Proceedings of the International Symposium on Biomedical Imaging 2007

DESIGN AND CALIBRATION OF A VIRTUAL TOMOGRAPHIC REFLECTION SYSTEM

Damion Shelton[1], *Bing Wu*[2], *Roberta Klatzky*[2], *and George Stetten*[1,3]

Carnegie Mellon University Robotics Institute[1] and Department of Psychology[2]
University of Pittsburgh Department of Bioengineering[3]
Pittsburgh, PA

ABSTRACT

Virtual Tomographic Reflection (VTR) is a new augmented reality technique that allows users to view volumetric image data using an interaction paradigm based on medical ultrasound. VTR is essentially a "virtual" version of the Sonic Flashlight, a device that permits real-time in situ visualization of ultrasound images by reflecting calibrated images displayed on a flat-panel monitor from a partially transparent half silvered mirror [1]. In VTR, as opposed to the actual Sonic Flashlight, the ultrasound scanner is replaced by an optically tracked dummy ultrasound scanner with a mirror and display capable of generating cross-sections through stored real or computer-generated image data. The design of this system was motivated by the need to conduct psychophysical and accuracy analysis of the real Sonic Flashlight in a highly controlled artificial environment [2]. Here we present the latest version of our VTR device and describe the overall system architecture as well as a highly accurate method for calibrating the system.

Index Terms— Virtual reality, Simulation, Biomedical imaging, Psychology, User interface human factors

1. INTRODUCTION

Real-time tomographic reflection (RTTR) is an augmented reality (AR) technique that relies on a specific configuration of a partially transparent half-silvered mirror and flat panel display in order to create the illusion that a tomographic (planar) slice of data is floating in space. When used in conjunction with an ultrasound scanner as a data acquisition source, the device is called the Sonic Flashlight [1]. RTTR has also been shown to work with Computed Tomography (CT) systems [3, 4]. The major advantage of RTTR over most other AR techniques is that it is viewpoint independent and does not require head-tracking in order to present the virtual image. As a result, it is very easy to make the Sonic Flashlight, or a number of other RTTR devices, entirely self contained and light weight.

Despite the initial clinical success of the Sonic Flashlight for the placement of peripherally inserted central catheter (PICC) lines in the deep veins of the arm [5], there are many unanswered questions about how and why the illusion of a planar slice floating in space is useful for localizing 3D structures. We are particularly interested in the perceptual issues surrounding the task of PICC line placement, since our first clinical trials focus on PICCs. To briefly summarize, the major difference between the Sonic Flashlight and conventional ultrasound (CUS) guidance is that of hand-eye coordination, which is displaced with CUS but not with Sonic Flashlight [6]. Using CUS, the user must look away from the patient workspace in order to view the ultrasound guidance information, while with the Sonic Flashlight the guidance information is presented *in situ*.

In the first test of psychophysical issues surrounding RTTR, Wu et al. measured user performance during perceptual localization (as assessed by needle pointing) and needle insertion tasks with both CUS and the Sonic Flashlight and found fundamental perceptual and performance differences between the two [6]. The experimental design used a magnetic tracking device attached to a pen tool and a number of water tanks containing ultrasound targets (commonly called phantoms).

The use of physical phantoms presents problems when attempting to assess the psychophysics of a variety of target parameters whose values (contrast, target shape, target size, etc.) would be very difficult to manage. In fact, prior to Wu's study we had already proposed a system for exploring psychophysical issues surrounding RTTR using a largely virtual test environment [2]. A virtual test environment allows rapid testing of a wide variety of virtual phantoms, and more importantly, provides precise control over phantom characteristics such as contrast, amount of noise, etc.

2. SYSTEM DESIGN

The goal of VTR is to generate a slice through a stored 3D image and present the slice to the user on a mock Sonic Flashlight display. An Optotrak 3020 optical tracking device (Northern Digital, Inc.) is used to track the position and orientation of the mock Sonic Flashlight, a pen or needle tool, and a physical container that provides a bounding box for the virtual dataset. The physical bounding box can be made of a variety of materials, but we anticipate that the most useful will be a "blank" gel phantom that provides haptic feedback similar to that experienced during an actual needle insertion. Note that we do not attempt to simulate the acoustic behavior of ultrasound; the purpose of this system is to model reslicing of a 3D volume without regard to imaging modality.

