cm
20	0.065cm	0.077cm	0.062cm
30	0.057cm	0.070cm	0.062cm
40	0.057cm	0.068cm	0.055cm

5 Projecting onto the Scene

Once the complete system has captured a point cloud, located the ultrasound probe, and determined a diagnosis from the ultrasound data, the system must display the results onto the patient’s skin. The position of the displayed results should be relative to the probe and undistorted, i.e., “rectified” with respect to the surface onto which they are being projected. To aid in this goal, display elements, such as text annotations, images, and arrows, are specified in 3-D world coordinates. We support commands such as ”Color yellow all points within a 4 cm radius of the world-point (60, -12, 3)” or ”Project the text ‘Poor Contact: Apply pressure’ at the point 5 cm in front of the ultrasound probe.”

To achieve this effect, we use a stored, high resolution point cloud known as the ”canvas” that is captured during initialization and updated only when the rapidly acquired reconstructions begin to deviate from the canvas. The canvas point cloud has up to 700 lines, whereas rapid point clouds used for tracking have only 45 lines.

Graphics primitives such as spheres or images as well as ray tracings and volume renderings are drawn onto the canvas, by updating the color associated with each point. Next, the canvas is transformed by the ”Camera Matrix” of the projector, which takes points from world coordinates to pixel coordinates on the projected image. Finally, the points of the canvas point cloud are splatted onto a bitmap and projected. This causes each point on a surface in the scene to be illuminated with the color light of its corresponding approximate nearest point in the canvas pointcloud.

The four frames per second reconstruction rate may seem slow, but our system does not need to rely on a high frame rate to avoid nausea or support smooth navigation. Only the annotations, not the full scene, are being driven by the system. Additionally, the ultrasound probe is typically a slow moving object and the environment is generally relatively static.

6 Conclusion

Ongoing work focuses on quantifying the performance of this algorithm. Experiments presented in this paper indicate 2mm consistency within a 30cm x 60cm x 60cm operating environment, updated at four frames per second. In other experiments, not presented in this paper, the system has been shown to be insensitive to a wide range of ambient light brightness and to the image being projected into the scene.

In future work, the probe tracker data will be used, in combination with a custom image reconstruction technique, to compound a sequence of imprecisely and sparsely tracked ultrasound images into a complete 3D Volume. In that volume, vessels, for example, could be identified and the scene could be augmented with instructions on where to insert the needle, at what angle, and to what depth, so as to best intersect with a target vessel. This is illustrated in Fig 7.

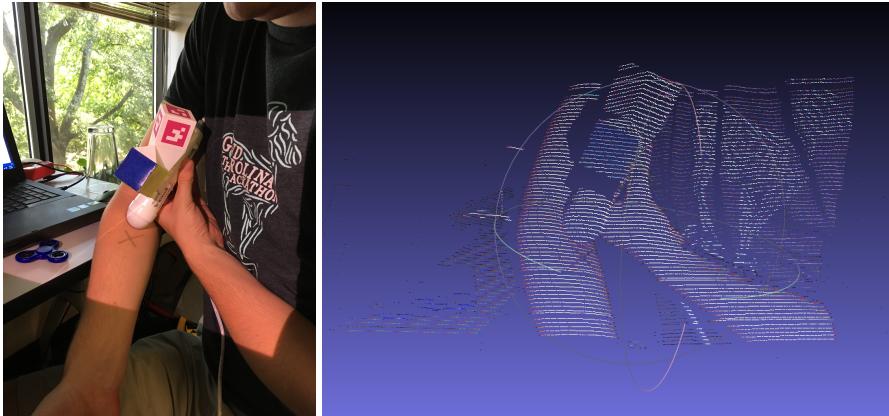


Fig. 7: An illustration of the end goal of the system. A target (black X) is projected onto the skin as the camera computes depth from structured light using the raster lines of the projected image as that image is drawn. Ultrasound can be used to detect peripheral vessels, select needle insertion locations and verify needle placement patency.

Acknowledgments

This work was funded, in part, by the following grants.

- NIH/NIBIB: In-field FAST procedure support and automation (R43EB016621)
- NIH/NINDS: Multimodality image-based assessment system for traumatic brain injury (R44NS081792)
- NIH/NIGMS/NIBIB: Slicer+PLUS: Collaborative, open-source software for ultrasound analysis (R01EB021396)

References

1. Stephen R Aylward, Matthew McCormick, HJ Kang, Sharif Razzaque, Roland Kwitt, and Marc Niethammer. Ultrasound spectroscopy. In *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, pages 1013–1016. IEEE, 2016.
2. Samuel Gerber, Maeliss Jallais, Hastings Greer, Matt McCormick, Sean Montgomery, Bradley Freeman, Deborah Kane, Deepak Chittajallu, Neal Siekierski, and Stephen Aylward. Automatic estimation of the optic nerve sheath diameter from ultrasound images. In *MICCAI Workshop*, 2017.
3. Christoph Mertz, Sanjeev J Koppal, Solomon Sia, and Srinivasa Narasimhan. A low-power structured light sensor for outdoor scene reconstruction and dominant material identification. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pages 15–22. IEEE, 2012.
4. Daniel Moreno and Gabriel Taubin. Simple, accurate, and robust projector-camera calibration. In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on*, pages 464–471. IEEE, 2012.