Projective Rotations Applied to a Pan-tilt Stereo Head
Andreas Ruf, Radu Horaud

To cite this version:
Abstract

A non-metric pan-tilt stereo-head consists of a weakly calibrated stereo rig mounted on a pan-tilt mechanism. It is called non-metric since neither the kinematics of the mechanism, nor camera calibration are required. The Lie group of “projective rotations” - homographies of projective space corresponding to pure rotations – is an original formalism to model the geometry of such a pan-tilt system. A Rodrigues alike formula as well as a minimal parameterization of projective rotations are introduced.

Based on this, the practical part devises a numerical optimization technique for accurately estimating projective rotations from point correspondences, only. This procedure recovers sufficient geometry to operate the system. The experiments validate and evaluate the proposed approach on real image data. They show the weak calibration, image prediction, and homing of a non-metric pan-tilt head.

1 Introduction

One of the most useful sensors in computer vision is a pan and tilt stereo head. Such a sensor is composed of a camera pair mounted on a pan and tilt mechanism. Such an active sensor has been proved to be useful for a number of applications including the tracking of moving objects, 3-D reconstruction, visual servoing, and so forth. For all these applications, controllability is a crucial issue: the stereo head must move according according to what it has previously seen. Therefore controllability consists of closing the loop between camera observations and actions onto the pan and tilt motors. Such a closed loop implies a model of the active stereo head including a purely geometric model of the camera pair configuration as well as direct and inverse kinematic models of the mechanism on which the cameras are mounted.

Whereas projective models exist for the camera geometry, pan-tilt mechanisms are classically modeled using Euclidean transformations. Therefore, the projective information recovered with the weakly calibrated stereo camera pair must be converted into Euclidean information.

In this paper we introduce a new approach and a new mathematical framework to represent both the camera-pair geometry and the pan-tilt mechanism in projective space.

The concept of projective displacement (a projective transformation of a rigid body) is first introduced and we show how to represent the pan and tilt stereo head within such a representation (sections 2 and 3). The particular case of projective rotations corresponding to pure rotational motions is considered in section 4. We propose a parameterization of projective rotations which is the projective equivalent of the well known Rodrigues formula. We describe some interesting properties of this parameterization and we reveal the Lie group structure of projective rotations. Section 5 addresses the problem of how to estimate projective rotations in practice and we introduce both a linear and a non-linear method for estimating such a transformation from images observations. Section 6 describes several experiments including the complete ”weak calibration” of a pan-tilt stereo head, feed-forward prediction useful for object tracking and a homing experiment. Finally section 7 contains a summary, some conclusions and some directions of future work.

1.1 Notations

Boldface types $\mathbf{H}, \mathbf{T}$ are used for matrices, bold italic types $\mathbf{M}, \mathbf{k}$ for vectors, calligraphic types $\mathcal{F}, \mathcal{P}$ for frames, and Roman types $a, b, \theta$ for scalars, angles. Vectors $\mathbf{k}$ are column vectors, and row vectors are written as a transpose $\mathbf{h}^T$. Corresponding coefficients $k_{12}, k_{13}$ are indexed as usual. Equality up to scale is denoted $\simeq$.

2 Preliminaries

A calibrated stereo rig is modeled as two pinhole cameras $\mathbf{K}, \mathbf{K'}$, that are rigidly linked by $(\mathbf{R'}, \mathbf{t'})$. A point $\mathbf{N}$
in Euclidean space projects onto the points \( m \) and \( m' \) in the left and right projective image plane [6]. The solution of the corresponding projection constraints (1) for \( N \) yields a Euclidean reconstruction relative to the Euclidean camera frame \( E \)

\[
\begin{align*}
m \simeq [K] \theta N, \\
m' \simeq [K'R'][K't'] N.
\end{align*}
\]  

(1)

A weakly calibrated stereo rig is modeled as a pair of cameras with known epipolar geometry \( F \) that is estimated from point correspondences only [8]. This allows two projection matrices to be calculated, such that the corresponding projection reconstruction solving (2)

\[
m \simeq PM, \\
m' \simeq P'M
\]  

(2)
is relative to a projective camera frame \( \mathcal{P} \), that is indeed defined by \( P, P' \) and thought to be associated with the stereo rig [7], [13].

