Recovering Projection Geometry: How a cheap camera can outperform an expensive stereo system

Matthias M. Mitschke

Siemens AG Medical Engineering
Henkestraße 127
91054 Erlangen, Germany
Matthias.Mitschke@med.siemens.de

Nassir Navab

Imaging & Visualization Department
Siemens Corporate Research
Princeton, NJ 08540-6632, USA
Nassir.Navab@scr.siemens.com

Abstract

Recovering the projection geometry of an X-ray system or an augmented reality video see-through Head Mounted Display (HMD) are mathematically quite similar. Recent work in both medical imaging and augmented reality use external optical sensors in order to recover the motion of the imaging system. In this paper, we take the example of the recovery of an X-ray projection geometry. We show that the mathematical problem, which needs to be solved, is equivalent to the hand-eye calibration well studied in both computer vision and robotics community. We present a comparative study for the recovery of the motion and therefore projection geometry using five different hand-eye calibration methods proposed in the literature. We compare the motion estimation results using expensive external stereo-based tracking systems with one obtained by using an integrated optical camera. The paper concludes by showing that even if the motion estimation is more accurate when using an external sensor, the projection geometry is better estimated by the integrated optical camera. These results are of crucial importance to both medical imaging and augmented reality communities.

1 Introduction

Real-time recovery of the projection geometry is a fundamental issue in different applications. In particular, both medical imaging and augmented reality communities are interested in this. In augmented reality applications the goal is to merge virtual objects onto real images as the viewer is moving and therefore changing his/her viewpoint [15, 1, 9, 7]. Other than for augmented reality applications [9, 7], medical imaging community is also interested in real-time recovery of X-ray projection geometry for 3D tomographic reconstruction [13, 10]. Both of these scientific communities show a great interest in using precise and commercially available external optical tracking and localization systems such as Optotrak and Polaris from Northern Digital, Inc. These external sensors have been used in many medical and non-medical applications during the last years. Chassat and Lavallée discuss different medical applications of these systems and present a comparative study of their accuracy [2].

In this paper we first discuss the mathematical equations which need to be solved in order to use an external sensor for tracking and localization. We take the example of the recovery of the projection geometry for interventional 3D tomographic reconstruction. Here we describe a mathematical framework for solving this problem. We show that the mathematical framework is the same as the well-known hand-eye calibration framework in computer vision and robotics community. We use it to recover the motion of the C-arm and therefore the X-ray projection geometry during its rotational run around an object of interest. The results are evaluated both in terms of the qualitative comparisons of 3D reconstruction results and numerical comparison of motion estimation parameters. Five different hand-eye calibration methods are used.

We then compare these results with the one obtained by using an integrated standard CCD camera, see [13]. We finally discuss the interesting observation that even though the motion parameters are better estimated using the external sensor, the projection geometry is more accurately estimated by the integrated optical camera.

The rest of the paper is organized as follows: section 2 provides a general description of the problem of geometrical calibration for 3D tomographic reconstruction. This is taken as the application example in this paper. Note that for HMD calibration, the frame-

work is quite similar. Section 3 first briefly describes the projection geometry for camera and X-ray imaging, and defines the world, camera, external sensor and X-ray coordinate systems. Then it introduces different calibration procedures, and the algorithms we use. Section 4 presents the experimental results. Our conclusions and suggestions for future work are presented in section 5.

2 Geometrical calibration of an X-ray C-arm for 3D tomographic reconstruction

For tomographic reconstruction, a sequence of images is captured during the C-arm rotation around the object of interest. For each image the precise projection geometry has to be determined for the 3D reconstruction process [11, 12]. Some C-arms do not provide an accurate projection geometry or pose information. Therefore, for each image frame the projection geometry has to be calibrated.

2.1 Direct estimation of X-ray projection geometry

In order to get the most accurate estimation of the projection geometry for each X-ray frame, a calibration phantom with X-ray opaque markers has to be used. It is positioned such that it is visible in each frame. The projection matrix representing the projection geometry can then be estimated from the set of detected markers in the X-ray image and its corresponding 3D model points. We will use the projection matrices obtained with this method as *gold standard* – all other methods are compared to it – for our evaluation experiments.

Using this very precise calibration method in applications causes the problem that both the marker points and the object that is to be reconstructed have to be visible in the sequence of X-ray images. Marker points that overlay the object (patient's anatomy) influence the result of the 3D reconstruction. Vice versa, the pattern recognition process for precisely locating the markers in the X-ray image is influenced by the overlaid object of interest.

To deal with this problem the recovery of the projection geometry can be done in an offline procedure if the motion of the C-arm is reproducible between the calibration and the patient run [11, 12]. Some C-arm systems however do not have this desirable property. The projection geometry then has to be determined online during the patient run. This is described in more detail in the next section.