The mock Sonic Flashlight (see Figure 1) is physically quite similar to the clinical version, with the exception of the tracking cluster add-on, and the lack of a functional ultrasound machine. The organic LED (OLED) display, mirror, transducer body[1], and geometry of the mirror-OLED alignment are all identical between the clinical and virtual versions of the device. The tracking cluster adds a slight amount of weight as compared to the clinical device, although not enough to be noticeable. In practice, one of the main concerns is minimizing the bulk of the cables exiting from the rear of the Sonic Flashlight, and although the tracking cluster does add an additional cable we were able to remove the transducer cable (not needed for the virtual device) and maintain a similar overall cable weight between the real and virtual devices. A thorough analysis of tracking cluster design can be found in the work of Hamza-Lup et al. [7, 8]

[1]We use a surplus nonfunctional transducer of the same model as in the clinical Sonic Flashlight.

Fig. 1. Mock Sonic Flashlight and pen tool in use.

Additional components of the system include the pen/needle tool and the tracked bounding box. The former was purchased from Northern Digital, the manufacturers of the Optotrak system, and the latter is simple to construct using a planar array of tracking markers rigidly fixed to whatever physical medium is desired for optimum haptic characteristics. All of the tracking markers used with the Optotrak are active infrared LEDs, grouped into clusters of up to 24 markers to define rigid bodies which report position and orientation in the global coordinate frame. Rigid body definitions are automatically generated using software provided by Northern Digital.

In order to provide a realistic illusion of viewing cross-sections through a physical object (the bounding box), the entire system must be calibrated to work in a common frame of reference, the Optotrak coordinate system, also referred to as global coordinates. This allows the simulated data set to be placed at a known location in space and also allows computation of the appropriate cross section depending on the position and orientation of the mock Sonic Flashlight. Assuming the relatively simple case of a bounding box in the shape of a rectangular prism, alignment of the image data within the bounding box is achieved by sampling clusters of points on each of three orthogonal planes on the surface of the bounding box using the pen tool.

Reslicing of the medical image data set is performed with a combination of custom software written in C++ and a high-performance 3D accelerator card[2] using the OpenGL 2.0 API. The software front-end is responsible for loading the medical image data from disk, interfacing with the Optotrak tracking unit, and allowing management of the psychophysical trial being conducted. We use the Insight Toolkit, an open source medical image analysis toolkit, to handle a variety of input image formats as well as to generate custom datasets using built-in modeling capabilities [9]. Rather than attempting to compute the correct cutting plane in software, the medical image dataset is converted to a 3D OpenGL texture and stored in the video card memory, allowing for very fast update rates. The video card we use provides multiple video outputs, so we use a multiple display setup where the test control interface is presented on a flat panel monitor and the simulated slice image is sent to the OLED interface board for display on the mock Sonic Flashlight.

A "virtual CUS" mode is also implemented for future psychophysical studies that compare the Sonic Flashlight to conventional ultrasound. To simulate the operation of a conventional ultrasound device, the OLED display is disabled and the slice image is instead sent to an external flat panel display that mimics a conventional ultrasound display. When operating in virtual ultrasound mode, the subject still uses the virtual Sonic Flashlight physical apparatus, eliminating any performance differences between the two display modalities which would result from variations in weight or shape.

3. CALIBRATION OF THE VTR DEVICE

In RTTR a combination of geometric constraints and a planar calibration technique ensure that the reflected virtual image and ultrasound slice are coincident. Proper calibration of an RTTR system guarantees that each pixel of rendered data exactly corresponds with the point in physical space from which the ultrasound data was acquired. Owing to the absence of an actual ultrasound machine, VTR can be thought of as the process of slicing a simulated dataset with the virtual image (i.e. reflection of the planar display) itself. Much as in RTTR, the reflected pixels in VTR can be thought of as occupying real physical space within the virtual dataset. In other words, the pixels are both the *representation* and *source* of the data.

The goal of the calibration is to take 2D screen coordinates used to display pixels on the flat-panel monitor (P^*) and translate them into 3D world coordinates shared by the computer-generated target and tools simulating invasive procedures (G). Alternatively, one may think of the calibration procedure as determining a transformation matrix which, when multiplied with that of a tracking device rigidly fixed to the mock Sonic Flashlight (V), results in the coordinate system of the virtual image (P) in global space. Figure 2 shows the relationships between the various coordinate systems (described below) and known data involved in the calibration process.

