

# Final Progress Report

Lukas Gemar

March 4, 2016

## 0.1 Tool Model

The position of the tip of the interaction tool is  $\mathbf{x}_T^w$ . This position is specified in world coordinates; the superscript  $w$  indicates that the coordinates of this position are specified relative to the origin of world space,  $\mathbf{O}_w$ . This origin of the world coordinate system is approximately the location of the users head. It is not exactly the location of the user's head because the user can look up and down and side to side while wearing the display. The origin of the world coordinate system should remain fixed to these rotations and translations. The rotation of the display is measured, but the translation of the display is not. However, the translation of the head should be a function of the rotation of the display. A simple model of how the head moves from side to side should relate the head's rotation to its movement relative to the neck and shoulders.

Since the camera is mounted to the display, the position of the world coordinate system is not exactly the position of the camera. The origin of the camera coordinate system,  $\mathbf{O}_c$ , is offset from the world coordinate system by the rotation and translation of the user's head. Assuming that the rotation and translation of the user's head is in fact minimal while the user is seated, then the translation of the camera,  ${}^c\mathbf{t}^w$ , should be a function of the rotation of the head mounted display,  ${}^c\mathbf{R}^w$ :  ${}^c\mathbf{t}^w = f({}^c\mathbf{R}^w)$ .

Let  ${}^c\mathbf{R}^w$  and  ${}^c\mathbf{t}^w$  be the rotation and translation matrices, respectively, that specify the transformation from the world coordinate system to the camera coordinate system. Then the position of the interaction tool in camera coordinates,  $\mathbf{x}_T^c$ , is given by the following relation:

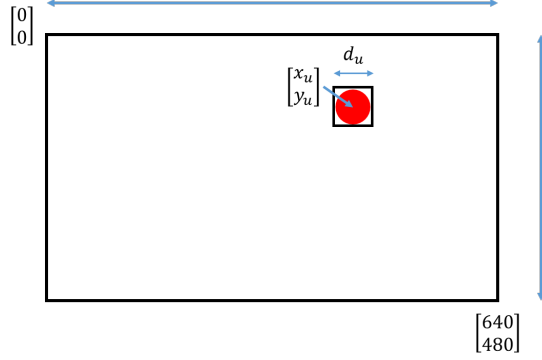
$$\mathbf{x}_T^c = {}^c\mathbf{R}^w \mathbf{x}_T^w + {}^c\mathbf{t}^w$$

The position of the tool relative to the camera,  $\mathbf{x}_T^c$ , is not observed directly. However, this position is inferred from the position of the back of the tool,  $\mathbf{x}_E^c$ , and the orientation of the tool. Call the back of the tool the eraser. The position of the eraser,  $\mathbf{x}_E^c$ , is found through computer vision. The orientation of the tool relative to the world coordinate system is specified by the orientation of the tool in body coordinates,  $\theta_T^b$ , and the rotation matrix from the body coordinates of the tool to the world coordinates,  ${}^w\mathbf{R}^b$ ; their product yields the absolute orientation of the tool in world coordinates. To correctly display the orientation of the tool to the user,  ${}^w\mathbf{R}^b$ , the rotation of the camera relative to the world, must also be known.  ${}^w\mathbf{R}^b$  is measured by the virtual reality display. The position of the tool in camera coordinates is found through the following relation:

$$\mathbf{x}_T^c = \mathbf{x}_E^c + {}^c\mathbf{R}^w ({}^w\mathbf{R}^b \theta_T^b)$$

## 0.2 Camera Model

The following is a model of a single frame from the camera. In this frame, the red circular object represents the back of the interaction tool. The back of the interaction tool is called the eraser. The width of the target along the  $x$  dimension of the image is  $d_u$ .



The midpoint of the eraser in the ideal undistorted image is

$\mathbf{x}_E^u = \begin{bmatrix} x_u & y_u \end{bmatrix}^T$ . The center of the undistorted image is  $c_x$  in the  $x$  direction and  $c_y$  in the  $y$  direction. The focal lengths in the  $x$  and  $y$  directions are  $f_x$  and  $f_y$ , respectively. These parameters are found through the camera calibration routine. The measurements of the center of the eraser and its width are related to the hidden state  $\mathbf{x}_C = \begin{bmatrix} x_C & y_C & z_C \end{bmatrix}^T$  in the following way:

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) = \begin{bmatrix} x_u \\ y_u \\ d_u \end{bmatrix} = \begin{bmatrix} \frac{f_x x_C}{z_C} + c_x \\ \frac{f_y y_C}{z_C} + c_y \\ \frac{d_C}{z_C \sqrt{\frac{1}{f_x^2} + \frac{1}{f_y^2}}} \end{bmatrix}$$

These relationships are found through using a simple pin-hole camera model [1].

## 0.3 Attitude Estimation

The attitude of the interaction tool relative to the world coordinate frame can be represented in a number of ways. The two representations that will be important here are the attitude matrix and the attitude quaternion. First, the attitude of the orientation tool can be described by the attitude matrix  $A$ . The attitude matrix typically maps a vector in the reference

frame to a vector in the the body frame [2]. This mapping is given by the relation,

$$\mathbf{b} = \mathbf{A}\mathbf{r}$$

Second, the attitude of the interaction tool can be represented as a quaternion. A rotation in three-dimensions is equivalent to an axis of rotation and an angle of rotation. If  $\hat{n}$  is the axis of a rotation and  $\theta$  is the angle of the rotation, then the quaternion  $\bar{q}$  represents the attitude:

$$\bar{q} = \begin{bmatrix} \vec{q} \\ q_4 \end{bmatrix}$$

with  $\vec{q} = \hat{n} \sin \frac{\theta}{2}$  and  $q_4 = \cos \frac{\theta}{2}$ . The quaternion must have length 1:  $\bar{q}^T \bar{q} = 1$ .

The attitude matrix and the quaternion representing attitude are related by the following equation, where  $A$  is a function of  $\bar{q}$  [3]:

$$A(\bar{q}) = (q_4^2 - \|\vec{q}\|^2)I_{3 \times 3} + 2\vec{q}\vec{q}^2 - 2q_4\vec{q}^X$$

where  $\vec{q}^X$  is the skew-symmetric cross product matrix,

$$\vec{q}^X = \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2q_1 & 0 & 0 \end{bmatrix}$$

The relationship between  $A$  and  $\bar{q}$  can also be expressed as the product of two matrices [2]:

$$A(\bar{q}) = \Omega(\bar{q})^T \Phi(\bar{q})$$

with

$$\Omega(\bar{q}) = \begin{bmatrix} q_4 \mathbf{I}_{3 \times 3} + \bar{\mathbf{q}}^X \\ -\bar{q}^T \end{bmatrix}$$

$$\Phi(\bar{q}) = \begin{bmatrix} q_4 \mathbf{I}_{3 \times 3} - \bar{\mathbf{q}}^X \\ -\bar{q}^T \end{bmatrix}$$

# Bibliography

- [1] Roger Y Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *Robotics and Automation, IEEE Journal of*, 3(4):323–344, 1987.
- [2] John L Crassidis, F Landis Markley, and Yang Cheng. Survey of nonlinear attitude estimation methods. *Journal of guidance, control, and dynamics*, 30(1):12–28, 2007.
- [3] Ern J Lefferts, F Landis Markley, and Malcolm D Shuster. Kalman filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 5(5):417–429, 1982.