3 Online recovery of projection geometry

In this section we want to define the framework for online recovery of X-ray projection geometry from external sensors. First, we briefly introduce the X-ray projection geometry and the concept of virtual detector plane. Then the two different approaches compared in this paper are presented.

3.1 Definition of X-ray projection geometry

The X-ray projection geometry is represented by $\boldsymbol{P}$ a 3×4 homogeneous matrix of projection.

$$\boldsymbol{u} \cong \boldsymbol{P}\,\boldsymbol{x}, \tag{1}$$

where $\boldsymbol{u} = [u, v, 1]^T$ and $\boldsymbol{x} = [x, y, z, 1]^T$ are the homogeneous coordinates of an image pixel (2D) and corresponding 3D voxel of the world coordinate system. The symbol $\cong$ is used to emphasize that the equality is up to scale. This matrix $\boldsymbol{P}$ represents all the imaging geometry parameters, which can be divided into two sets. The first set is called the extrinsic parameters, described by the 4×4 homogeneous transformation matrix $\boldsymbol{E}$. They define position ($\boldsymbol{t}$) and orientation ($\boldsymbol{R}$) of the imaging system in the world coordinate system. The second set is called the intrinsic parameters and is described by a (3×3) matrix $\boldsymbol{A}$. These parameters only depend on internal parameters of the imaging system.

$$\boldsymbol{P} = \boldsymbol{A} \cdot \boldsymbol{E} = \begin{bmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \tag{2}$$

It is interesting to note that the intrinsic parameters of a moving X-ray system depend in general on its extrinsic parameters. In the next section we describe this point in more detail and propose to use virtual detector [13] in order to fix the intrinsic parameters for the X-ray imaging system.

3.2 Virtual detector plane

Due to the weights of both X-ray source and detector the C-arm will undergo minor torsion during the motion. Because of that and the large distance between focal point (X-ray source) and image plane (detector) the intrinsic parameters will not be as constant as it can be assumed for standard CCD cameras. In fact they depend on the extrinsic parameters of the imaging system.

Figure 1 demonstrates the differences between the projection geometry for a CCD camera (left) and an X-ray system (right). The intrinsic parameters of a

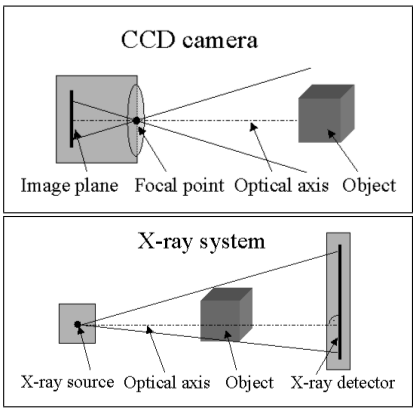

Figure 1: Pinhole camera model: Differences between CCD camera (top) and X-ray imaging system (bottom)

CCD camera only depend on the manufacturing of the camera and the relation between image plane, focal point, and optical axis stay fixed. For an X-ray imaging system, focal point (X-ray source) and image plane (X-ray detector) are far apart from each other, and the optical axis (normal dropping from X-ray source on the detector plane) depends on the orientation of the detector plane. Especially if the X-ray system is moving, the intrinsic parameters according to the pinhole camera model are not fix, thus cannot be determined once and assumed constant afterwards.

In order to use the framework of hand-eye calibration we need to ensure that our X-ray imaging system can be treated as a regular pinhole camera. We found a solution to solve this problem by introducing the concept of a *virtual detector plane* [13]. Using a small number of markers on a plate attached to the X-ray source that are projected close to the borders of the X-ray image we can define a common (virtual) detector plane for all images by warping the images such that these markers are projected to pre-defined positions. Thereby we virtually fix the intrinsic parameters of the X-ray imaging system. For the 3D reconstruction this warping has no influence as long as the projection matrices that are used reflect this changed image plane.

The warping is described by a 2D-2D planar transformation $\boldsymbol{H}$. Instead of actually modifying the images – which would be a time consuming procedure – the planar transformation can be combined with the original projection matrix in order to get a projection matrix with the changed intrinsic geometry:

$$\boldsymbol{P}_i^{\mathrm{VD}} := \boldsymbol{A}^{\mathrm{VD}} \cdot \boldsymbol{E}_i = (\boldsymbol{H}_i \cdot \boldsymbol{A}_i) \cdot \boldsymbol{E}_i = \boldsymbol{H}_i \cdot \boldsymbol{P}_i. \quad (3)$$

3.3 General concept

Now we can present the necessary components for a unified framework of recovering the projection geometry.