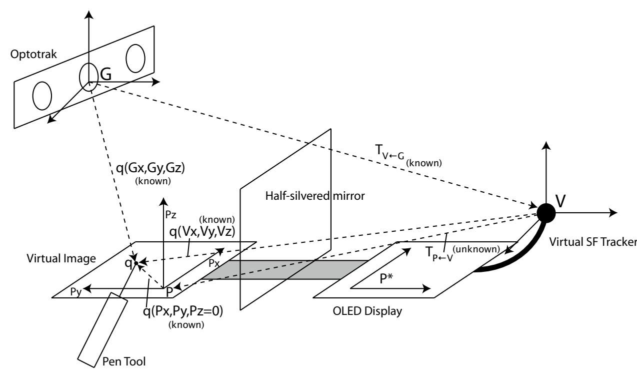

Fig. 2. Coordinate systems used in VTR calibration.

3.1. Known coordinate system transforms & data

1. The transform of the mock Sonic Flashlight tracker in global coordinates, $T_{V \leftarrow G}$.
2. The position of a point q in global coordinates, $q(G_x, G_y, G_z)$. q is the center of the spherical tip of a stylus that has previously been calibrated for use with the Optotrak.
3. Since $T_{V \leftarrow G}$ is known, it is straightforward to compute the position of point q in V coordinates, $q(V_x, V_y, V_z)$.

[2]Wildcat Realizm 200, 3Dlabs

4. The position of q in P, $q(P_x, P_y, P_z = 0)$, as described in the following section.

3.2. Calibration procedure

The goal of calibration is to determine $T_{P \leftarrow V}$, the position and orientation of the virtual image coordinate system P relative to the sonic flashlight coordinate system V. During operation of the virtual ultrasound system, this transform will be applied to $T_{V \leftarrow G}$ in order to yield the transformation of the virtual image in global coordinates, $T_{P \leftarrow G}$ (not shown in Figure 2 for purposes of clarity).

Calibration occurs by establishing correspondences between two sets of points $V(q_{i=1\ldots n})$ and $P(q_{i=1\ldots n})$ and recovering the optimal similarity transform which aligns corresponding points. Corresponding point pairs are determined by displaying a marker in the virtual image and visually aligning the tip of the pen tool with the virtual marker (see Figure 3). The initial selection of coordinate frame P is essentially arbitrary, although the most logical choice is to work in the pixel-based coordinate system native to the display. Starting with P in pixel coordinates, we recover an isotropic scale factor c which is the number of pixels per world coordinate unit (mm), the inverse of which is the pitch of the pixels on the display. Since pixel pitch is a known parameter of the OLED display, this is a useful check that the calibration is successful.

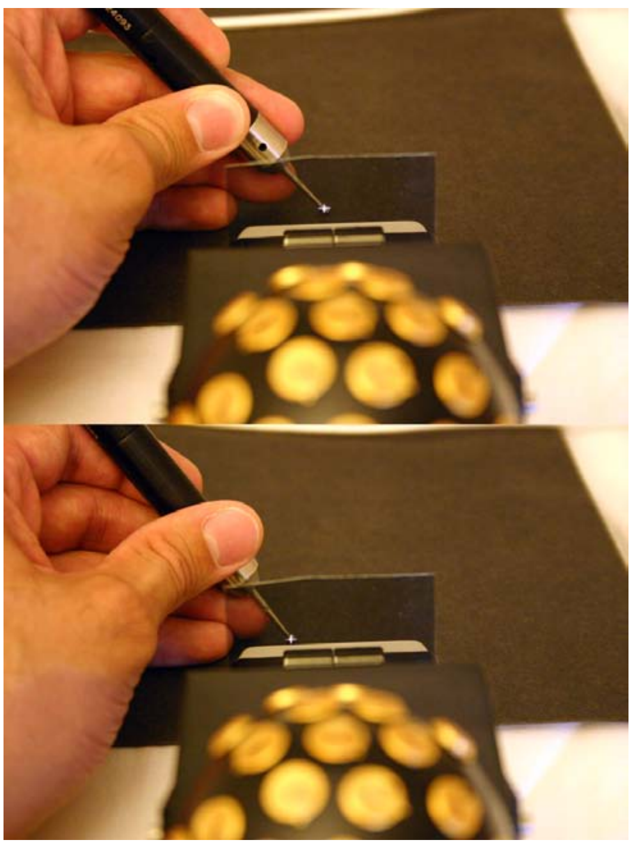

Fig. 3. Aligning the pen tool in V coordinates with two calibration targets in P coordinates.

