

Ultrasound tracking using ProbeSight: Camera pose estimation relative to external anatomy by inverse rendering of a prior high-resolution 3D surface map

Jihang Wang¹, Chengqian Che², John Galeotti^{1,3}, Samantha Horvath³, Vijay Gorantla⁴, George Stetten^{1,3}

¹ Department of Bioengineering, University of Pittsburgh

² Department of Biomedical Engineering, Carnegie Mellon University

³ Robotics Institute, Carnegie Mellon University

⁴ Division of Plastic Surgery, University of Pittsburgh Medical Center

Abstract

This paper addresses the problem of freehand ultrasound probe tracking without requiring an external tracking device, by mounting a video camera on the probe to identify location relative to the patient's external anatomy. By pre-acquiring a high-resolution 3D surface map as an atlas of the anatomy, we eliminate the need for artificial skin markers. We use an OpenDR pipeline for inverse rendering and pose estimation via matching the real-time camera image with the 3D surface map. We have addressed the problem of distinguishing rotation from translation by including an inertial navigation system to accurately measure rotation. Experiments on both a phantom containing an image of human skin (palm) as well as actual human skin (fingers, palm, and wrist) validate the effectiveness of our approach. For ultrasound, this will permit the compilation of 3D ultrasound data as the probe is moved, as well as comparison of real-time ultrasound scans registered with previous scans from the same anatomical location. In a broader sense, tools that know where they are by looking at the patient's exterior could have broad beneficial impact on clinical medicine.

1. Introduction

Ultrasound (US) is one of the most widely used medical imaging technologies [1]. Diagnostic ultrasound is low-cost, real-time, portable, and completely safe, with neither ionizing radiation nor strong magnetic fields. Ultrasound is also interactive, which aids real-time usage but poses challenges for later review. During imaging, the operator's mind continually tracks the location of the ultrasound images relative to the patient's external visible anatomy. However, after a free-hand scan is finished, even the most skilled clinicians find it challenging to effectively localize

the 2D images within the patient's anatomic 3D context. The existing techniques for probe localization, however, require external tracking devices, making them inconvenient for clinical use. In addition, the probe pose is typically measured with respect to a fixed coordinate system independent of the patient's anatomy, making it difficult to correlate ultrasound studies across time. This paper concerns the development and evaluation of a novel self-contained ultrasound probe tracking system, which navigates the probe in patient space using camera pose estimation relative to the anatomical context.

This paper is organized as follows. Section 2 reviews prior work on probe tracking systems. The technical approach is described in Section 3. Section 4 reports the experimental results, and we discuss our conclusions in Section 5.

2. Related Work

External equipment can be used to track and record an ultrasound probe's position, to aid ultrasound registration and post-acquisition review. There are two common types of tracking systems, optical and electromagnetic (EM). These systems use optical markers (active or passive) or sensor coils attached to the probe [2]–[7], but few researchers attempt to also attach these to the patient's skin. Without also tracking the patient, coordinates are reported in a fixed and arbitrary coordinate system rather than one relating to the patient's anatomy. Even if related initially to anatomical coordinates, errors are introduced when the patient moves. Optical tracking systems suffer from several other problems, including a limited field of view (FOV), the need for a clear line of sight, and a lack of orientation accuracy with small markers. EM tracking is sensitive to ferromagnetic objects and electronic equipment near the tracker, limiting its clinical usage. Furthermore, both types of tracking systems are bulky and not self-contained, making them inflexible for clinical use.

Self-contained methods for US probe localization by directly attaching video camera(s) on the US probe have been investigated by a number of researchers over the last decade. This theoretically permits determining the probe's pose relative to the patient's anatomy without requiring any external tracking equipment. Chan, et al., used stereo cameras mounted on the US probe with computer vision methods to determine needle location relative to probe [8]. Raffi-Tari, et al., attached strip pattern markers to the skin along patient's spine and used a single camera mounted on an ultrasound probe to register the US probe to the patient's anatomy, permitting accumulation of 3D ultrasound [9]. More recently, Sun, et al., developed a low-cost and unobtrusive method for 6-DoF probe tracking in freehand 3D US, by attaching trackable random binary patterns on the skin [10]. However, such artificial surface markers can be problematic during clinical procedures, especially if they must remain from one scan to the next. Longitudinal imaging studies can require co-registered images over the course of months or even years. Surface markers can also block the passage of US into patient, and may degrade or be distorted when exposed to US gel. Sun et al. have developed a camera based system that tracks skin features directly, similar to our work described here, except that they still require an artificial marker to be affixed to the skin for scale calibration [11].

Tracking skin features directly inherently avoids the problems mentioned above. We previously developed a new technique called ProbeSight for tracking an ultrasound probe by matching a probe-mounted camera's images of the patient's external anatomy against an (inverse) rendered high-resolution surface map model. The surface map is acquired in advance using a multiple-camera imaging device designed for reconstructive surgery. We previously used a plastic model animal containing rich skin features as a target to demonstrate the accuracy of the device direct approach, independently validating each degree of freedom (DOF) of camera motion [12].