1. *Virtual detector:* We need to model the X-ray imaging system by a pinhole camera model. The problem and the solution to it are described in detail in section 3.2. The resulting 2D-2D planar transformation $\boldsymbol{H}_i$ is applied to ensure fixed intrinsic parameters throughout the image sequence.

2. *Reference frame:* For one arbitrary image frame the projection geometry is determined simultaneously for both X-ray ($\boldsymbol{P}^X_{\mathrm{ref}}$) and the sensor system that is used for the online calibration procedure ($\boldsymbol{P}^S_{\mathrm{ref}}$).

3. *Motion estimation:* The motion between current and reference frame, described by $\boldsymbol{M}^S_i$, is computed using the sensor system for online calibration.

Combining these components leads to the following computation of X-ray projection matrix $\boldsymbol{P}^X_i$:

$$\boldsymbol{P}^X_i = \boldsymbol{H}_i \cdot \boldsymbol{P}^X_{\mathrm{ref}} \cdot \boldsymbol{M}^S_i \qquad \text{with } \boldsymbol{P}^S_i = \boldsymbol{P}^S_{\mathrm{ref}} \cdot \boldsymbol{M}^S_i \quad (4)$$

The motion estimation is the only part that will be different for various approaches. The next sections describe the application of the developed framework to the two methods using an integrated CCD camera or external tracking systems.

3.4 Camera-Augmented C-arm (CAMC)

This method for the recovery of projection geometry has been presented in [13]. A CCD camera is attached to the X-ray source of the C-arm.

An optical marker system – not visible in the X-ray image – is used to determine the projection geometry of the CCD camera for each frame. For a reference frame both X-ray and CCD camera projection geometries are determined simultaneously using a combined optical and X-ray calibration phantom. The motion between current and reference frame is computed from the estimated CCD camera projection matrices. Both marker systems are defined in the same world coordinate system. Thus the motion estimated from CCD camera projection matrices can be applied to the reference X-ray projection matrix in order to get the X-ray projection matrix for the current frame according to Equation 4.

The most relevant difference between this approach and the method using external tracking systems is that

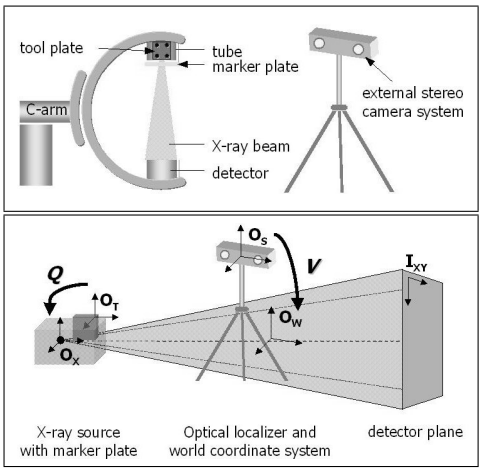

Figure 2: Schematic setup of our test system (top) and relations between used coordinate systems (bottom)

both the sensor for projection geometry recovery and the X-ray imaging system have almost the same viewing direction – both optical axes are almost parallel – for each image frame.

3.5 Recovering projection geometry using external tracking systems

We use the optoelectronic tracking system POLARIS from Northern Digital, Inc. in our experiments. This system consists of a double CCD camera system and marker plates with infrared LEDs. One such marker plate, also called *tool plate*, is attached to the C-arm. Its position $\boldsymbol{t}$ and orientation $\boldsymbol{R}$ relative to the stereo camera system is tracked with high accuracy in real-time. This transformation from sensor coordinate system $\mathcal{O}_S$ to tool plate coordinate system $\mathcal{O}_T$ is described by a 4×4 matrix of homogeneous coordinate transformation and will be denoted by $\boldsymbol{L}$.

The tool plate is rigidly attached to the frame of the X-ray source. The relation between X-ray coordinate system $\mathcal{O}_X$ and $\mathcal{O}_T$ is described by the homogeneous transformation $\boldsymbol{Q}$. The sensor coordinate system $\mathcal{O}_S$ and world coordinate system $\mathcal{O}_W$ have a fixed relation during the calibration setup which is described by the homogeneous transformation $\boldsymbol{V}$. See Figure 2 for a visualization of the different coordinate systems and their relations.

Both transformations $\boldsymbol{Q}$ and $\boldsymbol{V}$ need to be calibrated in an offline procedure. We show in the following that the mathematical framework attached to this system is the same as the one known as *hand-eye calibration* in the vision and robotics community.