Although the calibration requires only 4 point pairs, better results are obtained by sampling a wide variety of points over the range of possible positions in the virtual image. Several techniques have been proposed for solving point correspondence problems of this type [10, 11, 12]. We chose to implement a refined version of this algorithm developed by Umeyama [13].

Let $X = \{x_1, x_2, \cdots, x_n\}$ and $Y = \{y_1, y_2, \cdots, y_n\}$ be $n \times m$ matrices, where m is the dimensionality of the coordinate systems involved (in our case 3), and n is the number of point pairs acquired, with $x_i = V(q_i)$ and $y_i = P(q_i)$.

$$\mu_x = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1}$$

$$\mu_y = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{2}$$

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{n} |x_i - \mu_x|^2 \tag{3}$$

$$\sigma_y^2 = \frac{1}{n}\sum_{i=1}^{n} |y_i - \mu_y|^2 \tag{4}$$

$$\Sigma_{xy} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \mu_y)(x_i - \mu_x)^T \tag{5}$$

Let UDV^T be a singular value decomposition of Σ_{xy}. The optimal similarity transform which aligns the two point clouds can be decomposed into a rotation matrix R, scale factor c, and translation vector t. If rank$(\Sigma_{xy}) \geq m - 1$ (true in our case), then:

$$R = USV^T \tag{6}$$

$$t = \mu_y - cR\mu_x \tag{7}$$

$$c = \frac{1}{\sigma_x^2} tr(DS) \tag{8}$$

where S is chosen as

$$S = \begin{matrix} I & \text{if } \det(U)\det(V) = 1 \\ diag(1, 1, \cdots, 1, -1) & \text{if } \det(U)\det(V) = -1 \end{matrix} \tag{9}$$

Although we could apply the entire similarity transform at once, it is more intuitive to perform an initial calibration trial to recover c and then apply the scale factor in software when creating the display coordinate frame P. Applying c separately from the rest of the transform allows us to draw graphical objects in more natural physical coordinates rather than in pixel coordinates. The desired transformation $T_{P \leftarrow V}$ results from composing a 4x4 homogeneous transformation matrix:

$$T_{V \leftarrow P} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \tag{10}$$

3.3. Results of Calibration

We performed two successive calibrations of the VTR device, the first to recover the scale c and the second to verify repeatability of the calibration after accounting for c when defining the P coordinate frame. Initial calibration, assuming P in pixel coordinates and using 16 point pairs, yielded a scale c of 18.3 pixels/mm. Inverting this to yield pixel pitch we obtain a pitch of 0.0546 mm/pixel. The actual display pixel pitch, obtained from the Kodak OLED display specs, is 0.0548 mm/pixel, for a pixel pitch error of approximately 0.39%.

If the entire display is considered and we multiply the isotropic pixel pitch error by the length of each display dimension in pixels, we find a worst case error magnitude of 0.11 mm for the x display length and 0.18 mm for the y display length.

For purposes of comparison, the Optotrak static positioning accuracy as reported by Northern Digital is 0.1 mm. Although we might expect to exceed this accuracy by virtue of a large number of point pairs (and the resulting increase in signal-to-noise) during calibration, the overall interaction of the various tracked coordinate systems is non-trivial. A more complete analysis of error in our system will be the subject of future work.

4. CONCLUSION AND FUTURE WORK

We have developed a new technique for reslicing volumetric data and displaying the resulting tomographic image in situ within a physical bounding box. Preliminary calibration results demonstrate that the position and orientation of the virtual image within space can be determined by point correspondences between display and world coordinates. In the future, we would like to quantify error within the virtual workspace of the tracking system by means of a physical calibration jig.

We believe that VTR will be an important tool both for quantitative testing of the Sonic Flashlight and more generally for exploring large medical images in an intuitive fashion. As an example of the former, we would like to extend our first psychophysical study [6] of the Sonic Flashlight using a larger variety of target objects. This is very easy to do in our virtual test environment, since we are not restricted to working with physical phantoms. As an example of the latter, we present an image showing a greatly reduced MRI of the brain "contained" inside a plastic bowl, with the needle tool being used to mark a point a point on the displayed slice. Note that the needle disappears below the surface of the plastic grid; the intersection point displayed on the slice is simulated via our software.

