

# Determining the Point of Minimum Error for 6DOF Pose Uncertainty Representation

Daniel Pustka\*, Jochen Willneff, Oliver Wenisch, Peter Lükewille, Kurt Achatz  
Advanced Realtime Tracking GmbH

Peter Keitler†, Gudrun Klinker  
Technische Universität München

## ABSTRACT

In many augmented reality applications, in particular in the medical and industrial domains, knowledge about tracking errors is important. Most current approaches characterize tracking errors by  $6 \times 6$  covariance matrices that describe the uncertainty of a 6DOF pose, where the center of rotational error lies in the origin of a target coordinate system. This origin is assumed to coincide with the geometric centroid of a tracking target.

In this paper, we show that, in case of a multi-camera fiducial tracking system, the geometric centroid of a body does not necessarily coincide with the point of minimum error. The latter is not fixed to a particular location, but moves, depending on the individual observations. We describe how to compute this point of minimum error given a covariance matrix and verify the validity of the approach using Monte Carlo simulations on a number of scenarios. Looking at the movement of the point of minimum error, we find that it can be located surprisingly far away from its expected position. This is further validated by an experiment using a real camera system.

**Index Terms:** I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Tracking H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

## 1 INTRODUCTION

Recent years have seen more and more tracking applications requiring a permanent on-line control of registration error or measurement uncertainty. This kind of information is important for augmented reality in general, but in particular for areas such as surgical navigation or quality control in industrial measurement. In vendor-calibrated tracking systems with rigid configuration and pre-defined working volume this kind of information is readily available. In flexible configurations of multi-camera setups and under varying viewing conditions in dynamic scenes, however, a full error description becomes more difficult and cannot easily be interpreted anymore. Also, the real-time description of varying uncertainty in pose estimation has only recently been addressed in more detail ([17], [2], [10]). Most approaches make use of the covariance matrix as a means of quantifying measurement errors and for the visualization of error ellipsoids. Less attention has been paid to the structure and interpretation of the covariance and its meaning in different reference frames. This will be investigated in the paper.

**Multi-Camera Fiducial Tracking** The tracking system used in this investigation consists of an n-ocular CCD camera system. The cameras detect the pose of a measurement target consisting of a rigidly attached set of fiducials, i.e. small spheres with retro-reflective coating. These fiducials reflect the light of infrared flashes, which are attached to the tracking cameras. Besides the details

\*e-mail: daniel.pustka@ar-tracking.de

†e-mail: keitler@in.tum.de

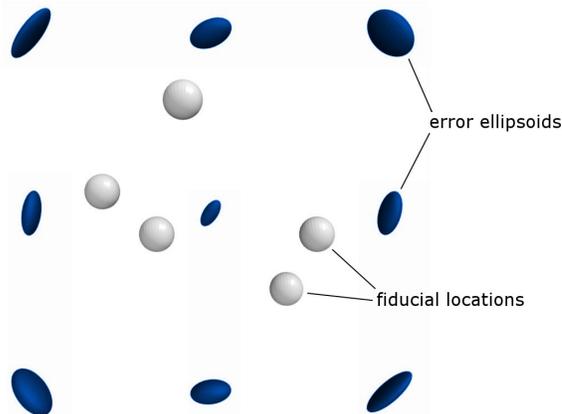


Figure 1: Distribution of translation errors (black) sampled at 9 different CREs for a fiducial target defined by the white spheres. It is directly evident that the samples more distant to the body's geometrical centroid represent CREs with a larger translational uncertainty due to the additional influence of rotation errors.

of the underlying computer vision algorithms, the accuracy of this kind of system depends on the particular geometry of the camera configuration, the size, number and spatial arrangement of fiducials on the target and the visibility conditions for each of the markers. As the marker visibility changes according to the movement and interaction of objects within the tracking volume, the error estimation for a target has to be updated dynamically.

**Center of Rotational Error** When computing the 6DOF object pose from the 2D marker reflexes measured by the cameras, the translation and orientation of the target is expressed with respect to a specific coordinate system, the target coordinate system. The origin of this reference frame can be chosen freely and defines the center of rotational error (CRE), for which the corresponding covariance matrix is computed. The matrix generally subsumes the translation and orientation uncertainty and by changing from one reference frame to another, the translatory error will be more or less polluted by rotational errors. This is depicted in Figure 1, where we show the covariance error ellipsoids for 9 different CRE locations, with all errors clearly larger than the one at the target center.

It is also obvious that for every visibility situation there is a distinguished reference frame which has minimum translation error. We call the origin of this coordinate system the point of minimum error (PME). In Figure 2 one can compare the different outcomes of covariance computation at different CREs for an elongated target which is not fully visible. As opposed to the large translational uncertainty at the geometrical centroid (GC) the error is considerably smaller at the PME.

**Contribution** In this paper, we show that, in case of a multi-camera fiducial tracking system, the center of rotational error can be set, in principle, into practically any point without obvious disadvantages. As the reference frame with minimum translational error

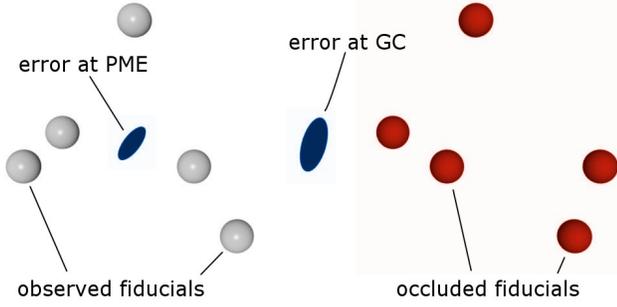


Figure 2: Comparison of covariances (black ellipsoids) computed at the geometric centroid (GC) and at the point of minimum error (PME) for a target whose fiducials are not fully visible to both of two cameras (white spheres visible, gray spheres occluded). Obviously, the translational error differs considerably between choosing the GC or the PME as the CRE.

is not fixed but changes according to the individual camera observations, however, there is a prominent location, namely the point of minimum error, which maximally separates translational from rotational errors.

After introducing this point of minimum error, we describe how to compute its position from a given covariance matrix. Therefore, we show that the covariance matrix contains all information to compute this distinguished point just by the knowledge of the transformation properties of the different parts of the covariance matrix.

The validity of the approach is verified using Monte Carlo simulations by comparing the analytical and numerical results of PME computation.

**Related Work** There have been a number of publications dealing with error estimation of tracking systems. One of the oldest works on pose error is by Woltring et al. [18] who analytically derive the effects of isotropic 3D error on an isotropic distribution of fiducials and who observe that the error is minimal at the centroid of the fiducials.

Fitzpatrick [7] give a formula for estimating the target registration error based on the simplifying assumption of an isotropic fiducial location error.