3.5.1 Calibration procedure

The transformation from world coordinate system $\mathcal{O}_W$ into X-ray camera coordinate system $\mathcal{O}_X$ for each image frame can be described by the external pose tracker:

$$\boldsymbol{E}_i = \boldsymbol{Q} \cdot \boldsymbol{L}_i \cdot \boldsymbol{V}^{-1}. \tag{5}$$

The determination of the two unknown homogeneous transformations $\boldsymbol{Q} : \mathcal{O}_X \longmapsto \mathcal{O}_T$ and $\boldsymbol{V} : \mathcal{O}_S \longmapsto \mathcal{O}_W$ defines a mathematical problem, which turns to be the same as a well-known calibration task in robotics called *hand-eye calibration*.

A camera is mounted at the gripper of a robot whose pose can be controlled. In order to move the robot such that the camera gets into a certain position and orientation relative to a target the relation of these two coordinate systems have to be determined. We describe the analogy between the two systems as follow: In our application the 3D-3D transformation measured by the pose tracker is the equivalent of the robot's gripper pose while the X-ray imaging system takes the place of the attached camera. This observation lets us use several algorithms, which have been proposed to solve the hand-eye calibration problem. Different authors proposed to solve it by linear and non-linear systems [14, 16, 3, 8, 4]. These algorithms use motion equations between pairs of image frames to solve for the two unknown transformations separately.

The motion of C-arm between two arbitrary image frames is described by

$$\begin{aligned} \boldsymbol{E}_i^{-1} \cdot \boldsymbol{E}_{\text{ref}} &= \boldsymbol{V} \cdot \boldsymbol{L}_i^{-1} \cdot \boldsymbol{Q}^{-1} \cdot \boldsymbol{Q} \cdot \boldsymbol{L}_{\text{ref}} \cdot \boldsymbol{V}^{-1} \\ &= \boldsymbol{V} \cdot \boldsymbol{L}_i^{-1} \cdot \boldsymbol{L}_{\text{ref}} \cdot \boldsymbol{V}^{-1}. \end{aligned} \tag{6}$$

The left side of (6) describes the motion of the X-ray camera between two image frames in world coordinates. This motion can be computed from the projection matrices. Traditionally this is done by decomposing the projection matrices [6, 5] into extrinsic and intrinsic parameters and then computing the motion from the extrinsic parameters. Different authors [8, 12] have emphasized that the decomposition of projection matrices is an unstable process, and an approach for computing the motion between two image frames from projection matrices without decomposing them has been proposed using the *direct motion estimation* algorithm [12].

Similar to (6) we can write a motion equation and eliminate $\boldsymbol{V}$:

$$\begin{aligned} \boldsymbol{E}_{\text{ref}} \cdot \boldsymbol{E}_i^{-1} &= \boldsymbol{Q} \cdot \boldsymbol{L}_{\text{ref}} \cdot \boldsymbol{V}^{-1} \cdot \boldsymbol{V} \cdot \boldsymbol{L}_i^{-1} \cdot \boldsymbol{Q}^{-1} \\ &= \boldsymbol{Q} \cdot \boldsymbol{L}_{\text{ref}} \cdot \boldsymbol{L}_i^{-1} \cdot \boldsymbol{Q}^{-1} \end{aligned} \tag{7}$$

Here a motion in camera coordinates is described. This motion cannot be estimated directly from the projection matrices. Using the following substitution proposed in [8] leads to an equation where a motion in world coordinate system is described: Let us assume that the extrinsic parameters $\boldsymbol{E}_{\text{ref}}$ of the projection matrix of the reference frame are known. The new unknown transformation $\boldsymbol{S}$ is introduced such that $\boldsymbol{Q} = \boldsymbol{E}_{\text{ref}} \cdot \boldsymbol{S}$. This new matrix $\boldsymbol{S}$ describes a transformation from world coordinate system $\mathcal{O}_W$ into the coordinate system of the tool plate $\mathcal{O}_T$ for the reference frame. Equation 7 then becomes:

$$\begin{aligned} \boldsymbol{E}_{\text{ref}} \cdot \boldsymbol{E}_i^{-1} &= \boldsymbol{E}_{\text{ref}} \cdot \boldsymbol{S} \cdot \boldsymbol{L}_{\text{ref}} \cdot \boldsymbol{L}_i^{-1} \cdot \boldsymbol{S}^{-1} \cdot \boldsymbol{E}_{\text{ref}}^{-1} \\ \boldsymbol{E}_i^{-1} \cdot \boldsymbol{E}_{\text{ref}} &= \boldsymbol{S} \cdot \boldsymbol{L}_{\text{ref}} \cdot \boldsymbol{L}_i^{-1} \cdot \boldsymbol{S}^{-1} \end{aligned} \quad (8)$$