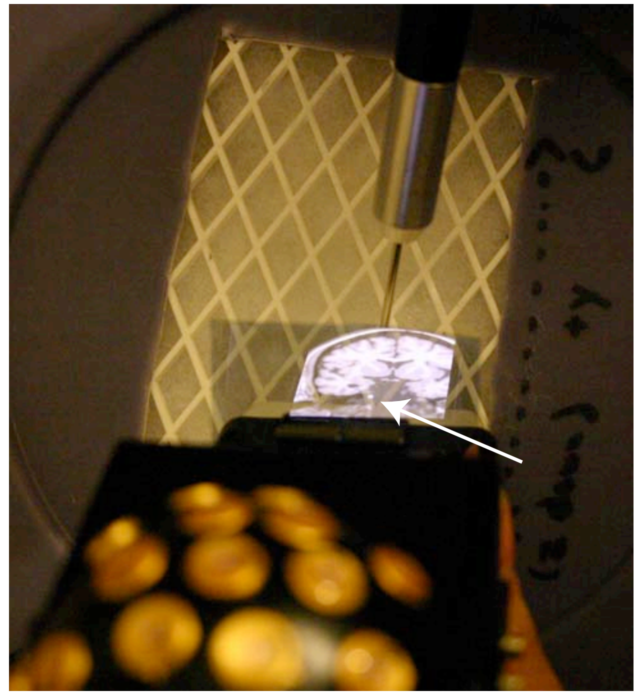

Fig. 4. VTR device being used to image a greatly reduced brain MRI. Simulated needle intersection point is noted by by the arrow.

5. REFERENCES

[1] G. Stetten and V. Chib, "Overlaying ultrasound images on direct vision," *Journal of Ultrasound in Medicine*, vol. 20, no. 1, pp. 235–240, 2001.

[2] D Shelton, R Klatzky, and G Stetten, "Method for assessing augmented reality needle guidance using a virtual biopsy task," in *IEEE International Symposium on Biomedical Imaging*, Arlington, 2004, vol. 1, pp. 273–276, IEEE.

[3] K. Masamune, Y. Masutani, S. Nakajima, I. Sakuma, T. Dohi, H. Iseki, and K. Takakura, "Three-dimensional slice image overlay system with accurate depth perception for surgery," in *Medial Image Computing and Computer-Assisted Intervention (MICCAI)*, Pittsburgh, 2000, vol. 1935, pp. 395–402, Springer.

[4] Ken Masamune, Gabor Fichtinger, Anton Deguet, Daisuke Matsuka, and Russell Taylor, "An image overlay system with enhanced reality for percutaneous therapy performed inside ct scanner," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2002: 5th International Conference*, T. Dohl and R. Kikinis, Eds. 2002, vol. 2489 of *Lecture Notes in Computer Science*, pp. 77–84, Springer-Verlag GmbH.

[5] W. Chang, N. Amesur, D. Wang, A. Zajko, and G. Stetten, "First human clinical trial of the sonic flashlight - guiding placement of peripherally inserted central catheters," in *2005 Meeting of the Radiological Society of North America*, Chicago, Illinois, 2005, number SSJ03-02.

[6] B Wu, R Klatzky, D Shelton, and G Stetten, "Psychophysical evaluation of in-situ ultrasound visualization," *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, vol. 11, no. 6, pp. 684–693, 2005.

[7] F. Hamza-Lup, L. Davis, C. Hughes, and J. Rolland, "Marker mapping techniques for augmented reality visualization," in *XVIIth Intl. Symposium on Computer and Information Sciences (ISCIS)*. 2002, pp. 152–156, CRC Press.

[8] L. Davis, F.G. Hamza-Lup, and J. Rolland, "A method for designing marker-based tracking probes," in *IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, 2004, pp. 120–129.

[9] L. Ibanez, W. Schroeder, L. Ng, J. Cates, and et al., *The ITK Software Guide: The Insight Segmentation and Registration Toolkit (version 1.4)*, Kitware, Inc., Albany, NY, 2003.

[10] K.S. Arun, T.S. Huang, and S.D. Blostein, "Least-squares fitting of two 3-d point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 5, pp. 698–700, 1987.

[11] Berthold K.P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of North America*, vol. 4, pp. 629–642, 1987.

[12] Berthold K.P. Horn, Hugh M. Hilden, and Shahriar Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *Journal of the Optical Society of North America*, vol. 5, pp. 1127–1135, 1988.

[13] Shinji Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 376–380, 1991.