If the stereo rig’s internal geometry is fix, i.e. for constant \( K, K' \), and \( (R', t') \), the so-called projective-Euclidean link \( H_{PE} \)

\[
N \simeq H_{PE}M, \\
H_{PE} \simeq \left[ K^{-1} \theta \begin{array}{c} 0 \\ a^T \\ 1 \end{array} \right].
\]  

(3)
is the well-defined homography, that upgrades the projection reconstruction (2) in \( \mathcal{P} \) to a Euclidean one (1). in \( E \). \( H_{PE} \) encapsulates the geometry of the stereo camera: the left intrinsic parameters \( K \) and the plane at infinity \( a^T \). Calculating \( H_{PE} \) amounts to calibrating the system [18], [5].

Suppose now, that the rig undergoes a rigid motion [10] described in \( E \) by \( N' = T_{RT}N \), and convince yourself that the corresponding homography in \( \mathcal{P} \) writes as

\[
H_{RT} = \gamma H_{PE}^{-1} T_{RT} H_{PE}.
\]  

(4)

Such homographies \( H_{RT} \), that are algebraically similar to displacement group \( SE(3) \), form also a group to be referred to as projective displacements [16].

Throughout this paper, we will use them, more precisely a class of subgroups: the projective rotations [3], to model the motions of the pan-tilt system relative to the projective frame \( \mathcal{P} \).

3 The non-metric pan-tilt stereo head

This mechanism is described as a RR-chain consisting of a pan joint with angle \( \psi \) that is serially linked to a tilt joint with angle \( \phi \). Furthermore, a zero-position of this chain is defined by \( \psi_0, \phi_0 \) and new joint-variables are reset accordingly \( (\Psi, \Phi) = (\psi - \psi_0, \phi - \phi_0) \). Two weakly calibrated stereo camera are rigidly mounted onto the end-effector (fig. 1).

Consider the standard camera frame \( E \) and a corresponding projective frame \( \mathcal{P} \), both rigidly moving with the cameras. The Euclidean and projective coordinates of a static point \( M \) in these frames are functions of the joint angles: \( N(\Psi, \Phi) \) and \( M(\Psi, \Phi) \). In the zero-position, an initial frame \( \mathcal{E}_0 \) and correspondingly \( \mathcal{P}_0 \) are defined to be coincident with them. In these frames, the world-coordinates of \( M \) are defined to be \( N(0, 0) \) and \( M(0, 0) \). The points \( N \) and \( M \) are related by (3). Rather than attaching moving frames to each link as does the Denavit-Hartenberg method [15], the so-called zero-reference method [14] models a mechanism’s geometry relative to the inertial frame by means of displacements \( T_{\Psi} \) and \( T_{\Phi} \), which represent revolutions of the pan and tilt joint. Most importantly, the joints have to be in zero-position relative to the moving frame \( E \) (fig. 1).

\[
N(\Psi, 0) = T_{\Psi} \cdot N(0, 0), \\
N(\Psi, \Phi) = T_{\Phi} \cdot N(\Psi, 0).
\]  

(5)

(6)

This is also reflected by the order in which the revolutions are multiplied in the forward kinematic map:

\[
N(\Psi, \Phi) = T_{\Phi}T_{\Psi}N(0, 0).
\]  

(7)

If we switch now from the Euclidean to the projective coordinates by substituting (3) into (7), we obtain the homography \( H(\Psi, \Phi) \) from \( \mathcal{P}_0 \) to \( \mathcal{P} \) that corresponds to the motion

\[
\begin{align*}
R_{\Phi} & = H_{PE} \left[ \begin{array}{c} 0 \\ a^T \\ 1 \end{array} \right], \\
H_{\Phi} & = R_{\Phi} H_{PE}^{-1}.
\end{align*}
\]  