Both unknown coordinate transformations $\boldsymbol{S}$ and $\boldsymbol{V}$ are expressed by similar equations. The motion described by the extrinsic parameters $\boldsymbol{E}$ can be directly computed from the projection matrices without decomposing them. With

$$\begin{aligned} \boldsymbol{E}_{(i)} &:= \boldsymbol{E}_i^{-1} \cdot \boldsymbol{E}_{\text{ref}} \\ \boldsymbol{L}_{(i)} &:= \boldsymbol{L}_i^{-1} \cdot \boldsymbol{L}_{\text{ref}} \end{aligned}$$

and $\boldsymbol{E}_{(i)}$, $\boldsymbol{L}_{(i)}$ and $\boldsymbol{V}$ being of the form $\begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix}$ Equation 6 – and similar to that Equation 8 – can be split into a matrix and a vector equation:

$$\boldsymbol{R}_{E_{(i)}} \boldsymbol{R}_V = \boldsymbol{R}_V \boldsymbol{R}_{L_{(i)}} \quad (9)$$

$$\left(\boldsymbol{R}_{E_{(i)}} - \boldsymbol{I}\boldsymbol{d}_3\right) \boldsymbol{t}_V = \boldsymbol{R}_V \boldsymbol{t}_{L_{(i)}} - \boldsymbol{t}_{E_{(i)}} \quad (10)$$

Two different strategies exist in order to solve the above two equation systems. Most approaches regard the two equations decoupled. First the rotation equation (9), and after that the translation equation (10) is solved using the resulting rotation.

Using special properties of rotation matrices, Equation 9 can be transformed into

$$\boldsymbol{r}_{E_{(i)}} = \boldsymbol{R}_V \cdot \boldsymbol{r}_{L_{(i)}}. \quad (11)$$

with $\boldsymbol{r}_{E_{(i)}}$ and $\boldsymbol{r}_{L_{(i)}}$ representing the eigenvectors (rotation axes) of $\boldsymbol{R}_{E_{(i)}}$ and $\boldsymbol{R}_{L_{(i)}}$. This equation system can be solved from two or more equations, either directly [14], or representing the unknown rotation matrix $\boldsymbol{R}_V$ as a quaternion [8]. Equation 9 directly can be transformed into a quaternion equation which is then solved using the constraint of a resulting unit quaternion [3]. The translation is estimated afterwards using a linear least square method solving an equation system according to Equation 10. Note that the matrix $(\boldsymbol{R}_{E_{(i)}} - \boldsymbol{I}\boldsymbol{d}_3)$ has rank 2 and can therefore not be inverted.

A few methods try to solve both equations simultaneously. Both equations 9 and 10 can be combined and solved by a non-linear minimization [8]. Danilidis [4] uses dual quaternions, an algebraic representation from screw theory, to describe the motions. This leads to a linear minimization of rotation and translation simultaneously.

From the determined transformations $\boldsymbol{S}$ and $\boldsymbol{V}$ we now want to compute the X-ray projection matrices. We derive the following equation starting with the definition of the X-ray projection matrix:

$$\begin{aligned} \boldsymbol{P}_i^X &= \boldsymbol{A}_i \cdot \boldsymbol{E}_i && \text{(see Eq. 2)} \quad (12) \\ &= \boldsymbol{A}_i \cdot \boldsymbol{Q} \cdot \boldsymbol{L}_i \cdot \boldsymbol{V}^{-1} && \text{(see Eq. 5)} \\ &= \boldsymbol{H}_i \cdot \boldsymbol{A}_{\text{ref}} \cdot \boldsymbol{Q} \cdot \boldsymbol{L}_i \cdot \boldsymbol{V}^{-1} && \text{(see Eq. 3)} \\ &= \boldsymbol{H}_i \cdot \boldsymbol{A}_{\text{ref}} \cdot \boldsymbol{E}_{\text{ref}} \cdot \boldsymbol{S} \cdot \boldsymbol{L}_i \cdot \boldsymbol{V}^{-1} && \text{(def. of } \boldsymbol{S}) \\ &= \boldsymbol{H}_i \cdot \boldsymbol{P}_{\text{ref}}^X \cdot \boldsymbol{S} \cdot \boldsymbol{L}_i \cdot \boldsymbol{V}^{-1} && (13) \end{aligned}$$

Note that although in Equation 12 the projection matrix is presented as product of intrinsic and extrinsic parameters, we do not have to compute any intrinsic or extrinsic parameters, but use a reference projection matrix $\boldsymbol{P}_{\text{ref}}$.