In the AR community, the field of error analysis of AR systems has drawn much attention. Holloway [11] gives a comprehensive overview and derives error bounds for various error sources. Error estimation for optical trackers was investigated by Hoff and Vincent [10], Davis et al. [5][6], Allen and Welch. [1] and Bauer et al. [2]. Most publications state the need for computing the error at the centroid, but do not investigate the choice of the CRE in more detail.

Coelho et al. [4] use the unscented transformation [12] (UT) to propagate tracking errors through a scene graph into the image shown to an AR user. UT is generally considered to be superior to the linear propagation of covariance that we use, due to better handling of non-linearities. However, for the small rotational errors that we consider, the error propagation is sufficiently linear. We will verify this using Monte Carlo simulations later. Furthermore, the general problem of CRE choice remains, independently of the method of error propagation.

## 2 ERROR REPRESENTATION

In order to address the problem of a moving center of rotational error, we extend the standard error representation to explicitly take this point into account. The resulting extended error representation is given by equation 1. It describes the transformation of a point  $x_T$  in target coordinates to its image  $x_W$  in world coordinates. The

transformation from target to world coordinates is given by the rotation matrix  $R$  and the translation vector  $t$ .

$$x_W = E_R(Rx_T + t - c) + c + e_t \quad (1)$$

The error itself is modeled as an additional rotation matrix  $E_R$  that rotates around a CRE  $c$ , and a translation vector  $e_t$ . The value of  $c$  can explicitly change from frame to frame. Note that  $E_R$ ,  $e_t$  and  $c$  are given in world coordinates.

Both  $E_R$  and  $e_t$  are assumed to be *small* random variables with expectations  $I$  and  $0$  respectively. For the error propagation, we approximate  $E_R$  using a small-angle version of Rodrigues' rotation formula:

$$E_R = I_{3 \times 3} + [e_R]_{\times} \quad (2)$$

where  $[v]_{\times}$  denotes the antisymmetric matrix corresponding to a 3-vector  $v$ :

$$[v]_{\times} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} \quad (3)$$

The pose uncertainty is given by a  $6 \times 6$  covariance matrix  $\Sigma$  that describes the joint Gaussian distribution of  $e_t$  and  $e_R$ :

$$\begin{bmatrix} e_t \\ e_R \end{bmatrix} \sim N(0, \Sigma) \quad (4)$$

Intuitively, the diagonal entries of  $\Sigma$  can be interpreted as the positional uncertainty in world coordinates ( $\Sigma_{1,1}$  to  $\Sigma_{3,3}$ ) and the rotational uncertainty around the three axes of the world coordinate system ( $\Sigma_{4,4}$  to  $\Sigma_{6,6}$ ), given in radians. Off-diagonal values describe the statistical dependency between two values.

According to this definition, our error representation consists of the tuple  $(\Sigma, c)$  where  $c$  defines the center of rotational error.

## 3 BACKWARD PROPAGATION OF COVARIANCE

In order to analytically compute the covariance of a 6DOF pose from uncertainties of multiple 2D measurements, we use the backward propagation of covariance (BPC) [8] approach. Since the BPC requires the same intermediate steps as non-linear least-squares optimization, it is easily integrated into a pose estimation algorithm.

In consequence, our system calculates 6DOF target poses  $p$  by a least squares adjustment, where a  $6 \times 6$  covariance matrix  $\Sigma_{6D}$  of the parameters is part of the result. The adjustment procedure minimizes the weighted residuals in 2D image space of the individually projected target fiducial positions  $u$ . This projection consists of two separate steps. The first step basically is a 3D similarity transformation without scale of 3D fiducial positions  $x_T$  from target to world coordinates  $x_W$  as shown in equation 1.

The second step is using the collinearity equation for a pinhole camera projection of this target fiducial location in world coordinates  $x_W$  to the 2D image space coordinates  $u$  for each camera, where this fiducial was detected [13]:

$$u = \begin{bmatrix} x_h \\ y_h \end{bmatrix} - \frac{cc}{r_3^T \cdot (x_W - x_0)} \cdot \begin{bmatrix} r_1^T \cdot (x_W - x_0) \\ r_2^T \cdot (x_W - x_0) \end{bmatrix} \quad (5)$$

where  $x_h$  and  $y_h$  are the coordinates of the principle point and  $cc$  the principle distance of the camera. The vector  $x_0$  represents the world coordinates of the camera projection center and  $r_i$  is the  $i$ -th column of the camera rotation matrix  $R$ :

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \end{bmatrix} \quad (6)$$

The position and orientation of a camera relative to the world coordinate system are generally known as its *extrinsic* parameters.

These are complemented with *intrinsic* parameters such as principle point, principle distance and corrections to compensate for lens distortion and other effects [3].

The derivatives of the 6DOF to 2DOF projection function w.r.t. the target pose  $p = (e_t, e_R)$  form the Jacobian matrix  $J$  for the minimization process.  $J$  is computed by applying the chain rule on the target-to-world transformation and the pinhole projection:

$$J = \frac{\partial u}{\partial p} = \frac{\partial u}{\partial x_W} \cdot \frac{\partial x_W}{\partial p} \quad (7)$$

The chain rule allows for the separation in two derivative matrices with dimensions of  $2 \times 3$  and  $3 \times 6$  respectively. This separation into two steps simplifies the calculation of the derivatives to two well-known approaches. The derivatives of the collinearity equation are equivalent to the ones applied for the forward intersection to calculate a 3D position [13].

As the rotations and translations are given in the world coordinate system, the Jacobian matrix of the similarity transformation part is a  $3 \times 3$  identity matrix for translation and an antisymmetric matrix  $[-v]_{\times} = [-Rx_T]_{\times}$  for rotation:

$$\frac{\partial x_W}{\partial p} = \begin{bmatrix} 1 & 0 & 0 & 0 & v_3 & -v_2 \\ 0 & 1 & 0 & -v_3 & 0 & v_1 \\ 0 & 0 & 1 & v_2 & -v_1 & 0 \end{bmatrix} \quad (8)$$

Multiplying the two derivative matrices shown in 7 yields the  $2 \times 6$  Jacobian matrix  $J$  for the iterative adjustment process [14],

$$\delta p = (J^T P J)^{-1} J^T P u \quad (9)$$

which provides the corrections  $\delta p$  of the target pose parameters as well as the covariance matrix  $\Sigma_{6D}$  given by

$$\Sigma_{6D} = \sigma_0^2 \cdot (J^T P J)^{-1} \quad (10)$$

The observations  $u$  refer to the measured fiducial position in image space and their 2D covariances  $\Sigma_{2D}$  are crucial as they determine the weight matrix  $P = \Sigma_{2D}^{-1}$  of each individual observation. For simplicity, all observations are introduced with the same weight. To avoid a possible mismatch in the stochastic model the standard deviation  $\sigma_0$  is required for scaling of the cofactor matrix  $(J^T P J)^{-1}$ .