(8)

(9)
of the pan-tilt system.

\[ H_{PE}(\Psi, \Phi) \simeq T_\Phi T_\Psi H_{PE} M(0, 0) \]

\[ M(\Psi, \Phi) \simeq \frac{H_{PE}^{-1} T_\Psi H_{PE}^{-1} T_\Phi H_{PE} M(0, 0)}{H_\Phi H_\Psi} \] (8)

As soon as we manage to write equation (8) without \( H_{PE} \), we have obtained a purely projective model of the system’s geometry, of both camera and mechanism, that is free of Euclidean frames and respective metric measures. In this sense, we will call it a non-metric pan-tilt stereohead.

Its non-metric forward kinematics follows directly from (8) and writes as

\[ H(\Psi, \Phi) = H_\Phi H_\Psi = \exp(\Phi \hat{H}_\Phi) \exp(\Psi \hat{H}_\Psi), \] (9)

a product of two homographies \( H_\Phi \) and \( H_\Psi \) (8). They are those elements of a class of projective rotations \( H_R \) (see section 4) that model pan and tilt revolutions for the given system. In particular the exponential representation (18) allows us to calculate the forward kinematics as a product of exponentials (9), [1], [17], and thus as an explicit function of the joint angles \( (\Psi, \Phi) \). This assumes the tangent operators \( \hat{H}_\Phi \), \( \hat{H}_\Psi \) of pan and tilt motions to be known, which amounts to a weak calibration of the mechanism (section 6.1). They can calculated from a single pan and tilt motion (29) or very accurately using non-linear optimization (section 5).

4 Theory of projective rotations

In the sequel, we will show that under assumptions laid out in section 2, a class of subgroups of projective displacements is well-defined which models the projective kinematics of revolute joints. Moreover, this subgroups will turn out to be a Lie-group cite [16].

4.1 Definition

A projective rotation (p-rotation) is a homography \( H_R' \) acting on \( P \), which corresponds to the pure rotation \( T_R \) acting on \( E \). Both transforms are algebraically similar by the PE-link (3) and scale \( \gamma \)

\[ H_R' = \gamma H_{PE}^{-1} T_R H_{PE}. \] (10)

A normalized projective rotation (np-rotation) \( H_R = H_R'/\gamma \) has \( \text{det} H_R = 1 \) after division by

\[ \gamma = (\text{det} H_R')^{1/4} \text{sign}(\text{trace} H_R') \] (11)

The angle of a np-rotations obeys as for \( T_R \)

\[ \cos(\theta) = 1/2(\text{trace} H_R - 2). \] (12)

4.2 Jordan Normal Form

A pure rotation has as eigenvalues the complex-conjugate pair \( \lambda_0 = \cos \theta \pm i \sin \theta \) and a double \( \lambda_1 = 1 \). In consequence, the real Jordan matrix has a rotation block \( J_2(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \) and a Jordan-Block \( J_2(1) \) of order two [11].

The Jordan matrix \( J_R \) of a np-rotation hence is

\[ J_R = \begin{bmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \] (13)

\[ H_R = H_J^{-1} J_R H_J. \] (14)

It is canonical up permutations of the blocks. Here, they are arranged like a rigid displacement, i.e. a rotation by \( \theta \) around the z-axis. The decomposition (14) itself is ambiguous in \( H_J \). This ambiguity is characterized by the commutator \( C_R \) of \( J_R \):

\[ C_R = \begin{bmatrix} a & -b & 0 & 0 \\ b & a & 0 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & d & c \end{bmatrix}, \] (15)

The complete family of Jordan decompositions thus is

\[ H_R = H_J^{-1} C_R^{-1} J_R C_R H_J. \] (16)

4.3 Rodrigues’ Parameterization

Proposition 1 A projective rotation is a projective transformation that is conjugated to a pure rotation.

a) The 4×4 homography describing this transformation can be parameterized as:

\[ H_R = I + \sin \theta \hat{H}_R + (1 - \cos \theta) \hat{H}_R^2 \] (17)

This will be referred to as the “Rodrigues formula of projective rotations”.

b) It equals the solution of the matrix exponential

\[ H_R = \exp \left( \theta \hat{H}_R \right) \] (18)

This will be referred to as “exponential form of projective rotations”.

c) This parameterization is unique.

