

Projective Rotations applied to a Pan-Tilt Stereo Head

Andreas Ruf

Radu Horaud

Andreas.Ruf@inrialpes.fr

Radu.Horaud@inrialpes.fr

GRAVIR-IMAG, INRIA Rhone-Alpes

655, avenue de l'Europe

38330 Montbonnot St.Martin, France

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 onto 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 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 Eu-

clidean 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 , θ 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 N

in Euclidean space projects onto the points \mathbf{m} and \mathbf{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* \mathcal{E}

$$\mathbf{m} \simeq [\mathbf{K}|\boldsymbol{\theta}] N, \quad \mathbf{m}' \simeq [\mathbf{K}'\mathbf{R}'|\mathbf{K}'\mathbf{t}'] N. \quad (1)$$

A weakly calibrated stereo rig is modeled as a pair of cameras with known epipolar geometry \mathbf{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)

$$\mathbf{m} \simeq \mathbf{P}\mathbf{M}, \quad \mathbf{m}' \simeq \mathbf{P}'\mathbf{M} \quad (2)$$

is relative to a *projective camera frame* \mathcal{P} , that is indeed defined by \mathbf{P} , \mathbf{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 \mathbf{K} , \mathbf{K}' , and $(\mathbf{R}', \mathbf{t}')$, the so-called *projective-Euclidean link* \mathbf{H}_{PE}

$$N \simeq \mathbf{H}_{PE}\mathbf{M}, \quad \mathbf{H}_{PE} \simeq \begin{bmatrix} \mathbf{K}^{-1} & \boldsymbol{\theta} \\ \mathbf{a}^T & 1 \end{bmatrix}. \quad (3)$$

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

Suppose now, that the rig undergoes a rigid motion [10] described in \mathcal{E} by $N' = \mathbf{T}_{RT}N$, and convince yourself that the corresponding homography in \mathcal{P} writes as

$$\mathbf{H}_{RT} = \gamma \mathbf{H}_{PE}^{-1} \mathbf{T}_{RT} \mathbf{H}_{PE}. \quad (4)$$

Such homographies \mathbf{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 ψ that is serially linked to a tilt joint with angle ϕ . Furthermore, a *zero-position* of this chain is defined by ψ_0, ϕ_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* \mathcal{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 $\mathbf{M}(\Psi, \Phi)$. In the zero-position, an *inertial 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 $\mathbf{M}(0, 0)$. The points N and \mathbf{M} are related by (3). Rather than attaching moving

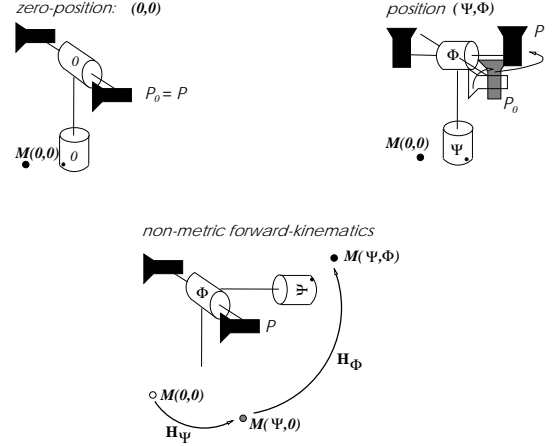


Figure 1: Non-metric pan-tilt stereo system:

Its kinematics is modeled by homographies that transform a projective reconstruction \mathbf{M} as a function of the joint angles (Ψ, Φ) .