For the present paper, we now track the US probe in all 6-DOF simultaneously, optimizing our inverse rendering through the incorporation of experimental software, OpenDR, to solve for camera pose against the surface map. OpenDR is an approximate differentiable renderer initially introduced by Loper and Black for estimating human body shape from Kinect depth and RGB data [13]. In addition to repurposing OpenDR for our unique application, we also use an inertial navigation system (INS) to continually measure precise angular pose, permitting disambiguation between pixel shifts caused by translation and rotation. We tested our system on targets including (1) a flat phantom with a printed image of human skin and (2) an in-vivo human hand.

Although ProbeSight has broad potential uses across clinical ultrasound, we have chosen a compelling niche for our initial application: the monitoring of rejection after hand transplant surgery. This application provides complex anatomic shape and skin texture, a relatively shallow depth to relevant nerves and vasculature, and a need for precise ultrasound localization across long-term monitoring.

3. Technical Approach

The current version of ProbeSight contains an Ethernet-powered machine vision camera (Prosilica GT1290C; Allied Vision Technologies, Canada) and an INS (VN-100 Rugged; VectorNav Technologies, USA), both rigidly mounted onto the ultrasound probe using a custom machined probe bracket. For prototype experimentation, an additional slide rail adjusts the distance between the camera and skin surface (Figure 1). Our ProbeSight apparatus is attached to the 30-70MHz linear array transducer of a high-frequency ultrasound imaging system (VisualSonics Vevo 2100; FUJIFILM VisualSonics, Inc, Netherlands). Validation is provided by a Micron Tracker optical tracking system (Sx60; Claron Technology, Canada), with passive optical markers attached to both ProbeSight and to the targets.

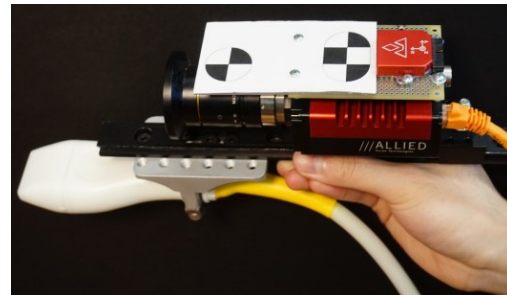


Figure 1: Ultrasound probe with mounting bracket and slide rail for the camera and INS.

Image acquisition and pre-processing software was written in C++ (Visual Studio 2010), using several libraries, including the Open Graphics Library (OpenGL), Open Source Computer Vision (OpenCV) library, Micron Tracker Software Development Kit (SDK), and the VectorNav INS data export SDK. Inverse rendering and optimization software was written in Python (Version 2.7.6), using both the OpenDR and Qt graphical user interface (GUI) framework.

3.1. Components and Coordinate Systems

Each component of ProbeSight has its own native coordinate system, requiring transformations between camera coordinates C_C and those of each of the other components. Figure 2 shows these coordinate systems and

the transformations between them. Since the INS is mounted on the camera, its rotations are inherently in camera coordinates C_C . Other important coordinate systems include anatomy (prior 3D scan) C_A , validation tracker C_T , and the validation optical marker attached to ProbeSight C_M . In order to provide simulated camera images of the 3D surface map to the OpenDR pipeline, all coordinates need to be transformed from anatomy coordinates C_A to camera coordinates C_C . Using pre-calibrated camera intrinsic parameters (camera matrix and distortion coefficients), we can then re-project the 3D surface map to simulate what would be seen by the real camera from any possible point of view, for comparison with the actual camera image to find the matching pose.

Diagram illustrating the experimental setup for motion capture and navigation. The setup includes a Vectra M3 system (tracker camera and cameras) and an INS camera with an optical marker and US probe. The hand is positioned on a molded hand splint. The diagram shows the coordinate systems and transformations: $T_{T \rightarrow M}$, $T_{M \rightarrow T}$, and $T_{M \rightarrow C}$.

subject is used to keep the gross hand pose constant between the prior scan and the experiment, as shown in Figure 2. The 3D coordinates of the exterior anatomy P_A are originally in the preoperative 3D scanner’s coordinates C_A . We perform a one-time transformation from C_A to C_T , denoted $T_{A \rightarrow T}$, which requires touching white-dot fiducials on the hand splint with a tracked tooltip, analogous to the phantom experiment. To fully transform P_A into the camera’s coordinates, we use $P_C = T_{M \rightarrow C} M^{-1} T_{A \rightarrow T} P_A$.

subdermal veins, which provide additional cues for tracking. Having all these features improves the results of subsequent optimization using OpenDR.