Proof: Eq. (14) writes:

\[ H_R = H_J^{-1} (J_R - I) H_J + I \]

Denote the rows of \( H_J \) by \( h_i^T \) and the columns of \( H_J^{-1} \) as \( k_i \). Remember also that

\[ h_i^T k_j^T = \delta_{ij} \] (19)
Then

\[
\begin{align*}
\mathbf{H}_R &= \begin{bmatrix} k_1 & k_2 \end{bmatrix} \begin{bmatrix} \cos \theta - 1 - \sin \theta & k_1^T \\ \sin \theta & \cos \theta - 1 \end{bmatrix} + \mathbf{I} \\
&= \mathbf{I} + \sin \theta \begin{bmatrix} k_2 h_1^T - k_1 h_2^T \\ * \end{bmatrix} + \ldots
\end{align*}
\]

(19)

Introduce the rank 2 matrix \( \mathbf{\hat{H}}_R \) and use (19) to show

\[
\begin{align*}
\mathbf{\hat{H}}_R &= k_2 h_1^T - k_1 h_2^T \\
\mathbf{\hat{H}}_R^2 &= k_1 h_1^T + k_2 h_2^T,
\end{align*}
\]

(20)

and further by induction that

\[
\begin{align*}
\mathbf{\hat{H}}_R^{2n+2} &= (-1)^{n+1} \mathbf{\hat{H}}_R^2 \\
\mathbf{\hat{H}}_R^{2n+1} &= (-1)^n \mathbf{\hat{H}}_R.
\end{align*}
\]

Substituting (21), (22) into (20) proves the part a).

Part b) is shown by collecting the odd and even powers in the exponential of \( \mathbf{\hat{H}}_R \)

\[
\exp(\theta \mathbf{\hat{H}}_R) = \mathbf{I} + \sum_{n=0}^{\infty} \frac{\theta^{2n+2} \mathbf{\hat{H}}_R^{2n+2}}{(2n+2)!} + \sum_{n=0}^{\infty} \frac{\theta^{2n+1} \mathbf{\hat{H}}_R^{2n+1}}{(2n+1)!},
\]

then substituting (23), (24), and the power series of \( \sin \) and \( \cos \).

Uniqueness is indeed an issue, since the square \( \mathbf{C}_A \) in the ambiguity (15), (16) propagates through (26) also into \( \mathbf{\hat{H}}_R \) (27). The commutativity of the two matrices

\[
\mathbf{C}_A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad \mathbf{J}_A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}
\]

(25)

however proves that \( \mathbf{\hat{H}}_R \) is unambiguous, part c).

\[
\begin{align*}
\mathbf{\hat{H}}_R &= \begin{bmatrix} k_1 & k_2 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix} \\
&= \begin{bmatrix} k_1 & k_2 \end{bmatrix} \mathbf{C}_A^{-1} \mathbf{J}_A \mathbf{C}_A \begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix}.
\end{align*}
\]

(26)

4.4 Lie group property

Consider the one-parameter group of np-rotations \( \mathbf{H}_R \) with a fix \( \mathbf{H}_f \) (14). It is a smooth, one-dimensional manifold \( \theta \rightarrow \mathbf{H}_R(\theta) \) in \( \mathbb{R}^{4 \times 4} \). Its tangent operator is an infinitesimally small group element in the vicinity of the identity \( \lim_{\theta \rightarrow 0} \frac{\mathbf{H}_R(\theta) - \mathbf{I}}{\theta} \):

\[
\lim_{\theta \rightarrow 0} \frac{(1 - \cos \theta) \mathbf{H}_R^2 + \sin \theta \mathbf{H}_R}{\theta} = \mathbf{H}_R.
\]

(28)

Note that since the scale ambiguity is removed, this limit is well defined (11).