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 \mathbf{T}_Ψ and \mathbf{T}_Φ , which represent *revolutions* of the pan and tilt joint. Most importantly, the joints have to be in zero-position relative to the moving frame \mathcal{E} (fig. 1).

$$N(\Psi, 0) = \mathbf{T}_\Psi \cdot N(0, 0), \quad (5)$$

$$N(\Psi, \Phi) = \mathbf{T}_\Phi \cdot N(\Psi, 0). \quad (6)$$

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

$$N(\Psi, \Phi) = \mathbf{T}_\Phi \mathbf{T}_\Psi N(0, 0). \quad (7)$$

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

of the pan-tilt system.

$$\begin{aligned} \mathbf{H}_{PE} \mathbf{M}(\Psi, \Phi) &\simeq \mathbf{T}_\Phi \mathbf{T}_\Psi \mathbf{H}_{PE} \mathbf{M}(0, 0) \\ \mathbf{M}(\Psi, \Phi) &\simeq \underbrace{\mathbf{H}_{PE}^{-1} \mathbf{T}_\Phi \mathbf{H}_{PE}}_{\mathbf{H}_\Phi} \underbrace{\mathbf{H}_{PE}^{-1} \mathbf{T}_\Psi \mathbf{H}_{PE}}_{\mathbf{H}_\Psi} \mathbf{M}(0, 0) \quad (8) \end{aligned}$$

As soon as we manage to write equation (8) without \mathbf{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 stereo-head*.

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

$$\mathbf{H}(\Psi, \Phi) = \mathbf{H}_\Phi \mathbf{H}_\Psi = \exp(\Phi \hat{\mathbf{H}}_\Phi) \exp(\Psi \hat{\mathbf{H}}_\Psi), \quad (9)$$

a product of two homographies \mathbf{H}_Φ and \mathbf{H}_Ψ (8). They are those elements of a class of *projective rotations* \mathbf{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 (Ψ, Φ) . This assumes the tangent operators $\hat{\mathbf{H}}_\Phi$, $\hat{\mathbf{H}}_\Psi$ of pan and tilt motions to be known, which amounts to a weak calibration of the mechanism (section 6.1). They can be 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 \mathbf{H}'_R acting on \mathcal{P} , which corresponds to the pure rotation \mathbf{T}_R acting on \mathcal{E} . Both transforms are algebraically similar by the PE-link (3) and scale γ

$$\mathbf{H}'_R = \gamma \mathbf{H}_{PE}^{-1} \mathbf{T}_R \mathbf{H}_{PE}. \quad (10)$$

A *normalized projective rotation* (*np-rotation*) $\mathbf{H}_R = \mathbf{H}'_R/\gamma$ has $\det \mathbf{H}_R = 1$ after division by

$$\gamma = (\det \mathbf{H}'_R)^{1/4} \text{sign}(\text{trace} \mathbf{H}'_R) \quad (11)$$

The *angle* of a np-rotations obeys as for \mathbf{T}_R

$$\cos(\theta) = 1/2(\text{trace} \mathbf{H}_R - 2). \quad (12)$$

4.2 Jordan Normal Form

A pure rotation has as eigenvalues the complex-conjugate pair $\lambda_\theta = \cos \theta \pm i \sin \theta$ and a double $\lambda_1 = 1$. In consequence, the real Jordan matrix has a rotation block $\mathbf{J}_2(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ and a Jordan-Block $\mathbf{J}_2(1)$ of order two [11].

The *Jordan matrix* \mathbf{J}_R of a np-rotation hence is

$$\mathbf{J}_R = \begin{bmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (13)$$

$$\mathbf{H}_R = \mathbf{H}_J^{-1} \mathbf{J}_R \mathbf{H}_J. \quad (14)$$

It is canonical up permutations of the blocks. Here, they are arranged like a rigid displacement, i.e. a rotation by θ around the z -axis. The decomposition (14) itself is ambiguous in \mathbf{H}_J . This ambiguity is characterized by the commutator \mathbf{C}_R of \mathbf{J}_R :

$$\mathbf{C}_R = \begin{bmatrix} a & -b & 0 & 0 \\ b & a & 0 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & e & f \end{bmatrix}, \quad \text{rank}(\mathbf{C}_R) = 4. \quad (15)$$

The complete *family of Jordan decompositions* thus is

$$\mathbf{H}_R = \mathbf{H}_J^{-1} \mathbf{C}_R^{-1} \mathbf{J}_R \mathbf{C}_R \mathbf{H}_J. \quad (16)$$