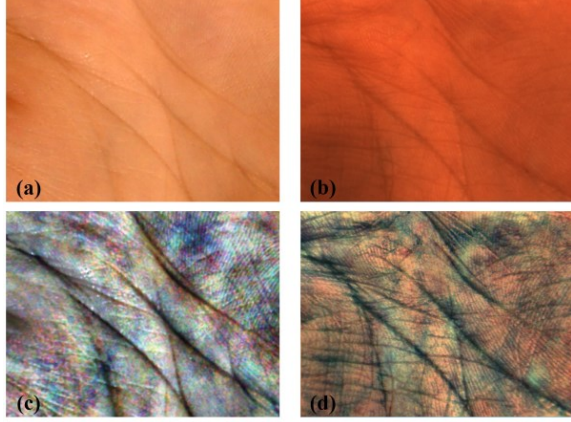


Figure 3a and 3b: The original re-projected surface map and camera image. **Figure 3c and 3d:** The color enhanced re-projected surface map and camera image.

3.3.2 Image Foreground-background Segmentation

Background information in either the 3D surface map or the camera image that does not correspond to the other image can interfere with image matching, reducing OpenDR’s accuracy. Therefore, during preprocessing we segment the foreground (anatomy) pixels from both the original 3D surface map and the raw camera images. In the case of the 3D surface map, this segmentation need only be performed once, and is simplified by knowing the color of the foreground vs. the background beforehand. For the raw camera images, because the hand splint is blue, a typical “blue screen” segmentation can be performed using color-thresholding in HSV (Hue, Saturation, Value) space. The resultant foreground masks are then applied to the CLAHE-enhanced image.

3.3.3 Cropping and Downsampling

After each 1936×1456 pixel camera image is acquired, a 640×480 central pixel region (RGB) is cropped from the camera’s sensor. Since the underlying camera raw data has a lower color resolution, inherent to all Bayer-pattern sensors, we then downsample the cropped region to 320×240 pixels. This cropping and downsampling substantially speeds up subsequent processing by OpenDR, while preserving most of the original color information.

3.4. OpenDR pipeline for 3D pose localization

We use standard computer graphics to create realistic images for a particular camera pose and 3D scene, where the scene includes the surface geometry, surface reflectance, and illumination. Because these images can accurately describe the physical quantities that would be imaged from a real scene, computer graphics can also be used to work backward from an observed image to the

physical attributes of the scene from which it was acquired [15]. The latter is called “inverse rendering,” which generally involves the recovery of scene geometry, camera motion, and illumination. ProbeSight currently provides camera motion, but could potentially also infer changes in geometry (i.e., skin deformation). Matching computing illumination could also presumably be useful, as a means of optimizing recovery of motion and geometry.

With the coordinate transformations described in section 3.1 and a pre-calibrated camera, ProbeSight is able to re-project the 3D surface map to simulate what would be seen at the initial camera’s pose. For the present work, we are temporarily using the optical tracker to quickly initialize ProbeSight’s pose for the first frame only. When the US probe moves in space, its on-board camera image (observed image) is continually updated. We built an OpenDR pipeline to explicitly relate these updates in the observed image with changes in the parameters in the 3D surface map (3D model) OpenDR uses to render, allowing us to estimate motion of the camera and attached ultrasound probe.

We define $f(\theta)$ as the forward rendering process that depends on the parameters θ . The simplest optimization between changes in observed image and 3D model would solve for the parameters minimizing the difference between the rendered and observed image intensities as $E(\theta) = \|f(\theta) - I\|^2$. OpenDR separates θ into 3D vertex locations V of the surface (approximately 40,000 vertexes) in camera coordinates, per-vertex brightness values A that depend on the illumination model, and pre-calibrated camera parameters C , so that $\theta = \{V, A, C\}$. The differentiable rendering procedure is shown in Figure 4, where U indicates the 2D projected coordinates of 3D vertices. Using auto differentiation techniques, OpenDR

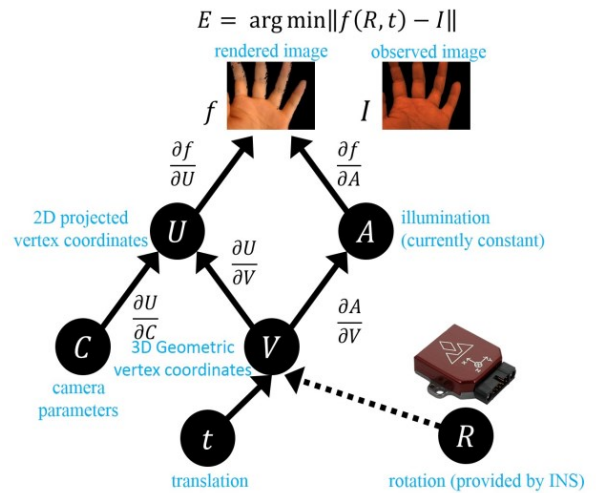


Figure 4: Partial derivative structure of the OpenDR renderer (This figure is adapted from [13]).

computes derivatives of the rendered image with respect to θ [16].