The resulting equation (13) reflects the general framework we developed in section 3.3. The motion between current and reference frame in world coordinate system for our approach using external tracking system is:

$$\boldsymbol{M}_i^S = \boldsymbol{S} \cdot \boldsymbol{L}_i \cdot \boldsymbol{V}^{-1} \quad (14)$$

Due to the definition of $\boldsymbol{S}$, describing the transformation from the tool plate coordinate system of the reference frame to the world coordinate system, above equation indeed describes a motion in the world coordinate system measured by the external tracking system.

4 Experimental Results

This section presents the experimental results. We first describe the special setup that was necessary for the hand-eye calibration. We then evaluate the two approaches for online recovery of X-ray projection geometry. As we have tested several hand-eye calibration algorithms, we also want to compare their performance for the external tracking system approach. We evaluate these methods both numerically and based on the visual result of 3D tomographic reconstruction.

4.1 Calibration setup

During the examination the C-arm rotates around an axis going through the center of the volume of interest which is to be reconstructed. However, in order

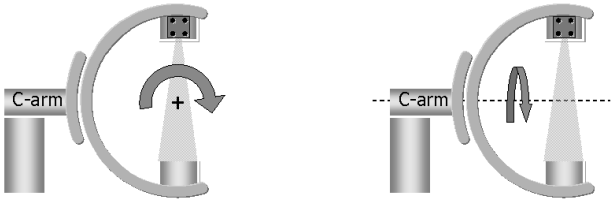

Figure 3: Setup for calibration: C-arm in normal position during rotation (left) and in special position (angulated) for the calibration procedure (right)

to be able to compute the necessary transformation matrices, we need the C-arm to have at least two rotational motions around different rotation axes [16]. This is not a natural motion for a C-arm used for 3D reconstruction. Fortunately, this must only be done during the off-line calibration process. This is visualized in Figure 3.

The image series used for calibration consist of 85 images of the normal patient run plus 8 images where the C-arm is angulated. The direct recovery of X-ray projection geometry – computing projection matrices $\boldsymbol{P}_i^X$ – is done using a cylindrical X-ray phantom as in [12], which can be seen on the table in Figure 3. This set of projection matrices is used as *gold standard* to which we will compare the other methods for both reconstruction results and numerical evaluation.

4.2 Reconstruction results

The set of projection matrices that has been computed for each method is used to reconstruct a test phantom. It consists of two cylindrical objects in an acrylic cover. Attached to it is a ring made from titanic alloy also in an acrylic cover. The results are shown in figures 4 and 5. Visualized are the 3 orthogonal Maximum Intensity Projections (MIP) of the reconstructed volume. These are the methods that have been evaluated:

- CAMC: Camera-augmented C-arm, described in section 3.4
- CHOU: method of Chou and Kamel [3], implementation courtesy of K. Daniilidis
- DANI: method of Daniilidis [4], implementation courtesy of K. Daniilidis
- HORlin: linear method proposed by Horaud and Dornaika in [8], based on Shiu and Ahmad [14]
- HORnlin: nonlinear method proposed by Horaud and Dornaika [8], implementation courtesy of K. Daniilidis

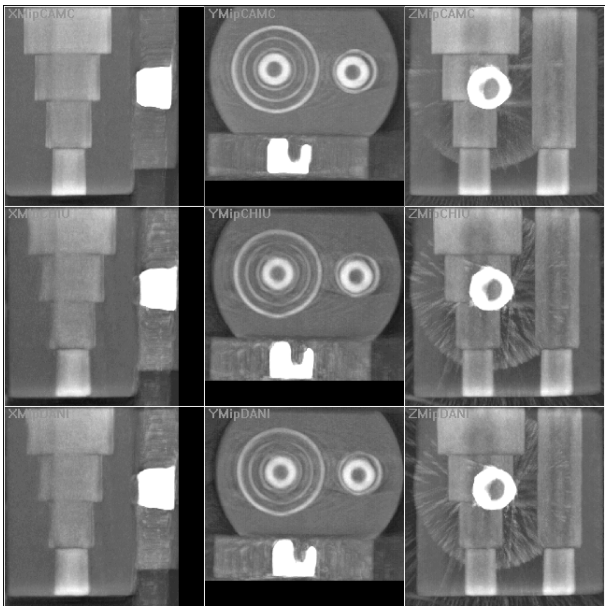

Figure 4: Experimental results of 3D reconstruction using different calibration methods (from top to bottom): CAMC, CHOU, DANI

- TSAI: method of Tsai and Lenz [16]
- XRAY: *Gold standard*, X-ray projection matrices obtained during calibration run are used for 3D reconstruction

All methods lead to satisfying results, and visually the differences in reconstruction quality are small. Nevertheless, both the gold standard of using the X-ray projection matrices and the CCD camera approach lead to better reconstruction quality than the external tracking system method. Differences between the various hand-eye calibration methods for the external tracking system approach are negligible.