Hence, after fixing the family \( \mathbf{H}_f \), it is straight-forward to see that the np-rotations \( \mathbf{H}_R \) form a one-parameter Lie group that is isomorphic to \( \text{SO}(2) \) and that has \( \mathbf{\hat{H}}_R \) as a corresponding one-dimensional, linear Lie algebra [17], [2].

The exponential (18) is the mapping from the Lie algebra into the Lie group. An inverse, the map from group to algebra can be directly calculated from a given np-rotation using its inverse, like a rotation axis is calculated using the transpose.

\[
\mathbf{\hat{H}}_R = \frac{1}{2 \sin \theta} (\mathbf{H}_R - \mathbf{H}_R^{-1}),
\]

(29)

which is easily verified using (17). It can be thought of as the matrix logarithm of a np-rotation \( \log \mathbf{H}_R \).

5 Estimation of projective rotations

In this section, we show how to accurately estimate the tangent operator \( \mathbf{\hat{H}}_R \) of observed projective rotations \( \mathbf{H}_R \) from joint angle and image measurements, only. The goal is to weakly self-calibrate our mechanism from pan and tilt motions observed with the stereo head.

5.1 Objective function

Take for instance the projective pan rotations \( \mathbf{H}_R(\Psi_i) = \mathbf{H}_\Psi \) and \( n + 1 \) respective stereo-images \( \mathbf{m}_i \) and \( \mathbf{m}_i' \) taken for angles \( \Psi_i \). In practice, the stereo head is moved by operating the pan joint to \( \mathbf{\Psi}_i \) and interest points \( \mathbf{m}_i, \mathbf{m}_i' \) are matched between all the images (fig. 2). Let \( \mathbf{M}_0 \) be a projective reconstruction\(^1\) at \( \mathbf{\Psi}_0 \). It is transformed to \( \mathbf{M}_i = \mathbf{M}(\mathbf{\Psi}_i, 0) \) using (9) and backprojected using (2). Now, the total reprojection error of the sequence is minimized over a 10-parameter vector \( \mathbf{x} \), that is indeed an local minimal parameterization of \( \mathbf{\hat{H}}_R \) (section 5.2)

\[
\min_{\mathbf{x}} \sum_{i=1}^{n} d \left( \mathbf{m}_i, \mathbf{P} \exp(\mathbf{\Psi}_i \cdot \mathbf{\hat{H}}_\Psi(\mathbf{x})) \mathbf{M}_0 \right) + \ldots
\]

\[
+ d \left( \mathbf{m}_i', \mathbf{P}' \exp(\mathbf{\Psi}_i \cdot \mathbf{\hat{H}}_\Psi(\mathbf{x})) \mathbf{M}_0 \right).
\]

(30)

The error measure \( d \) is a suitably chosen image error, e.g. the Euclidean distance in pixel between extracted and backprojected point.

5.2 Minimal parameterization

The definition of a projective rotation (10) introduced 13 parameters, 5 intrinsic parameters in \( \mathbf{K} \), 3 in the plane at infinity \( \mathbf{a'} \), and 5 in the pure rotation \( \mathbf{T}_R \). It is known, that from a single rotation, the plane at infinity is not fully

\(^1\)The vectors \( \mathbf{m}_i, \mathbf{m}_i', \mathbf{M}_i \) generically stand for a whole set of matched points and corresponding reconstructions.
defined [3]. Moreover, just a subset of the internal parameters can be calibrated and the respective constraints are of non-trivial orders [4]. The “physical” parameters are hence not suitable for minimization, where the vectors in (26) will allows us to express $\hat{H}_R(x)$ depending a vector $x$ of 10 parameters.