We reparametrize V with translation t , which must be optimized by OpenDR, and rotation R , which is provided by the INS. This makes optimal use of the rotation accuracy of the INS, and avoids ambiguity between image translation caused by camera translation vs. rotation. We express the error to be minimized as a difference between Gaussian pyramids, and we find a local minimum in the energy function with simultaneous optimization of these parameters. We perform nonlinear optimization using Powell’s dogleg method [17] which is implemented in Chumpy, a python-based framework designed to handle the auto-differentiation problem, [18] to estimate the translation between subsequent frames by iteratively matching the prior model with the observed image.

4. Experiment and Results

4.1. Tracking on Flat Phantom

ProbeSight was initially tested on a flat phantom containing a 2D printed image of actual palm skin. The image was acquired using a digital single lens reflex (dSLR) camera (Canon EOS 5D Mark I, Canon Inc., USA) with a macro-telephoto lens. A central rectangle was cropped from the palm area, and printed at 7 by 4 inches. (The cropped rectangle also served as the 3D anatomical model, as previously described in section 3.1) The phantom was used to validate both translation-only camera motion and freehand camera motion.

To achieve translation-only motion, ProbeSight was mounted on a 3-axis linear stage, allowing it to translate but not rotate within a one-cubic-inch 3D working space. During the experiment, a subject was instructed to randomly translate ProbeSight around within the working space. The Micron Tracker provided a ground-truth record of each movement with sub-millimeter accuracy. ProbeSight’s 3-DOF motion trace computed by OpenDR

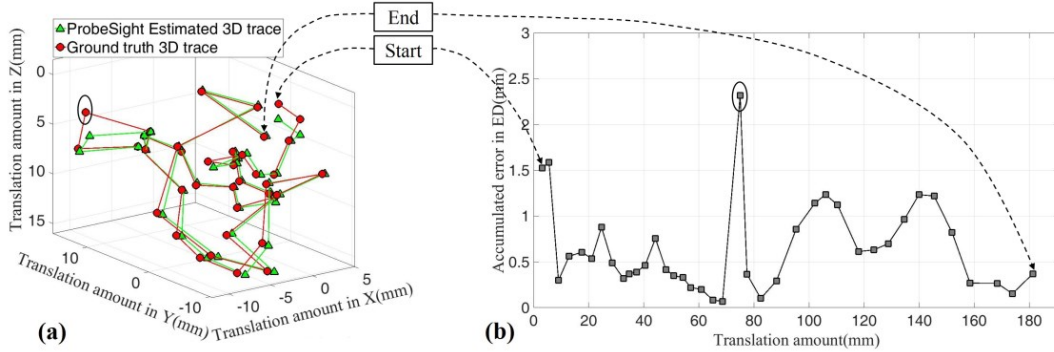


Figure 5a: Estimated probe trajectory (green triangles) compared with ground truth data (red circles) for the pure translation experiment data. **Figure 5b:** Accumulated error in Euclidian Distance. The outlier is circled in both graphs.

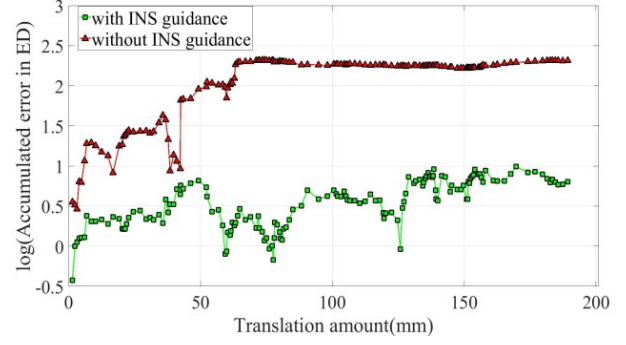


Figure 6: Performance comparison of ProbeSite on phantom freehand experiment data with or without INS guidance.

was compared with ground truth. An example trace is shown in Figure 5a and Figure 5b shows the accumulated Euclidian Distance (ED) error.

For the 6-DoF freehand experiment with the flat phantom, we analyzed two possible ProbeSight configurations, either using the INS or OpenDR to track 3D rotation (OpenDR was used in both cases to track translation). The accumulated error was 141.44 ± 69.91 mm over the first 100 mm of travel which obviously indicates that OpenDR failed to accurately track rotation, being unable to disambiguate between rotation and translation (a well-known problem in computer vision).

When OpenDR determined translation and the INS determined rotation, however, the accumulated error was only 4.06 ± 2.38 mm over the first 100 mm of travel. Figure 6 shows both results. To accommodate the extremely large error accumulated without INS guidance, a logarithmic scale is used.