4.3 Numerical Evaluation

Apart from the visual results we also want to compare the methods numerically. We compute the transformation (in world coordinate system) between projection matrices of each method and the corresponding gold standard X-ray projection matrix. The resulting errors angle of rotation and norm of translation are visualized in tables 1 and 2. At first glance the errors for the sensor methods are smaller in both rotation and translation than for CAMC. This is surprising, because the quality of 3D reconstruction is better for CAMC than any of the external tracking system methods.

According to an experimental evaluation of different optical tracking systems in [2] we expect the mean

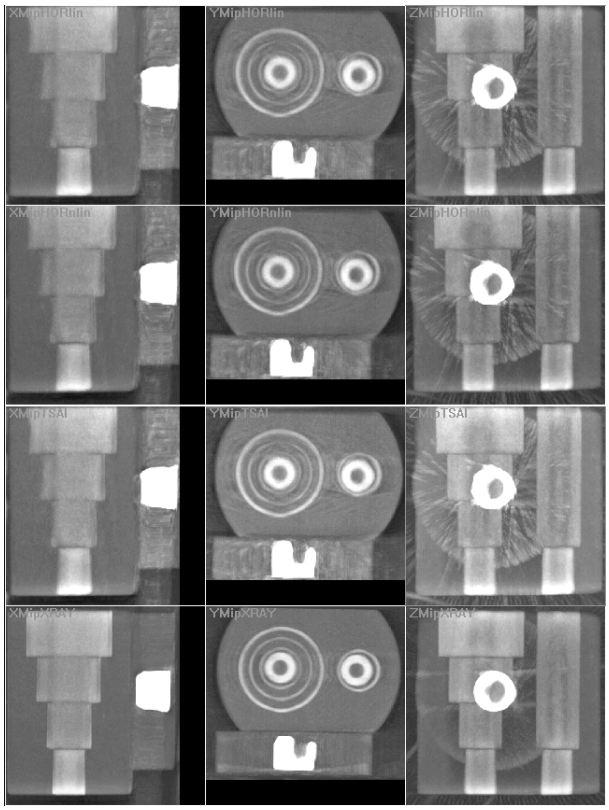

Figure 5: Experimental results of 3D reconstruction using different calibration methods (from top to bottom): HORlin, HORnlin, TSAI, XRAY (gold standard)

error of the POLARIS system to be about $t_{err} \approx 0.6$ mm and $\theta_{err} \approx 0.25$ deg. This covers the error in rotation, but only a part of translation. The differences between the various hand-eye calibration methods are negligible concerning the rotational error.

The translation error can be decomposed into two components, translation in the direction of the optical axis of the X-ray system, and translation perpendicular to this axis. Translation along the optical axis has very little effect on the recovery of projection geometry, this error parallel to the optical axis is of much less importance than the error perpendicular to the optical axis. If we compare the different methods based on the

	CAMC	CHOU	DANI	HORlin	HORnlin	TSAI
μ	0.50	0.16	0.22	0.16	0.30	0.14
σ	0.20	0.06	0.08	0.06	0.10	0.07

Table 1: Comparison between external sensor and integrated camera method: Rotation errors

method	t_{err}		$t_{\parallel opt}$		$t_{\perp opt}$	
	μ	σ	μ	σ	μ	σ
CAMC	1.39	0.71	1.13	0.87	0.58	0.21
CHOU	0.72	0.35	0.30	0.22	0.62	0.36
DANI	0.69	0.35	0.29	0.21	0.59	0.36
HORlin	0.72	0.35	0.30	0.22	0.61	0.36
HORnlin	0.73	0.34	0.30	0.22	0.62	0.36
TSAI	0.73	0.35	0.31	0.22	0.62	0.36

Table 2: Comparison between external sensor and integrated camera method: Mean and standard deviation of translation error t_{err}, error in direction of the optical axis of the X-ray system $t_{\parallel opt}$, and error perpendicular to optical axis of X-ray system $t_{\perp opt}$

translation error perpendicular to the optical axis of the X-ray system, the CAMC method leads to better numerical results than the external sensor methods, which explains the better quality of reconstruction. The mean errors are very similar but the standard deviation is significantly smaller for the CAMC method. A large amount of error for CAMC is along the optical axis. This is not surprising as we expect the most error to be along the optical axis of the attached CCD camera, which is almost parallel to the X-ray's optical axis. Although the stereo camera system of the pose tracker is much more precise in recovering the motion, its optical axis is always perpendicular to the optical axis of the X-ray system. So only a small fraction of the error is along the direction of the optical axis of the X-ray system, where it would have the least influence on the quality of 3D reconstruction. Similar to the rotation errors there are only small differences between the hand-eye calibration methods for the external tracking system approach.