The adjustment process uses predictions from previous time steps as approximation for the target pose  $p$ , which generally leads to convergence in only a few iterations.

#### 4 POINT OF MINIMUM ERROR COMPUTATION

In the previous sections, we presented an error representation that explicitly deals with a changing center of rotational error. The goal of the error description is to allow us to place the error into the point of minimum error, where rotational and translational errors are most naturally separated.

Considering the occlusion scenarios mentioned in the introduction, the most natural approach to compute the point of minimum error is to compute the geometric centroid of all observed fiducials. Unfortunately, as our experiments later on will show, this is only an approximation of the PME's location, which in extreme cases can be far from its true location. In the rest of this section, we will present an analytical approach to compute the PME as an offset, given the CRE and the  $6 \times 6$  covariance matrix in any location.

##### 4.1 Forward Propagation of Covariance

The discussion in this section is based on forward propagation of covariances [8]. Let  $f : \mathbb{R}^N \rightarrow \mathbb{R}^M$  be a differentiable function and  $x \in \mathbb{R}^N$  a random variable with expectation  $\bar{x}$  and covariance  $\Sigma_x$ . According to the forward propagation rule,  $y = f(x)$  is a random

variable with expectation  $\bar{y} = f(\bar{x})$ . The covariance  $\Sigma_y$  is computed as:

$$\begin{aligned} \Sigma_y &= J \Sigma_x J^T \\ J &= \frac{\partial f}{\partial x} \end{aligned} \quad (11)$$

with  $J$  being the Jacobian of  $f$ . This approximation is valid for Gaussian distributions when  $f$  is approximately linear within reasonable extent around  $x$ . This is the case for small angular errors, which are typical for the mentioned tracking systems.

##### 4.2 Change of CRE

Using the forward propagation method, we first describe how to propagate covariance from one CRE to another. Assume in the above error representation (eq. 1) that  $c$  is shifted by some offset  $\Delta c$ , i.e.

$$x_W = E_R(Rx_T + t - (c + \Delta c)) - E_R \Delta c + c + e_t \quad (12)$$

By integrating the term  $-E_R \Delta c$  into the translational error  $e_t$ , we compute the error propagation

$$\Sigma_{\Delta c} = J_{\Delta c} \Sigma J_{\Delta c}^T \quad (13)$$

$$J_{\Delta c} = \begin{bmatrix} I & [\Delta c]_{\times} \\ 0 & I \end{bmatrix} \quad (14)$$

where  $\Sigma_{\Delta c}$  is the new covariance matrix at the shifted CRE.

We make the following observations:

**Commutativity** Examining at the structure of  $J_{\Delta c}$ , we can see that the concatenation of two CRE changes  $J_{\Delta c_1} J_{\Delta c_2}$  can be expressed as a single change, as  $J_{\Delta c_1} J_{\Delta c_2} = J_{\Delta c_1 + \Delta c_2}$ .

**Reversibility** As  $J_{\Delta c}$  is invertible, we can revert the change of CRE without loss of information, by applying the propagation again using  $J_{\Delta c}^{-1}$ . We can also show that  $J_{\Delta c}^{-1} = J_{-\Delta c}$ .

##### 4.3 Point of Minimum Error Extraction from Covariance

In order to compute the location of the point of minimum error, we start from a covariance matrix  $\Sigma$ , evaluated for an arbitrary CRE  $c$ . The idea is to find an offset vector  $\Delta c$  for which the CRE change results in a minimal translation error. The resulting  $c + \Delta c$  describes the PME location. Minimizing translational error is the natural choice for finding the PME, as our error model basically defines translational error at some POI as the sum of translational error at the PME plus some rotational error.

We start the computation by looking at the block structure of  $\Sigma$  and how this evaluates into the translational error:

$$\begin{aligned} \Sigma'_t &= \begin{bmatrix} I & [\Delta c]_{\times} \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} I & [\Delta c]_{\times} \end{bmatrix}^T \\ &= A + [\Delta c]_{\times} B^T + B [\Delta c]_{\times}^T + [\Delta c]_{\times} C [\Delta c]_{\times}^T \end{aligned} \quad (15)$$

In order to determine the offset  $\Delta c$  for which  $\Sigma'_t$  is minimal, we have to choose a suitable matrix norm to minimize. In this case, we select the trace norm, as the trace of a covariance matrix is proportional to the RMS error. We therefore seek to minimize the following expression:

$$\arg \min_{\Delta c} \text{tr}(\Sigma'_t) \quad (16)$$

inserting equation 15, computing the derivative and solving for the  $\Delta c$  which minimizes the trace yields the following equation:

$$(\text{tr}(D)I - D)\Delta c = -\text{SKEW}^{-1}(B - B^T) \quad (17)$$

where  $\text{SKEW}^{-1}$  is a function that inverts the effect of  $[\dots]_{\times}$  by extracting a 3-vector from a skew-symmetric matrix as defined by eq. 3.

## 5 VALIDATION

In this section we want to verify the analytic methods described in the previous two sections using a Monte Carlo simulation of the tracking process. Monte Carlo simulation is a quite flexible but computationally expensive means to solve optimization and integration problems [16]. Our problem is of the latter type, since we want to compute the expected deviation (second statistical moment) of some point of interest (PoI) from its mean value. Monte Carlo Simulation was previously used successfully to investigate various influences in photogrammetric systems [9].

### 5.1 Monte Carlo Simulation

The Monte Carlo simulation is based on the assumption of known ground-truth data from which the pose at a desired PoI can be derived. This original data is then perturbed with Gaussian noise and its influence on the PoI observed. The goal is to compute a  $6 \times 6$  covariance matrix that can be compared to the analytic BPC result as described in section 3. We assume the following ground-truth data to be known:

- extrinsic camera parameters
- intrinsic camera parameters
- target geometry
- target pose
- point of interest

The target geometry simply is a list of 3D positions in target coordinates, which describe the constellation of fiducials. The point of interest is the 3D point for which the covariance shall be estimated. This corresponds to the CRE in the BPC approach.

The first step of our Monte Carlo simulation projects fiducials onto the image plane of each camera, based on the known target geometry and pose, as well as the known intrinsic and extrinsic camera parameters. This yields a list of synthetic 2D positions for each camera, perfectly fulfilling the geometric constraints imposed by the assumed ground-truth data.