4.2. Tracking on Human Upper Extremity

We next proceed to scanning an actual human subject’s upper extremity (hand, wrist, and a portion of the forearm). This target, with a 3D shape with varying physical texture, is more complex than the flat image of skin on the

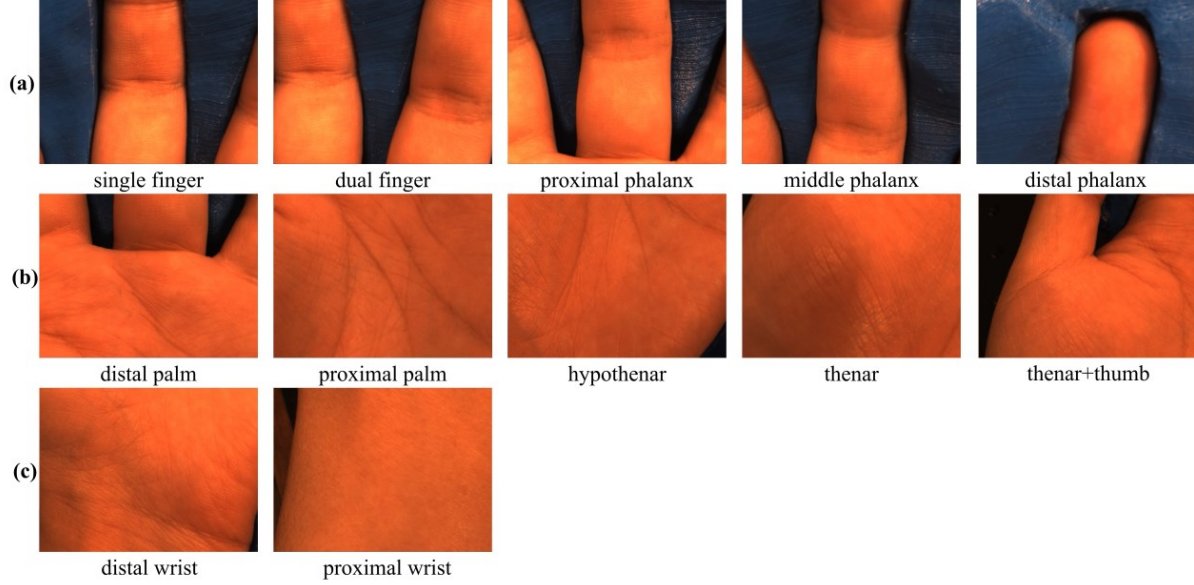


Figure 7: Example frames of three datasets: (a) Finger Dataset, (b) Palm Dataset and (c) Wrist Dataset.

phantom. Its textures are more visible when they cast tiny shadows into their “valleys,” depending on the direction of illumination. We used bright diffuse lighting to minimize this variable. We evaluated ProbeSight on three datasets containing various skin textures and geometry from different anatomical regions of the upper extremity: finger, palm and wrist. Some representative frames from these datasets are shown in Figure 7.

4.2.1 Datasets

Our *Finger* dataset contains 392 images continuously acquired from ProbeSight’s camera while moving across the fingers. The dataset consists of two parts, (1) for side-to-side motion across the proximal region of the fingers, and (2) for motion along the length of the fingers. In the first case, the *finger medial-lateral* (FML) dataset, the probe shifts from viewing both the middle and index fingers (*dual finger*), to viewing just the index finger (*single finger*), and then back to viewing both fingers (*dual finger*). In the second case, the *finger proximal-distal* (FPD) dataset, the probe moves from the middle finger’s proximal phalanx to the middle phalanx to the distal phalanx and then back again (with a view that also partially includes the index finger). Both finger edges and coarse creases are visible, including the *distal digital crease*, *middle digital crease*, and *proximal digital crease*. CLAHE preprocessing accentuates these creases in the images, thereby facilitating their use during OpenDR optimization.

The *Palm Dataset* contains 610 images continuously acquired from ProbeSight’s camera while moving across five regions of the palm. Again there are two parts to this

dataset, *palm medial-lateral* (PML) and *palm proximal-distal* (PPD). PML motion is from the hypothenar to thenar to thenar+thumb regions, and PPD motion is from the distal to proximal palm region. Visible coarse creases include the *distal transverse crease*, the *proximal transverse crease* and the *radial longitudinal crease* on the palm area. Only the thenar region lacks obvious coarse creases and edge information.

The *Wrist Dataset* contains 161 images continuously acquired from ProbeSight’s camera while moving along the wrist. This *wrist proximal-distal* (WPD) motion is from the distal wrist region to the proximal wrist region. Almost no coarse creases or obvious edges are visible in these regions. All upper extremity regions and creases mentioned above are defined and labeled in Figure 8.