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:*

$$\mathbf{H}_R = \mathbf{I} + \sin \theta \hat{\mathbf{H}}_R + (1 - \cos \theta) \hat{\mathbf{H}}_R^2 \quad (17)$$

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

b) *It equals the solution of the matrix exponential*

$$\mathbf{H}_R = \exp(\theta \hat{\mathbf{H}}_R) \quad (18)$$

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

c) *This parameterization is unique.*

Proof: Eq. (14) writes:

$$\mathbf{H}_R = \mathbf{H}_J^{-1} (\mathbf{J}_R - \mathbf{I}) \mathbf{H}_J + \mathbf{I}$$

Denote the rows of \mathbf{H}_J by \mathbf{h}_i^T and the columns of \mathbf{H}_J^{-1} as \mathbf{k}_i . Remember also that

$$\mathbf{h}_i^T \mathbf{k}_j^T = \delta_{ij} \quad (19)$$

Then

$$\begin{aligned}\mathbf{H}_R &= \begin{bmatrix} k_1 & k_2 \end{bmatrix} \begin{bmatrix} \cos \theta - 1 & -\sin \theta \\ \sin \theta & \cos \theta - 1 \end{bmatrix} \begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix} + \mathbf{I} \\ &= \mathbf{I} + \sin \theta (\mathbf{k}_2 \mathbf{h}_1^T - \mathbf{k}_1 \mathbf{h}_2^T) + \dots \\ &\dots + \cos \theta - 1 (\mathbf{k}_1 \mathbf{h}_1^T + \mathbf{k}_2 \mathbf{h}_2^T). \end{aligned} \quad (20)$$

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

$$\hat{\mathbf{H}}_R = \mathbf{k}_2 \mathbf{h}_1^T - \mathbf{k}_1 \mathbf{h}_2^T \quad (21)$$

$$\hat{\mathbf{H}}_R^2 = \mathbf{k}_1 \mathbf{h}_1^T + \mathbf{k}_2 \mathbf{h}_2^T, \quad (22)$$

and further by induction that

$$\hat{\mathbf{H}}_R^{2n+2} = (-1)^{n+1} \hat{\mathbf{H}}_R^2 \quad (23)$$

$$\hat{\mathbf{H}}_R^{2n+1} = (-1)^n \hat{\mathbf{H}}_R. \quad (24)$$

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

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

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

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 $\hat{\mathbf{H}}_R$ (27). The commutativity of the two matrices

$$\mathbf{C}_A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad \mathbf{J}_A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (25)$$

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

$$\hat{\mathbf{H}}_R = \begin{bmatrix} k_1 & k_2 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix} \quad (26)$$

$$= \begin{bmatrix} k_1 & k_2 \end{bmatrix} \mathbf{C}_A^{-1} \mathbf{J}_A \mathbf{C}_A \begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix}. \blacksquare \quad (27)$$

4.4 Lie group property

Consider the one-parameter group of np-rotations \mathbf{H}_R with a fix \mathbf{H}_J (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 $\lim_{\theta \rightarrow 0} \frac{\mathbf{H}_R(\theta) - \mathbf{I}}{\theta}$:

$$\lim_{\theta \rightarrow 0} \frac{(1 - \cos \theta) \hat{\mathbf{H}}_R^2 + \sin \theta \hat{\mathbf{H}}_R}{\theta} = \hat{\mathbf{H}}_R. \quad (28)$$

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

Hence, after fixing the family \mathbf{H}_J , 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 $\hat{\mathbf{H}}_R$ 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 the algebra can be directly calculated from a given np-rotation using its inverse, like a rotation axis is calculated using the transpose.

$$\hat{\mathbf{H}}_R = \frac{1}{2 \sin \theta} (\mathbf{H}_R - \mathbf{H}_R^{-1}), \quad (29)$$