Given an initial guess, a two-parameter family of solutions is introduced by the square $[a \ b \ a]$ (27). In order to resolve this ambiguity, $a$ and $b$ are fixed by setting $h_{11} = 1$, $h_{21} = 0$. In practice, this constraint is imposed by $QR$-decomposition $Q \cdot W$ with column-permutations $P$ to ensure non-zero, dominant elements on the diagonal and a good numerical condition.

$$\begin{bmatrix} h_1^T \\ h_2^T \\ \end{bmatrix} = Q \cdot W \cdot P, \ W = \begin{bmatrix} 1 & h_{12} & h_{10} \\ 0 & h_{22} & h_{24} \end{bmatrix} = \begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix}$$

Now, a well-conditioned form of (26) is $\hat{H}'_R = \hat{H}_R P^{-1}$

$$\hat{H}'_R = \begin{bmatrix} k_1 & k_2 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix}, \quad \begin{bmatrix} k_1 & k_2 \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} \end{bmatrix} Q.$$

(31)

With a good guess, just one initial $QR$-step is sufficient. Otherwise $h_{11}^1, h_{21}^1$ have to be monitored during the minimization.

After $h_{11}$ and $h_{21}$, further four parameters are eliminated by imposing the bilinear constraints (19), such that $\hat{H}_R(x)$ depends only on a vector $x$ of the 10 parameters $h_{12}^1, h_{13}^1, h_{14}^1, h_{22}^1, h_{23}^1, h_{24}^1, k_{13}^1, k_{14}^1, k_{23}^1, k_{24}^1$.

5.3 Initialization

It is straight-forward, that any Jordan decomposition (14) gives an initial guess. However, the tangent operator $\hat{H}_R$ calculated with (29) and rank two imposed by SVD usually yields a better guess. It has the polar decomposition

$$\hat{H}_R = \begin{bmatrix} f_{-i} \\ f_i \end{bmatrix} \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} e_{r}^T \\ e_{o}^T \end{bmatrix}.$$ 

The left and right eigenvectors $f_{-i}, f_i$ and $e_{r}^T, e_{o}^T$ to the eigenvalues $-i$ and $i$ can be recombined using the matrix $Y = \begin{bmatrix} 1 & -i \end{bmatrix}$ in order to obtain the same form as (21).

6 Experiments

A stereo head consisting of two CCD-cameras with 12.5mm lenses, baseline 35cm and vergence-angle 15° is mounted onto a robot end-effector. The two revolute axes of a 5-dof Carthesian robot by SINTERS represent the pan-tilt mechanism. As a repeatability of 0.01° is factory-given, the encoder readings are taken as ground-truth for joint angles ($\Psi, \Phi$), more precisely for their differences.

The real image data shows a small model of a mountain taken with tilt 20° and distance 2m with respect to the ground-plane. The extracted point-features are about a 200 of Harris’ interest-points. They are tracked or matched using SSD-based cross-correlations and refined to subpixel accuracy using paraboloid interpolation. Although an accuracy of 0.3px is theoretically predicted, errors sum-up during tracking such that noticeable drift of point features is observed, mainly because correlations are concatenated sequentially.

Figure 2: Calibration of pan(top) and tilt(bottom) joint: The images show the mountain model and superimposed point trajectories of pan and tilt motion, respectively. The missing steps are those used to calibrate. The + mark the backprojected points, the o mark the actual extracted points.

6.1 Non-metric calibration

In the calibration phase, two separate trajectories, the first stepping the pan-joint 20 times by 1°, and the second stepping the tilt-joint 10 times also by 1°, are driven. The corresponding image sequences show the mountain mov-
ing in the image horizontally \( u \) for pan and vertically \( v \) for tilt along a path of 300px length. A subset of steps and interest points was taken to establish a non-metric calibration of the system. This means estimating \( \hat{H}_\psi \) and \( \hat{H}_\phi \) (9) using our approach proposed in section (5). The result is evaluated against the the steps and points not used for calibration (fig. 2). In order to eliminate the interference caused by the drift, a reconstruction \( M_i \) at step \( i \) is transformed using the estimated model and backprojected just onto the following image \( i + 1 \). We evaluate backprojected against extracted points by the mean differences \( \Delta u, \Delta v \) in their \( u \)- and \( v \)-coordinates, and by the mean Euclidean distance \( \sqrt{\Delta u^2 + \Delta v^2} \) in left- and right image \( \Delta L, \Delta R \). The results show that the estimation is unbiased and accurate to subpixels (fig. 3, table 1)

![Figure 3: Calibration of pan(left) and tilt(right) joint:](image)

(top) Histogram of calibration errors \( \Delta u, \Delta v \). They seem to be normally distributed with zero mean, which indicates that the estimation is unbiased.