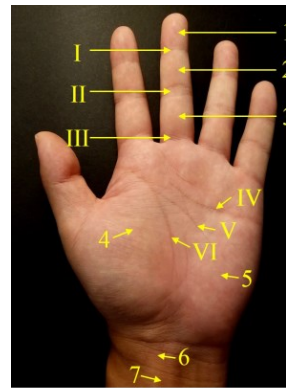


Figure 8: Definition of upper extremity regions and creases. The regions are defined as (1) distal phalanx, (2) middle phalanx, (3) proximal phalanx, (4) thenar, (5) hypothenar, (6) distal wrist and (7) proximal wrist. Major creases are defined as (I) distal digital crease, (II) middle digital crease, (III) proximal digital crease, (IV) distal transverse creases, (V) proximal transverse crease and (VI) radial longitudinal crease.

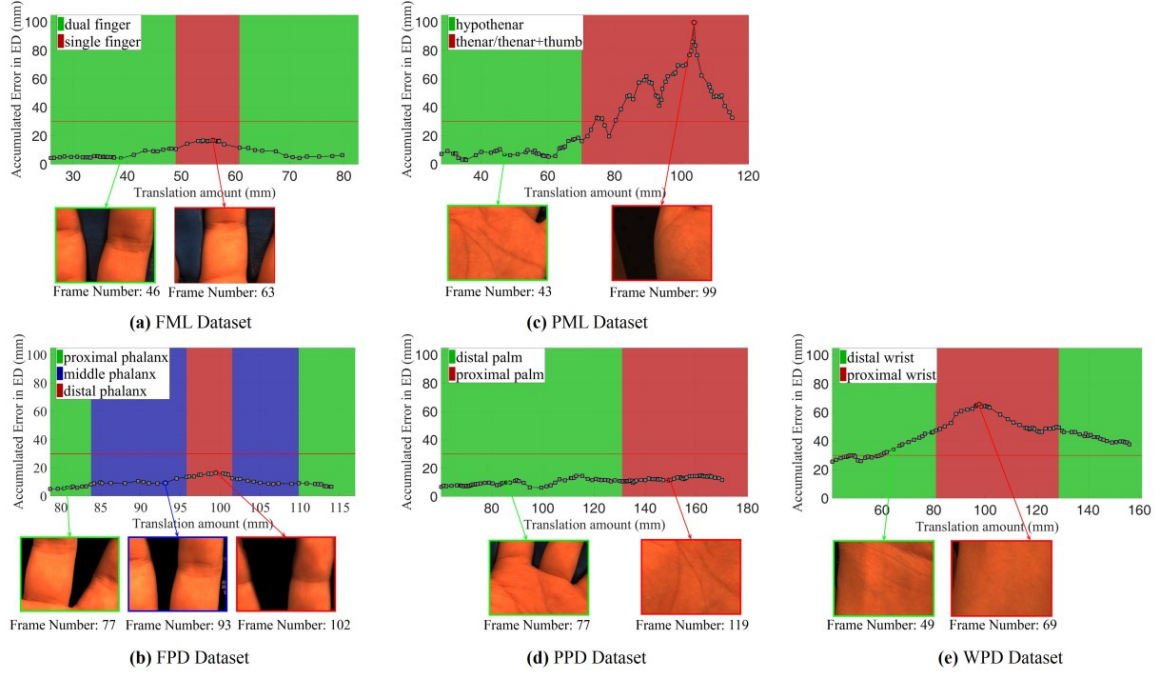


Figure 9: ProbeSight performance when tracking in different regions.

4.2.2 Results and Discussions

The efficacy of ProbeSight varies with the particular anatomical region and direction of camera motion.

Results from the FML dataset are shown in Figure 9a. The accumulated error climbs when the probe moves from dual fingers to a single finger, but then drops when the probe moves back to the dual finger region. Both of these regions have similar finger crease patterns (see representative frames 46 and 63 in the figure), and so we conclude that ProbeSight performs better when more features (e.g., coarse creases, edges and skinfold) are visible.

FPD results are shown in Figure 9b. The accumulated error increases when the probe moves from the proximal phalanx to the distal phalanx, but decreases when the probe goes back towards the proximal phalanx. The proximal and middle phalanx contain strong and clear creases, more than the distal phalanx (see representative frames 77, 93 and 102 in the figure). Again, we conclude that ProbeSight performs better when tracking strong and

clear features.

PML results are shown in Figure 9c. The accumulated error in the hypothenar region is much lower than the in thenar+thumb region. Looking at frames 43 and 99, we observe much less coarse creasing in the thenar region compared with hypothenar region. ProbeSight shows poor tracking ability without these coarse creases.

PPD results are shown in Figure 9d. Since both distal palm and proximal palm regions contain similar strong crease patterns, as shown in example frames 77 and 119, ProbeSight has low accumulated error when tracking in both regions.