5 Conclusion

In this paper we presented a method for the use of external tracking or localization sensors for recovering projection geometry of a moving X-ray system. This projection geometry is then used for 3D tomographic reconstruction. We first show that using external sensors for motion estimation is mathematically equivalent to the well-known hand-eye calibration problem. Then, we compare the results of X-ray geometrical calibration using external sensors with one using a CCD camera that is attached to the X-ray source. These results are evaluated both in terms of the accuracy of motion estimation, and the quality of the final 3D reconstruction. Experimental results show that external sensors using hand-eye calibration methods provide more accurate motion estimation. However, the

quality of the 3D reconstruction is better when using the integrated CCD camera. This work clearly shows that the main part of the translation error for the integrated CCD camera is always towards the optical axis of both systems, and therefore has only a small influence on the quality of 3D reconstruction. For the external sensor, the main part of translation error is always perpendicular to the optical axis of the X-ray system, and therefore has a large influence on the quality of 3D reconstruction. This conclusion is of major importance for both medical imaging and augmented reality communities. The main message here is that the accuracy of calibration methods and apparatus should be defined only in terms of the goals of each particular application. In particular, for image back-projection for tomographic reconstruction an integrated monocular imaging system, performs the necessary geometrical calibration better than a professional and highly accurate (stereo-based) external sensor.

Acknowledgements

Authors want to thank Kostas Daniilidis, Frank Sauer, Mariappan Nadar, Oliver Schütz, Rainer Graumann and Alok Gupta for active support and fruitful discussions. The authors would also like to thank Karl Wiesent and Wolfgang Seißler for providing the 3D reconstruction software.

References

[1] R. Azuma. A survey of augmented reality. *Presence: Teleoperators and Virtual Environment*, 6(4):355 – 385, 1997.

[2] F. Chassat and S. Lavallée. Experimental protocol of accuracy evaluation of 6-D localizers for computer-integrated surgery: Application to four optical localizers. In *First International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 277–284, Cambridge, MA, USA, 1998.

[3] J. Chou and M. Kamel. Finding the position and orientation of a sensor on a robot manipulator using quaternions. *International Journal of Robotics Research*, 10(3):240–254, 1991.

[4] K. Daniilidis. Hand-eye calibration using dual quaternions. *International Journal of Robotics Research*, 18(3):286–298, 1999.

[5] O.D. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Cambridge, MA, 1993.

[6] S. Ganapathy. Decomposition of transformation matrices for robot vision. In *Proceedings International Conference on Robotics and Automation*, pages 130–139, 1984.

[7] W. A. Hoff. Fusion of data from head-mounted and fixed sensors. In *IEEE International Workshop on Augmented Reality*, San Francisco, CA, USA, 1998.

[8] R. Horaud and F. Dornaika. Hand-eye calibration. *International Journal of Robotics Research*, 14(3):195–210, 1995.

[9] http://www.mrcas.ri.cmu.edu/projects/overlay.html.

[10] M. Mitschke, N. Navab, and O. Schütz. Online calibration of a mobile C-arm using external sensors. In *Proceedings of SPIE Medical Conference*, San Diego, CA, USA, 2000. Springer.

[11] N. Navab, A. Bani-Hashemi, M. Mitschke, D. W. Holdsworth, R. Fahrig, A. J. Fox, and R. Graumann. Dynamic geometrical calibration for 3D cerebral angiography. In *Proceedings of SPIE Medical Conference*, pages 361–370, Newport Beach, California, USA, 1996.

[12] N. Navab, A. Bani-Hashemi, M. S. Nadar, K. Wiesent, P. Durlak, T. Brunner, K. Barth, and R. Graumann. 3D reconstruction from projection matrices in a C-arm based 3D-angiography system. In *First International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 119–129, Cambridge, MA, USA, 1998.

[13] N. Navab, M. Mitschke, and O. Schütz. Camera augmented mobile C-arm (CAMC) application: 3D reconstruction using a low-cost mobile C-arm. In *Second International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 688–697, Cambridge, England, 1999.

[14] Y. C. Shiu and S. Ahmad. Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form AX=XB. *IEEE Transactions on Robotics and Automation*, 5(1):16–29, 1989.

[15] H. Tamamura, H. Yamamoto, and A. Katayama. Steps towards seamless mixed reality. In *Proc. ISMR*, 1999.

[16] R. Y. Tsai and R. K. Lenz. A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. *IEEE Transactions on Robotics and Automation*, 5(3):345–358, 1989.