Based on the assumption of Gaussian pixel noise, we can perturb these 2D positions to obtain artificial measurement data for the layed out geometric situation. This perturbation can be repeated many times, with different noise instantiations. A Gaussian distribution might not represent well the structured noises that are usually present in optical tracking systems, caused e.g. by mis-calibration of the cameras or insufficiencies in the camera model. Nevertheless, a pragmatic approach commonly followed in the field of metrology is to subsume both, systematic error and random noise, under the common notion *uncertainty* and to represent it by a Gaussian [14].

From the known point of interest and the known target pose, we can also derive the ground truth position of interest in the local target coordinate frame. This completes the prerequisites for the actual simulation experiment.

Using the artificial “noisy” 2D measurements of each camera, the known target geometry, as well as the known intrinsic and extrinsic parameters, we now estimate the target pose. This reflects the normal mode of operation of our tracking system, indeed the same algorithms are used also for the simulation. This estimation of the point of interest in world coordinates is repeated many times, using different noise instantiations for the 2D measurements. Given such a large set of poses (1,000,000 samples for each point of interest), we compute the mean pose and a  $6 \times 6$  covariance matrix, which is compared to the analytically computed counterpart.

### 5.2 Reference Scenarios

In order to evaluate the error computation approach, we have defined a number of reference scenarios which we will refer to later on. These consist of five synthetic scenarios that are used to systematically evaluate some effects and one realistic constellation, taken from a calibrated real setup.

**Scenario 1: single camera, planar target** A single camera is placed at a distance of 0.5m away from a planar target. The target consists of four fiducials arranged as a square with 0.1m side-length. The target is seen under an angle of  $45^\circ$ .

**Scenario 2: single camera, 3D target** Similar to the previous scenario, but a fifth fiducial is added such that the target forms a pyramid.

**Scenario 3: two cameras, planar target** Same as scenario 1, but the target is observed by a stereo camera pair with 0.2m base length at a distance of 0.5m.

**Scenario 4: two cameras, 3D target** Stereo camera pair observing the same target as in scenario 2.

**Scenario 5: two cameras, real scenario** Two cameras with arbitrary placement and a target consisting of ten fiducials with non-planar arrangement.

**Scenario 6: two cameras, real scenario, half-occlusion** Same as scenario 5, but half of the fiducials are occluded.

**Scenario 7: two cameras, real scenario, alternating occlusions** Same as scenario 5, but each fiducial is observed by exactly one camera.

### 5.3 Backward Propagation

We first use the Monte Carlo Simulation (MCS) to verify the backward propagation of covariance (BPC) approach. For this, we choose for each scenario a few sample points. These sample points are used as the CRE in the BPC and as the point of interest in the MCS.

An example of two  $6 \times 6$  covariance matrices generated by BPC and MCS is shown in table 1. When comparing both, we find that the entries of the MCS matrix agree with those of the BPC in the first 2-3 digits. However, for the BPC entries that are close to zero (usually  $< 10^{-17} \text{mm}^2$ ), MCS gives much bigger values (around  $10^{-6} \text{mm}^2$ ). As these values also differ strongly between multiple runs of the MCS (even changing the sign), we assume that this behavior can be explained by noise in the MCS, having an insufficient number of iterations (currently 1,000,000 for six DOFs).

### 5.4 Forward Propagation

Next, we empirically evaluate how well the covariance at a particular target CRE can be computed by forward propagation of the covariance from an arbitrary source CRE, i.e. how good the observations from section 4.2 are. For this, we use BPC to directly compute the covariance matrix at the target CRE. This is compared with the covariance computed by forward-propagation (eq. 13) of the covariance matrix evaluated for another CRE.

We find that, even when the distance between the two starting CREs is large, the resulting covariance matrices at the target CRE are equal, up to small numerical differences.

This empirical result strengthens our assumption that the measurement error for this kind of measurement system can be modeled as a covariance at a point of minimum error plus some part that is only dependent on the rotational error. For small rotational uncertainties, we therefore make no significant error when applying the linear(!) forward propagation and all non-linearity is contained in the backward propagation step.

$$\begin{bmatrix} 0.036735 & -2.5672 \cdot 10^{-17} & 7.1835 \cdot 10^{-17} & -4.7622 \cdot 10^{-20} & -0.00015546 & -3.8355 \cdot 10^{-5} \\ -2.5672 \cdot 10^{-17} & 0.010308 & -0.00058889 & -8.2805 \cdot 10^{-5} & 2.0947 \cdot 10^{-19} & -1.4972 \cdot 10^{-20} \\ 7.1835 \cdot 10^{-17} & -0.00058889 & 0.44315 & -0.00013705 & 1.0872 \cdot 10^{-19} & -2.2394 \cdot 10^{-19} \\ -4.7622 \cdot 10^{-20} & -8.2805 \cdot 10^{-5} & -0.00013705 & 1.9162 \cdot 10^{-6} & -5.5887 \cdot 10^{-22} & 4.1259 \cdot 10^{-22} \\ -0.00015546 & 2.0947 \cdot 10^{-19} & 1.0872 \cdot 10^{-19} & -5.5887 \cdot 10^{-22} & 2.0852 \cdot 10^{-6} & -7.0393 \cdot 10^{-7} \\ -3.8355 \cdot 10^{-5} & -1.4972 \cdot 10^{-20} & -2.2394 \cdot 10^{-19} & 4.1259 \cdot 10^{-22} & -7.0393 \cdot 10^{-7} & 6.833 \cdot 10^{-7} \end{bmatrix}$$

(a)

$$\begin{bmatrix} 0.0367844 & 2.64194 \cdot 10^{-5} & -0.000119706 & 8.33421 \cdot 10^{-8} & -0.000155759 & -3.84876 \cdot 10^{-5} \\ 2.64194 \cdot 10^{-5} & 0.010323 & -0.000588892 & -8.27622 \cdot 10^{-5} & -2.21609 \cdot 10^{-7} & 3.46285 \cdot 10^{-8} \\ -0.000119706 & -0.000588892 & 0.442735 & -0.000135955 & 9.13904 \cdot 10^{-7} & -2.02893 \cdot 10^{-7} \\ 8.33421 \cdot 10^{-8} & -8.27622 \cdot 10^{-5} & -0.000135955 & 1.91456 \cdot 10^{-6} & -1.52874 \cdot 10^{-9} & 2.6768 \cdot 10^{-10} \\ -0.000155759 & -2.21609 \cdot 10^{-7} & 9.13904 \cdot 10^{-7} & -1.52874 \cdot 10^{-9} & 2.08748 \cdot 10^{-6} & -7.03794 \cdot 10^{-7} \\ -3.84876 \cdot 10^{-5} & 3.46285 \cdot 10^{-8} & -2.02893 \cdot 10^{-7} & 2.6768 \cdot 10^{-10} & -7.03794 \cdot 10^{-7} & 6.83906 \cdot 10^{-7} \end{bmatrix}$$

(b)

Table 1: Covariance matrices produced by backward propagation (a) and Monte Carlo simulation (b) for scenario 1 at the PME.

$$\begin{bmatrix} 0.036735 & 0 & 0 \\ 0 & 0.010308 & -0.00058889 \\ 0 & -0.00058889 & 0.44315 \end{bmatrix}$$

(a)

$$\begin{bmatrix} 0.0067017 & 0 & 0 \\ 0 & 0.0067294 & -0.0063575 \\ 0 & -0.0063575 & 0.61402 \end{bmatrix}$$

(b)

Table 2: Translational part of the covariance matrices produced by backward propagation at the PME (a) and at the geometric centroid (b).