WPD results are shown in Figure 9e. The accumulated error markedly increases when the probe moves from distal to proximal across the wrist region. Looking at frames 49 and 69, we observe that neither region has sufficient coarse creases, while the proximal wrist region contains almost no coarse creases. This explains why the accumulated error in both regions are very high, while the proximal wrist region has the highest error.

Table 1. Average accumulated error in different regions

Datasets	Finger					Palm				Wrist	
	FML		FPD			PML		PPD		WPD	
	Single Finger	Dual Finger	Proximal Phalanx	Middle Phalanx	Distal Phalanx	Hypo Thenar	Thenar	Distal Palm	Proximal Palm	Distal Wrist	Proximal Wrist
Err (mm)	15.63	6.63	6.96	9.95	14.85	8.93	51.66*	9.76	12.57	54.82*	37.25*

Err: The average accumulated error in Euclidean Distance. * The average accumulated error is above 30 mm.

Compared to the system of Sun, et al., which also tracks skin features directly [11], ProbeSight's tracking error does not accumulate, because ProbeSight is able to self-correct.

Table 1 uses average accumulated error to further illustrate how ProbeSight's performance varies across different regions. In general, the accumulated error increases in the regions where there are fewer clear features such as coarse creases and edges (e.g., thenar, distal wrist and proximal wrist regions). In extreme cases, ProbeSight's error may drift outside OpenDR's convergence range, requiring its estimated pose to be reset before ProbeSight can regain tracking, even when it moves back to a relatively easy tracking region. Conversely, ProbeSight's error generally decreases when viewing regions with clear coarse creases and strong edges (e.g., dual finger, proximal phalanx and distal palm regions).

We consider regions in which ProbeSight's average accumulated tracking error is above 30 mm to be *challenging regions*, while other regions to be *unchallenging regions*. This 30mm error threshold is graphed as a horizontal red line in Figures 9.

In an attempt to improve ProbeSight's performance in challenging regions, we explored giving ProbeSight a wider field of view, so as to always have more creases and features in view. Instead of directly center-cropping the 640×480 region from the camera's Bayer-pattern sensor, we first center-cropped a 1280×960 double-size region from the original camera image and then down-sampling it to 640×480 pixels before preprocessing (as previously described in section 3.3.3). It is obvious that, in both the challenging thenar and wrist regions, image processed by the direct center-cropping method has insufficient coarse creases and also lacks prominent edges as Figure 10a and 10b show. By comparison, Figure 10c and 10d show the

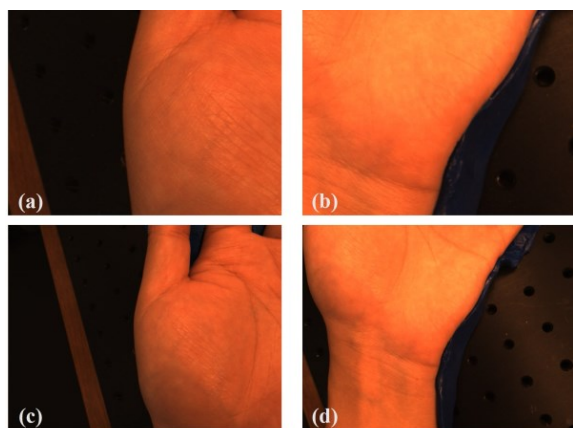


Figure 10a and 10b: The thenar and wrist region images acquired by direct center-cropping method. **Figure 10c and 10d:** The thenar and wrist region images acquired by double size center-cropping and downsampling method.

same original image processed by the method of double size center-cropping followed by downsampling incorporates coarse creases and edges from nearby regions, which may potentially improve the tracking performance.

Future work will involve data acquisition using the two different cropping methods, with a following up comparative study to show whether the proposed solution is able to significantly improve ProbeSight's tracking performance in the challenging regions. Toward a clinical application of ProbeSight, we will develop automatic initialization technology to no longer need the optical tracker, making ProbeSight a fully self-contained device.

5. Conclusions

In this paper, we have presented an effective 6-DOF probe tracking framework based on an OpenDR pipeline for inverse rendering and pose optimization. Results are presented for tracking natural skin features of upper extremities (arms and hands). A prior high-resolution 3D surface map is applied as an atlas of the anatomy. We demonstrate that with the assistance of INS, OpenDR is able to use a variety of skin features (e.g., coarse creases, edges and skinfold) to iteratively match the prior 3D surface map with a real-time 2D camera image, resulting in an accurate pose estimation for the camera. We have systematically evaluated our approach on three datasets from the finger, palm, and wrist, and the corresponding results show that the proposed approach successfully tracks the probe in a variety of regions in these datasets, without requiring any external tracking device. We have addressed the problem of distinguishing rotation and translation and demonstrated the advantages of including an INS to disambiguate these motions. ProbeSight's corrections lead to accurate tracking in unchallenging regions, even after temporarily accumulating small amounts of error in challenging regions. To our knowledge, ProbeSight is the first system to successfully accomplish US probe pose estimation by tracking skin and subdermal cues directly without requiring any artificial skin markers, and the first to include a pre-scan 3D surface map and INS. We believe the proposed solution may be valuable in many other types of surgical tool tracking applications, potentially leading to important benefits in clinical medicine.