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 $\hat{\mathbf{H}}_R$ 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 Ψ_i . In practice, the stereo head is moved by operating the pan joint to Ψ_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 reconstruction¹ at Ψ_0 . It is transformed to $\mathbf{M}_i = \mathbf{M}(\Psi_i, 0)$ 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 $\hat{\mathbf{H}}_R$ (section 5.2)

$$\begin{aligned} \min_{\mathbf{x}} \sum_{i=1}^n & d(\mathbf{m}_i, \mathbf{P} \exp(\Psi_i \cdot \hat{\mathbf{H}}_\Psi(\mathbf{x})) \mathbf{M}_0) + \dots \\ & + d(\mathbf{m}'_i, \mathbf{P}' \exp(\Psi_i \cdot \hat{\mathbf{H}}_\Psi(\mathbf{x})) \mathbf{M}_0). \end{aligned} \quad (30)$$

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}^T , and 5 in the pure rotation \mathbf{T}_R . It is known, that from a single rotation, the plane at infinity is not fully

¹The 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 allow us to express $\hat{\mathbf{H}}_R(\mathbf{x})$ depending on a vector \mathbf{x} of 10 parameters.

Given an initial guess, a two-parameter family of solutions is introduced by the square $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ (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 $\mathbf{Q} \cdot \mathbf{W}$ with column-permutations \mathbf{P} to ensure non-zero, dominant elements on the diagonal and a good numerical condition.

$$\begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix} = \mathbf{Q} \cdot \mathbf{W} \cdot \mathbf{P}, \quad \mathbf{W} = \begin{bmatrix} 1 & h'_{12} & h'_{13} & h'_{14} \\ 0 & h'_{22} & h'_{23} & h'_{24} \end{bmatrix} = \begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix}$$

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

$$\hat{\mathbf{H}}'_R = \begin{bmatrix} k'_1 & k'_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} h_1^T \\ h_2^T \end{bmatrix}, \quad \begin{bmatrix} k'_1 & k'_2 \end{bmatrix} = \begin{bmatrix} k_1 & k_2 \end{bmatrix} \mathbf{Q}. \quad (31)$$

With a good guess, just one initial QR -step is sufficient. Otherwise h'_{11} , h'_{21} 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 $\hat{\mathbf{H}}_R(\mathbf{x})$ depends only on a vector \mathbf{x} of the 10 parameters $h'_{12}, h'_{13}, h'_{14}, h'_{22}, h'_{23}, h'_{24}, k'_{13}, k'_{14}, k'_{23}, k'_{24}$.

5.3 Initialization

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

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

The left and right eigenvectors f_{-i}, f_i and e_{-i}^T, e_i^T to the eigenvalues $-i$ and i can be recombined using the matrix $\Upsilon = \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}$ 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 Cartesian 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 (Ψ, Φ) , 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.