## 5.5 PME Computation

To investigate whether the PME as computed by equation 17 really has minimum error, we use the MCS to sample the volume at the computed PME and at several points in a close vicinity. The results show that, within the precision of our MCS, the trace of the resulting covariance matrices at the computed PME location actually is smaller in all cases.

Table 2 shows the translational part of the covariance matrices produced by backward propagation at the PME (a) and at the geometric centroid (b). When comparing the two, we see that the x and y components actually have higher error at the PME compared to the GC. However, the total error (trace) is smaller at the PME, due to a much reduced z component.

## 6 PME LOCATION

In the last part of this paper, we more closely investigate the location of the PME in the various scenarios and describe the observations.

### 6.1 Comparison PME and GC location

Table 3 shows the error at PME and GC, as well as the location of the PME in the scenarios described in section 5.2. In order to be independent from the actual coordinate systems, the error is given as the square root of the three eigenvalues of the translational covariance. As we can see from the table, the PME not only minimizes the trace norm, but also the largest eigenvalue, which corresponds to the error along the “worst” axis. Another interesting observation is that the smaller eigenvalues become larger at the PME in some cases. This observation suggests that the PME somehow distributes the error among the axes (from z to x and y in the one-camera case) while still minimizing the overall error.

The addition of a second camera (scenarios 3 and 4) or a fifth non-planar fiducial (scenarios 2 and 4) seems to reduce the offset between PME and GC, but do not completely eliminate the problem. In the “real” scenario 5, no big difference can be seen between the PME and the GC when all fiducials are visible to all cameras. When half of the target is occluded (scenario 6), the PME strongly deviates from the GC, as expected. In the case where each fiducial is observed by only a single camera (scenario 7), but observations are distributed equally among the cameras, a slight difference between PME and GC locations can be observed, but not enough to change the estimated error significantly.

### 6.2 Scenario 1

As scenario 1 shows the most extreme difference between GC and PME locations, we will more closely investigate this situation. The scenario consists of a single camera that observes four fiducials, arranged in a square. Therefore, the observations made here can eventually also be applied to the AR toolkit-like square marker trackers frequently found in simple AR applications.

In this investigation, we first systematically shifted the CRE through a volume around the target and computed for each location the trace norm of the translational covariance, computed by BPC. Figure 3 shows a volume rendering of the result. The slice lying in the target’s plane is color-coded according to the error at each location. The gray ellipsoid describes an iso-surface of equal error at the center of which the PME is located. As already suggested by table 3, we can clearly see that in this scenario, the PME significantly moves out of the target plane and deviates from the centroid of the four fiducials.

In a second step, we have a closer look at how the PME’s location depends on the orientation of the planar target towards the camera. Therefore, we systematically rotate the target around its x-axis, which is parallel to the camera’s x-axis. The path of the PME is shown in figure 4. We observe that when the target plane is parallel to the image plane, the PME lies slightly (5mm) behind the target. As the target is rotated, the PME moves away quickly, reaching its furthest position at an angle of 45°. At 90°, the PME lies in the target plane, about 7mm closer to the camera than the GC.

The corresponding RMS errors at the GC and the PME are shown in figure 5c. We see that the error at the GC looks like a straight line with dents at 0° and 180°, whereas the error at the PME has more of a sine-like shape. At 0° and 90°, where both points have the closest location, they also have closest RMS error. The largest difference is observed at an angle of about 19°, where the GC error is about 25% larger than the PME error.

	error PME			error GC			PME $\rightarrow$ GC offset (mm)		
	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_1$	$\sigma_2$	$\sigma_3$	x	y	z
Scenario 1	0.1015	0.1917	0.6657	0.0816	0.0819	0.7836	0.0	98.5	-68.1
Scenario 2	0.0748	0.1220	0.6191	0.0729	0.0730	0.6509	0.0	57.2	-48.2
Scenario 3	0.0586	0.0640	0.2425	0.0577	0.0579	0.2484	0.0	19.6	-13.0
Scenario 4	0.0516	0.0537	0.2184	0.0515	0.0516	0.2201	0.0	10.0	-10.5
Scenario 5	0.0621	0.0648	0.1470	0.0622	0.0648	0.1470	0.7	1.9	1.0
Scenario 6	0.0934	0.0980	0.2397	0.1450	0.2430	0.3303	-107.9	0.7	0.2
Scenario 7	0.0875	0.0920	0.2097	0.0887	0.0921	0.2097	4.8	4.5	-1.2

Table 3: Comparison of PME and GC error in all scenarios. The first six columns contain the square roots of the eigenvalues of the translational covariance for the PME and GC locations. The last three columns describe the PME location relative to the GC.

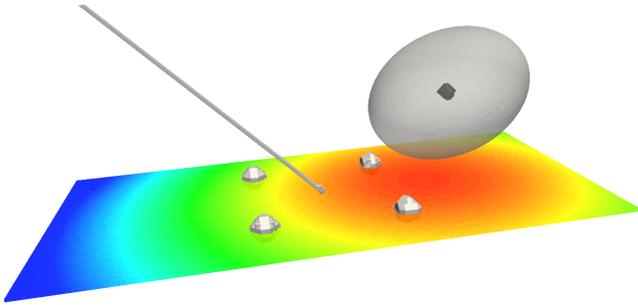


Figure 3: Volume rendering of the error for scenario 1. The color-coded slice goes through the plane defined by the target's fiducials. The line from the target centroid points towards the camera. The gray ellipsoid describes an iso-surface of equal error at the center of which the PME is located.