6. Acknowledgement

This work was funded by NIH R01 grant 1R01EY021641, National Library of Medicine contract HHSN27620100058OP, and DoD Peer Reviewed Medical Research Program (PR130773, HRPO Log No. A-18237) award from the U.S. Department of Defense. Dr. Matthew Loper is thanked for his assistance.

References

- [1] N. J. Hangiandreou, "AAPM/RSNA Physics Tutorial for Residents: Topics in US," *RadioGraphics*, vol. 23, no. 4, pp. 1019–1033, Jul. 2003.
- [2] P. H. Mills and H. Fuchs, "3D ultrasound display using optical tracking," in *[1990] Proceedings of the First Conference on Visualization in Biomedical Computing*, 1990, no. July, pp. 490–497.
- [3] G. Flaccavento, P. Lawrence, and R. Rohling, "Patient and Probe Tracking During Freehand Ultrasound," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer Berlin Heidelberg, 2004, pp. 585–593.
- [4] T. C. Poon and R. N. Rohling, "Tracking a 3-D ultrasound probe with constantly visible fiducials," *Ultrasound Med. Biol.*, vol. 33, no. 1, pp. 152–157, Jan. 2007.
- [5] M. Nakamoto, K. Nakada, Y. Sato, K. Konishi, M. Hashizume, and S. Tamura, "Intraoperative Magnetic Tracker Calibration Using a Magneto-Optic Hybrid Tracker for 3-D Ultrasound-Based Navigation in Laparoscopic Surgery," *IEEE Trans. Med. Imaging*, vol. 27, no. 2, pp. 255–270, Feb. 2008.
- [6] M. Hastenteufel, M. Vetter, H.-P. Meinzer, and I. Wolf, "Effect of 3D ultrasound probes on the accuracy of electromagnetic tracking systems," *Ultrasound Med. Biol.*, vol. 32, no. 9, pp. 1359–1368, Sep. 2006.
- [7] Q. H. Huang, Y. P. Zheng, M. H. Lu, and Z. R. Chi, "Development of a portable 3D ultrasound imaging system for musculoskeletal tissues," *Ultrasonics*, vol. 43, no. 3, pp. 153–163, Jan. 2005.
- [8] C. Chan, F. Lam, and R. Rohling, "A needle tracking device for ultrasound guided percutaneous procedures," *Ultrasound Med. Biol.*, vol. 31, no. 11, pp. 1469–1483, Nov. 2005.
- [9] H. Raffi-Tari, P. Abolmaesumi, and R. Rohling, "Panorama Ultrasound for Guiding Epidural Anesthesia: A Feasibility Study," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6689 LNCS, 2011, pp. 179–189.
- [10] S.-Y. Sun, M. Gilbertson, and B. W. Anthony, "6-DOF probe tracking via skin mapping for freehand 3D ultrasound," in *2013 IEEE 10th International Symposium on Biomedical Imaging*, 2013, pp. 780–783.
- [11] S.-Y. Sun, M. Gilbertson, and B. W. Anthony, "Probe Localization for Freehand 3D Ultrasound by Tracking Skin Features," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8674 LNCS, no. PART 2, 2014, pp. 365–372.
- [12] J. Wang, V. Shivaprabhu, J. Galeotti, S. Horvath, V. Gorantla, and G. Stetten, "Towards Video Guidance for Ultrasound, Using a Prior High-Resolution 3D Surface Map of the External Anatomy," in *Lecture Notes in Computer Science*, 2014, pp. 51–59.
- [13] M. M. Loper and M. J. Black, "OpenDR: An approximate differentiable renderer," in *European Conference on Computer Vision*, 2014, vol. 8695 LNCS, no. PART 7, pp. 154–169.
- [14] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Comput. Vision, Graph. Image Process.*, vol. 39, no. 3, pp. 355–368, Sep. 1987.
- [15] S. R. Marschner, "Inverse rendering for computer graphics," 1998.
- [16] W.-C. Chiu and M. Fritz, "See the Difference: Direct Pre-Image Reconstruction and Pose Estimation by Differentiating HOG," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 468–476.
- [17] M. L. A. Lourakis and A. A. Argyros, "Is Levenberg-Marquardt the most efficient optimization algorithm for implementing bundle adjustment?," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, 2005, vol. II, p. 1526–1531 Vol. 2.
- [18] M. M. Loper, "<https://github.com/mattloper/chumpy/blob/master/optimization.py>," [Online]. Available: <https://github.com/mattloper/chumpy/blob/master/optimization.py>.