(bottom) Histogram of image distances \( \Delta L, \Delta R \). Left and right images show equal error distributions which accord with an assumed additive Gaussian noise.

### 6.2 Feed-forward prediction

In the prediction experiments, the mountain is observed from 9 general positions of the pan-tilt head, which are not on the calibration trajectories i.e. both joints are moved. For each pair of positions, the just established weak calibration is used to predict points in the second image pair from joint angles and a reconstruction down for the first image pair using (9), (2). Again, predicted points are evaluated against extracted points. The results show that the mean prediction error is about 1px (fig. 4, table 1).

![Figure 4: Feed-forward prediction over 35 trials:](image)

The mean differences \( \Delta u, \Delta v \) indicates the lateral offset, whereas \( \Delta LR \) gives the mean distance between predicted and extracted points.

<table>
<thead>
<tr>
<th>mean/sdv</th>
<th>(( \Psi, 0 ))</th>
<th>(0, ( \Phi ))</th>
<th>(( \Psi, \Phi ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta u )</td>
<td>-0.8e4/0.36</td>
<td>0.02/0.42</td>
<td>-0.04/0.73</td>
</tr>
<tr>
<td>( \Delta v )</td>
<td>-0.01/0.71</td>
<td>-0.04/0.50</td>
<td>-0.02/0.77</td>
</tr>
<tr>
<td>( \Delta LR )</td>
<td>0.66/–</td>
<td>0.51/–</td>
<td>0.77/–</td>
</tr>
</tbody>
</table>

Table 1: Image errors in calibration and prediction:
The columns from left to right show result of the calibration of pan-joint, tilt-joint, and the feed-forward experiment. The rows give mean values and standard deviations of the error measures \( \Delta u, \Delta v, \Delta LR \). All values are in pixel [px].

in image-based tracking, trajectory planning, and avoidance of occlusions, collisions, and targets out-of-view.

### 6.3 Homing

In the homing experiments, a stereo image i.e. at least a single reconstructed point \( M_h \) defines a home position of the pan-tilt system, for which the non-metric kinematics (9) is available. Given an unknown configuration of the pan-tilt head and the current image \( M_c \) of the reference point, which is the displacement in joint-space that moves the head to is home position? This amounts to solving the inverse kinematic map from the constraint

\[
M_h = H_\psi H_\phi M_c \rightarrow (\Psi, \phi) .
\]  

(32)

In the experiments, we used again the reprojection error as objective function (30) now to minimize over the two parameters \( (\Psi, \phi) \) (32), [9], [12]. The results are evaluated against the joint-angles measured when the home image was taken. The results presented in Figure 5 show that a single point is sufficient to home a pan-tilt head with a precision better than 0.1°.

Practical applications are visual reset of the pan-tilt system during power-up, visual measurement of joint-angles, visual homing, or target tracking with the pan-tilt mechanism.
7 Summary and Conclusions

In this paper, we presented an original approach to model the geometry a pan-tilt mechanism relative to a weakly calibrated stereo-rig, mounted on top of it. In contrast to existing methods for stereo self-calibration, the proposed method allows to avoid the very difficult metric calibration but still proves to be highly accurate – 1px in image-space and 0.1° in joint-space – even on real image data. Such a calibration allows to build fully operational systems e.g. for homing, surveillance or target tracking applications. The proposed mathematical formalism for modelling a kinematic chain – projective rotations and their product-of-exponentials – is sound and complete. It extends directly to 6R robot arms. Future work will concentrate on using this approach to model, calibrate and control general robots using a non-metric visual servoing approach.

References