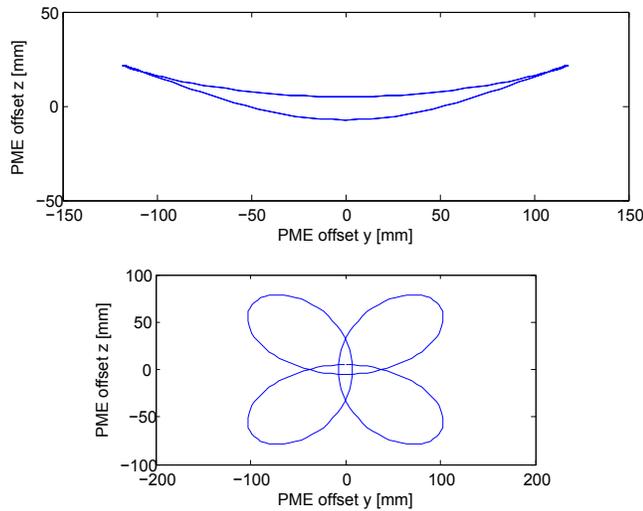


Figure 4: Traces of the PME location as the target is rotated around its x axis. The location is shown in both world (top) and target coordinates (bottom).

## 7 EXPERIMENTAL VALIDATION

In order to verify that the described effects also appear in a real system, we conducted an experiment with our ARTrack TP system, using two FireWire cameras. A planar target, consisting of four retro-reflective markers was placed in front of the two cameras at an angle of about  $45^\circ$ , and a long sequence of measurements (70,000 frames) was recorded without moving the target.

As Monte Carlo simulation had revealed the strongest effects in one-camera setups, we removed all but the first frame of the second camera. This allowed for stereo detection of the target in the first frame, with continuing tracking during the rest of the sequence.

Under normal circumstances, Gaussian sensor noise is negligible. We therefore artificially increased sensor gain to the maximum level, yielding an unrealistically high 2D feature noise with a standard deviation of about  $\frac{1}{10}$  pixel.

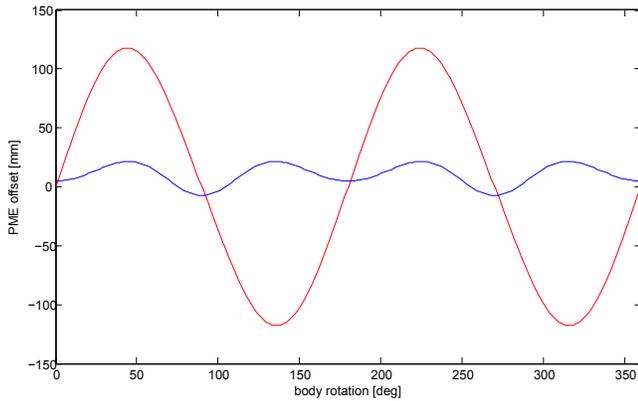
For each of the 70,000 recorded frames, a 6DOF pose was estimated from the four detected marker locations. In order to visually determine the PME location, we systematically sampled a volume around the target. For each point in the volume, its target coordinates were multiplied with all 70,000 6DOF poses and the resulting jitter in world coordinates was determined using the trace of the covariance norm of the resulting 3D coordinates.

Figure 6 shows a volume rendering of the resulting data set, including the location of the computed PME. The result undermines our claim that the PME can be far from the GC. As can be seen, there is some offset between the computed and the measured PME locations. In order to explain where this offset comes from, further investigations are necessary. One possible explanation could be static errors in either the camera calibration or the assumed target geometry.

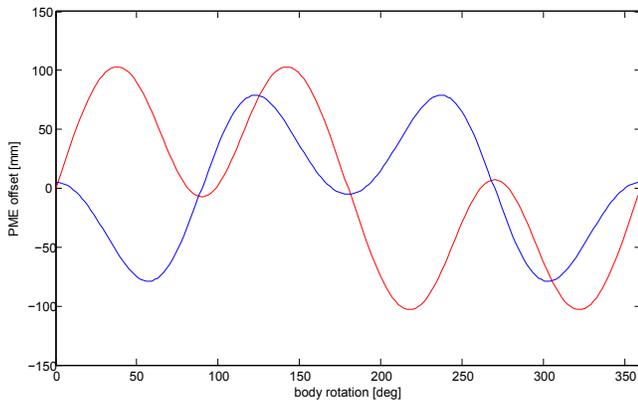
## 8 CONCLUSION

In this paper we have investigated how the choice of a center of rotational error (CRE) influences the result of error propagation techniques for optical tracking systems. We found out that, in order to properly characterize the error of a tracking system, the location of the CRE does not matter, as the error can always be propagated to any other point without loss of information. This is due to the usually unintuitive correlations between rotation and translation.

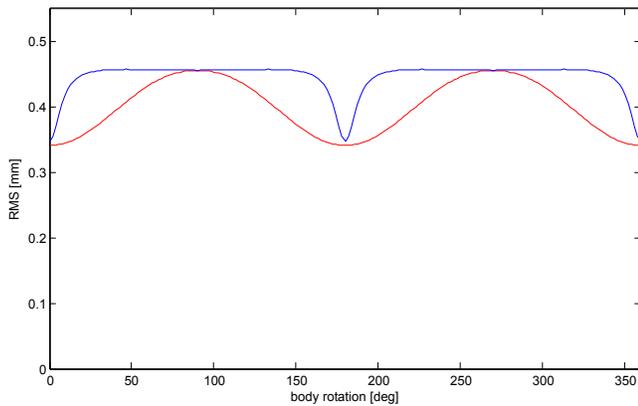
However, when it comes to the interpretation of the diagonal entries of the covariance matrix in terms of rotational and translational error, the choice of a CRE matters. Only in the point of minimum error (PME), the trace of the covariance matrix is minimal and rotational and translational errors are separated as much as possible. Contrary to common belief, the PME in general does not coincide with the geometric centroid (GC) of a tracking target. We showed, that the pose covariance matrix  $\Sigma$  contains enough information to compute the PME from  $\Sigma$  alone and provided the necessary equations. The approach was verified using Monte Carlo Simulation on a number of camera setups. We found out, that in some cases, the PME can be surprisingly far away from the GC, which was also validated in a real system experiment.



(a)



(b)



(c)

Figure 5: PME location (red:y, blue:z) plotted against the target's rotation angle around the x axis. Locations are shown in world (a) and target coordinates (b). (c) shows the RMS translational error estimated at the GC (blue) and at the PME (red) depending on the orientation.

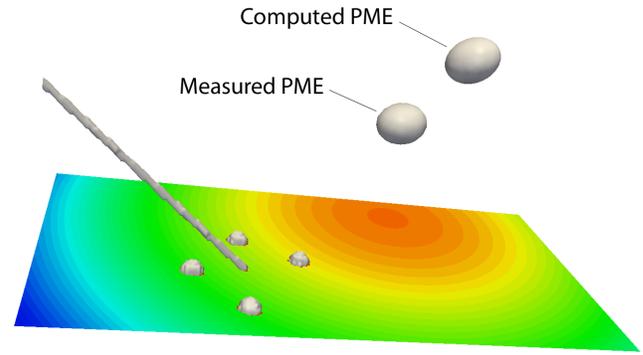


Figure 6: Results of the real system experiment. There is a deviation between computed and measured PME locations that needs further investigations. The effect that the PME is not identical to the GC, however, is clearly visible.