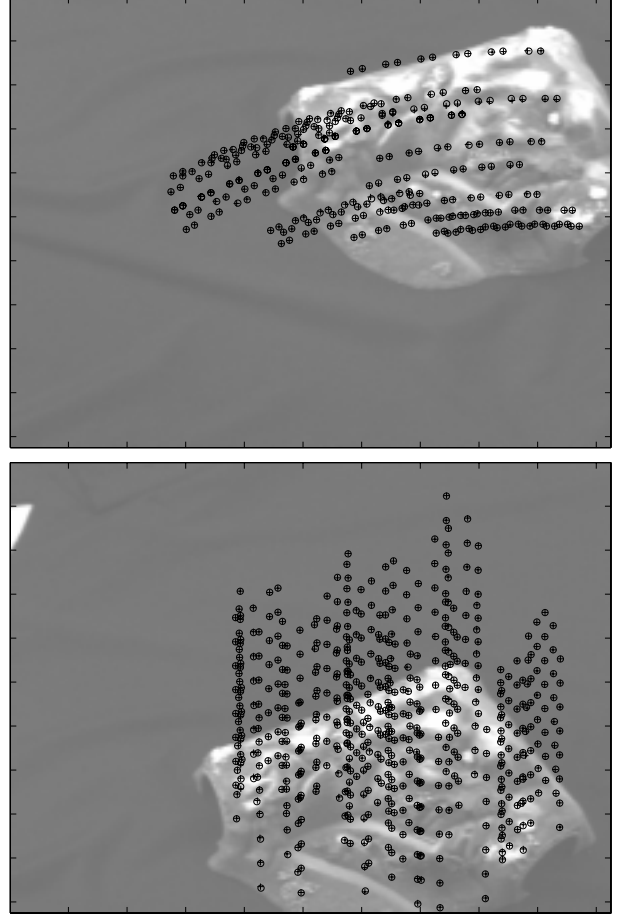


Figure 2: Calibration of pan(top) and tilt(bottom) joint: The images show the mountain model and superimposed point trajectories of pan and tilt motion, respectively. The missing steps are those used to calibrate. The + mark the backprojected points, the o mark the actual extracted points.

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 $\hat{\mathbf{H}}_\Psi$ and $\hat{\mathbf{H}}_\Phi$ (9) using our approach proposed in section (5). The result is evaluated against the steps and points not used for calibration (fig. 2). In order to eliminate the interference caused by the drift, a reconstruction \mathbf{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 Δu , Δv in their u - and v -coordinates, and by the mean Euclidean distance $\sqrt{\Delta u^2 + \Delta v^2}$ in left- and right image ΔL , ΔR . The results show that the estimation is unbiased and accurate to subpixels (fig. 3, table 1)

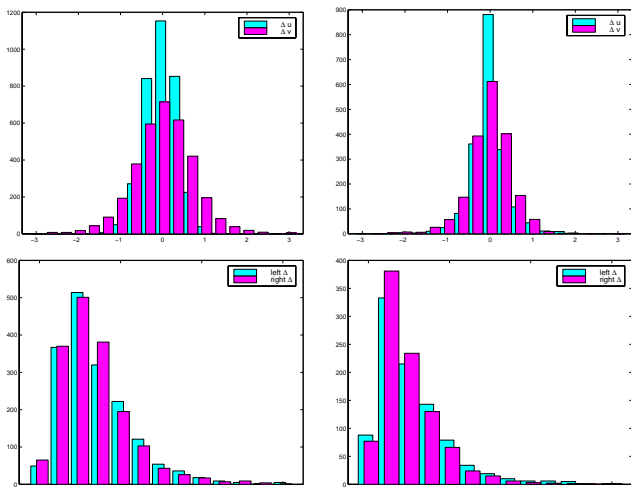


Figure 3: Calibration of pan(left) and tilt(right) joint: (top) Histogram of calibration errors Δu , Δv . They seem to be normally distributed with zero mean, which indicates that the estimation is unbiased. (bottom) Histogram of image distances ΔL , ΔR . Left and right images show equal error distributions 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).

Practical applications of the feed-forward prediction are

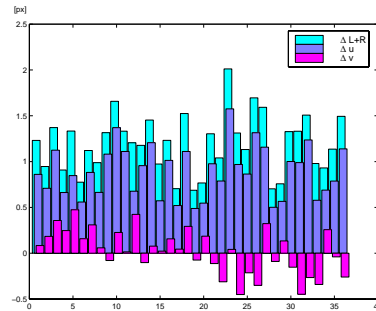


Figure 4: Feed-forward prediction over 35 trials: The mean differences Δu , Δv indicates the lateral offset, whereas ΔLR gives the mean distance between predicted and extracted points.

mean/sdv	$(\Psi, 0)$	$(0, \Phi)$	(Ψ, Φ)
Δu	-0.8e4/0.36	0.02/0.42	-0.04/0.73
Δv	-0.01/0.71	-0.04/0.50	-0.02/ 0.77
ΔLR	0.66/-	0.51/-	0.77/-

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 Δu , Δv , ΔLR . 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 \mathbf{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 \mathbf{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

$$\mathbf{M}_h = \mathbf{H}_\Phi \mathbf{H}_\Psi \mathbf{M}_c \rightarrow (\Psi, \phi). \quad (32)$$

In the experiments, we used again the reprojection error as objective function (30) now to minimize over the two parameters (Ψ, ϕ) (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.