When reporting covariance matrices to the user, we suggest that tracking systems use an error representation with variable CRE, such as the one presented in section 2. Unless the actual point of interest (such as the tip of a tool) is known to the system, this CRE should coincide with the PME. This has the advantage that rotational and translational errors are separated as much as possible and, in consequence, users can interpret the diagonal entries of the covariance matrix as pure rotation and translation. Furthermore, explicit CRE representations have the advantage that this separation can be maintained under a number of coordinate system transformations, by simply adjusting the CRE location (see appendix).

Unfortunately, the current approach requires the covariance matrix to be computed for an arbitrary CRE, before the PME can be computed. After this, the error can be propagated to the PME. It probably would be more efficient to have an approach where the PME location can be computed beforehand. This, however, is considered future work.

This paper has to be seen as only the first step in investigating the role of the CRE in 6DoF error representations for optical tracking systems. While we have theoretically and experimentally validated that the PME in general does not coincide with the GC, more insight needs to be gained why this is the case and what the exact relationship is between rotational and translational errors. We hope that this paper will serve as a good starting point for future research in this direction.

## ACKNOWLEDGEMENTS

This work was supported by the Bundesministerium für Bildung und Forschung, project AVILUS - Angewandte Virtuelle Technologien im Produkt- und Produktionsmittellebenszyklus (FKZ: 01IM08001C).

## REFERENCES

- [1] B. D. Allen and G. Welch. A general method for comparing the expected performance of tracking and motion capture systems. In *VRST '05: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 201–210, New York, NY, USA, 2005. ACM Press.
- [2] M. Bauer, M. Schlegel, D. Pustka, N. Navab, and G. Klinker. Predicting and estimating the accuracy of vision-based optical tracking systems. In *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR '06)*, Santa Barbara (CA), USA, October 2006.
- [3] D. C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37(8):855–866, 1971.
- [4] E. M. Coelho, B. MacIntyre, and S. Julier. OSGAR: A scenegraph with uncertain transformations. In *Proc. IEEE International Sym-*

posium on Mixed and Augmented Reality (ISMAR'04), pages 6–15, Washington, DC, 2004. IEEE.

- [5] L. Davis, E. Clarkson, and J. P. Rolland. Predicting accuracy in pose estimation for marker-based tracking. In *Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'03)*, page 28, 2003.
- [6] L. Davis, F. G. Hamza-Lup, and J. P. Rolland. A method for designing marker-based tracking probes. In *ISMAR '04: Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 120–129, Washington, DC, USA, 2004. IEEE Computer Society.
- [7] J. M. Fitzpatrick, J. B. West, and C. R. M. Jr. Predicting error in rigid-body, point-based registration. *IEEE Transactions on Medical Imaging*, 17(5):694–702, 1998.
- [8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, June 2000.
- [9] H. Hastedt. Monte-carlo-simulation in close-range photogrammetry. In *ISPRS Symposium Commission V, Istanbul, Turkey, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science*, volume 35, pages 18–23, 2004.
- [10] W. Hoff and T. Vincent. Analysis of head pose accuracy in augmented reality. In *IEEE Transactions on Visualization and Computer Graphics*, volume 6(4), pages 319–334. IEEE Computer Society, 2000.
- [11] R. L. Holloway. *Registration errors in augmented reality systems*. PhD thesis, Chapel Hill, NC, USA, 1995.
- [12] S. Julier and J. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, March 2004.
- [13] K. Kraus. *Photogrammetry, Geometry from Images and Laser Scans*. de Gruyter Textbook, 2nd edition, 2007.
- [14] W. Niemeier. *Ausgleichsrechnung, Statistische Auswertemethoden*. de Gruyter Textbook, 2nd edition, 2008.
- [15] D. Pustka, M. Huber, M. Bauer, and G. Klinker. Spatial relationship patterns: Elements of reusable tracking and calibration systems. In *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR'06)*, October 2006.
- [16] C. Robert and G. Casella. *Monte Carlo statistical methods*. Springer Verlag, 2004.
- [17] T. Sielhorst, M. Bauer, O. Wenisch, G. Klinker, and N. Navab. Online estimation of the target registration error for -ocular optical tracking systems. In *MICCAI (2)*, pages 652–659, 2007.
- [18] H. J. Woltring, R. Huiskes, A. de Lange, and F. E. Veldpaus. Finite centroid and helical axis estimation from noisy landmark measurements in the study of human joint kinematics. *Journal of Biomechanics*, 18(5):379–389, 1985.

## A TRANSFORMATION RULES FOR EXPLICIT CRE ERROR REPRESENTATIONS

Mathematically, an arbitrary center of rotational error can be chosen for 6DOF pose error description, and the resulting  $6 \times 6$  covariance matrix can always be propagated to any point of interest without loss of information. However, when no PoI is known to the tracking system, it might make sense to choose the PME as the CRE, because the PME provides the best separation between translational and rotational errors and gives the most intuitive results when users just look at the diagonal entries of the covariance matrix without further propagation.

Since the PME depends on the current set of observations, the error description needs to explicitly take a dynamic CRE into account. For this purpose, we propose to use the error representation defined in equation 1 and the tuple  $(c, \Sigma)$  as defined by equation 4 as an error representation at runtime.

Error representations without explicit CRE location have problems when the coordinate systems of the underlying transformation are changed. Consider the case where the target coordinate system is moved from the centroid to some more convenient location. When the error is implicitly centered around the new target coordinate frame, the covariance also needs to be updated and rotational error is added to the translational error. While the underlying error still is the same, the translational error that can be read from the

upper-left block of the covariance matrix now seems to be much higher. By having an explicit CRE, we can just update the CRE location to the new coordinate systems without factoring rotational into translational error.

In this appendix, we describe the effect that certain transformations have on the proposed error representation. We also give equations for practical computation of 3D covariances in arbitrary points of interest.

**Spatial Relationship Patterns** In order to more clearly define the meaning of the different transformations, each of them is illustrated by a spatial relationship pattern [15]. In these diagrams, nodes represent coordinate systems or orientation-free points, whereas edges represent measurements such as position or pose. Input measurements are drawn as solid lines whereas output edges are drawn as dashed lines.

Each edge is attributed with the type of measurement. An edge  $A \rightarrow B$  of type  $6DOF$  represents the following transformation of a point  $x_B$  in  $B$  to a point  $x_A$  in  $A$ :

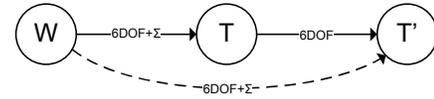
$$x_A = R x_B + t$$

For edges of type  $3DPos$ , rotation is missing and the position of  $B$  in  $A$  is given.