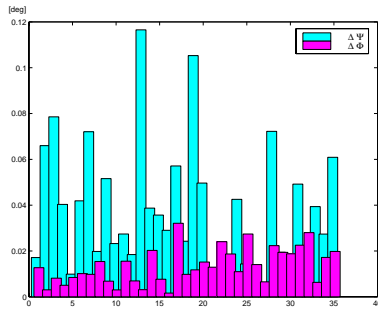


Figure 5: Visual homing using a single point: Joint-space error over 35 trials.

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

- [1] R. W. Brockett. Robot manipulators and the product of exponentials formula. In *Proc. MTNS-83 Int. Symposium*, pages 120–120, Beer-Sheva, Israel, 1983.
- [2] I. Bronstein and K. Semendjajew. *Teubner-Taschenbuch der Mathematik*, volume I. B.G. Teubner, Stuttgart Leipzig, 1996.
- [3] G. Csurka, D. Demirdjian, A. Ruf, and R. Horaud. Closed-form solutions for the euclidean calibration of a stereo rig. In *Proc. 5. ECCV*, volume I, pages 426–442. Springer, 1998.
- [4] D. Demirdjian, G. Csurka, and R. Horaud. Autocalibration in the presence of critical motions. In *Proc. BMVC*. Springer, 1998.
- [5] F. Devernay and O. Faugeras. From projective to Euclidean reconstruction. In *Proc. IEEE Conf. CVPR*, pages 264–269, San Francisco, CA., June 1996.
- [6] O. D. Faugeras. *Three Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Boston, 1993.
- [7] R. I. Hartley. Euclidean reconstruction from uncalibrated views. In M. Z. Forsyth, editor, *Applications of Invariance in Computer Vision*, pages 237–256. Springer Verlag, Berlin Heidelberg, 1994.
- [8] R. I. Hartley. In defence of the 8-point algorithm. In *Proc. 5. ICCV*, pages 1064–1070, Cambridge, Mass., June 1995. IEEE Computer Society Press.
- [9] N. Hollinghurst and R. Cipolla. Uncalibrated stereo hand-eye coordination. *Image and Vision Computing*, 12(3):187–192, March 1994.
- [10] R. Horaud and G. Csurka. Self-calibration and euclidean reconstruction using motions of a stereo rig. In *Proc. 6. ICCV*, pages 96–103. IEEE Computer Society Press, January 1998.
- [11] R. A. Horn and C. A. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1994.
- [12] M. Jagersand and R. Nelson. Acquiring visual-motor models for precision manipulation with robot hands. In Buxton-Cipolla, editor, *Proc. 4. ECCV*, volume II, pages 603–612. Springer, April 1996.
- [13] Q.-T. Luong and T. Viéville. Canonic representations for the geometries of multiple projective views. *Computer Vision and Image Understanding*, 64(2):193–229, September 1996.
- [14] J. McCarthy. *An Introduction to Theoretical Kinematics*. MIT Press, 1996.
- [15] B. Mooring, Z. Roth, and M. Driels. *Fundamentals of Manipulator Calibration*. John Wiley & Sons, New York, 1991.
- [16] A. Ruf, G. Csurka, and R. Horaud. Projective translations and affine stereo calibration. In *Proc. IEEE Conf. CVPR*, pages 475–481. IEEE Computer Society Press, June 1998.
- [17] J. Selig. *Geometrical Methods in Robotics*. Springer, 1996.
- [18] A. Zisserman, P. A. Beardsley, and I. D. Reid. Metric calibration of a stereo rig. In *Proc. IEEE Workshop on Representation of Visual Scenes*, pages 93–100, June 1995.