Measurements with associated uncertainty information are denoted by adding  $+\Sigma$  to their type, e.g.  $6DOF+\Sigma$ . For  $6DOF$  edges, uncertainty is represented by the tuple  $(\Sigma, c)$  as described above. For  $3DPos$  types, only a  $3 \times 3$  covariance matrix is given.

Edges in the pattern may also describe additional constraints that must be fulfilled between two objects. In the style of UML notation, the type definitions are surrounded by double angle brackets ( $\ll \dots \gg$ ).

### A.1 Change of Target Coordinate System

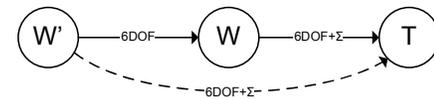


Let  $(R, t)$  be a transformation from target to world coordinates, as described above. In homogeneous matrix notation, the target coordinate system can be changed by a right-hand-side transformation  $(R_R, t_R)$  using the following matrix product:

$$\begin{bmatrix} R' & t' \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_R & t_R \\ 0 & 1 \end{bmatrix} \quad (18)$$

Assuming that  $(R_R, t_R)$  is an error-free transformation, the error description  $(\Sigma, c)$  of  $(R, t)$  is equally valid for  $(R', t')$ , i.e.  $(\Sigma', c') = (\Sigma, c)$ .

### A.2 Change of World Coordinate System



Let  $(R, t)$  be a transformation from target to world coordinates, as described above. In homogeneous matrix notation, the world coordinate system can be changed by a left-hand-side transformation  $(R_L, t_L)$  using the following matrix product:

$$\begin{bmatrix} R' & t' \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_L & t_L \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \quad (19)$$

Assuming that  $(R_L, t_L)$  is an error-free transformation, the new error description  $(\Sigma', c')$  for  $(R', t')$  is computed using the following equations:

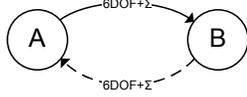
$$\Sigma' = A_L \Sigma A_L^T \quad (20)$$

$$c' = R_L c + t_L \quad (21)$$

where  $A_L$  is a  $6 \times 6$  matrix defined as follows:

$$A_L = \begin{bmatrix} R_L & 0 \\ 0 & R_L \end{bmatrix} \quad (22)$$

### A.3 Pose Inversion



Exchanging the world and target coordinate systems corresponds to a matrix inversion:

$$\begin{bmatrix} R' & t' \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} R^{-1} & -R^{-1}t \\ 0 & 1 \end{bmatrix} \quad (23)$$

Transforming both covariance matrix and centroid from world into the target coordinate frame yields the new error description  $(\Sigma', c')$ :

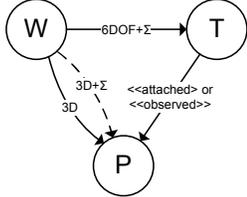
$$\Sigma' = A_I \Sigma A_I^T \quad (24)$$

$$c' = R^{-1}(c - t) \quad (25)$$

where  $A_I$  is a  $6 \times 6$  matrix defined as follows:

$$A_I = \begin{bmatrix} R^{-1} & 0 \\ 0 & R^{-1} \end{bmatrix} \quad (26)$$

### A.4 Transformation of a Point of Interest



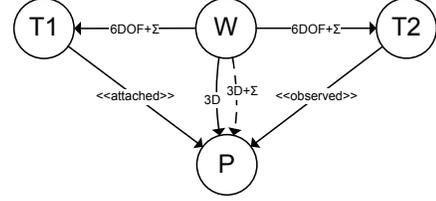
Let  $p$  be the world coordinate location of a point of interest that is rigidly attached to the target. The uncertainty in the location of  $p$  caused by the measurement uncertainty of the target is computed using the following equation:

$$\Sigma_p = \begin{bmatrix} I_{3 \times 3} & [p - c]_{\times} \end{bmatrix} \Sigma \begin{bmatrix} I_{3 \times 3} & [p - c]_{\times} \end{bmatrix}^T \quad (27)$$

where  $\Sigma_p$  is a  $3 \times 3$  covariance matrix that describes the location uncertainty of  $p$  in world coordinates and  $[p - c]_{\times}$  is a  $3 \times 3$  anti-symmetric matrix as defined earlier.

Note that the same equation holds when  $p$  is a point in the world that is observed by a camera attached to the target. In this case, the resulting covariance describes the uncertainty in the observation of  $p$  (in world coordinates).

### A.5 Combination of two Pose Measurements



A slightly more complicated case typically appears in AR applications: a point of interest is rigidly attached to one target and observed by another. For instance, a doctor may look through a tracked HMD onto a PoI that lies on a patient with an attached tracking target. In this case we are interested in the combined error that results from tracking errors in both targets, propagated to the point of interest.

To compute the resulting position of the PoI relative to the HMD (T2), one normally would first invert the transformation  $W \rightarrow T2$  multiply the result with  $W \rightarrow T1$  and then use the result to transform the PoI into the coordinate system T2. Unfortunately, we cannot easily apply our error representation onto the multiplication of two uncertain poses, as the product would no longer have well-defined center of rotational error.

However, given a single point of interest  $p$ , whose coordinates relative to T1 and T2 are given by  $p^{T1}$  and  $p^{T2}$  respectively, we can express the observation process by the following equation:

$$p^{T2} = \tilde{E}_{R2} (\tilde{R}_2 (E_{R1} (R_1 p^{T1} + t_1 - c_1) + c_1 + e_{t1}) + \tilde{t}_2 - \tilde{c}_2) + \tilde{c}_2 + \tilde{e}_{t2} \quad (28)$$

where  $\tilde{R}_2$ ,  $\tilde{t}_2$ ,  $\tilde{E}_{R2}$ ,  $\tilde{e}_{t2}$  and  $c_2$  are the parameters of the inverted transformation  $W \rightarrow T2$  as defined in section A.3. Applying the forward propagation of covariance and transforming the result back to the world coordinate system, we obtain the following simple equation:

$$\Sigma'_p = \Sigma_p^{T1} + \Sigma_p^{T2} \quad (29)$$

where  $\Sigma_p^{T1}$  and  $\Sigma_p^{T2}$  are the  $3 \times 3$  covariance matrices resulting from the PoI transformation as given by equation 27, using the parameters of  $W \rightarrow T1$  and  $W \rightarrow T2$  respectively.

In short, the combined error at a point of interest is computed by summing up the covariance matrices resulting from the individual PoI propagations.