Conjugate Rotation: Parameterization and Estimation from an Affine Feature Correspondence

Kevin Köser, Christian Beder and Reinhard Koch
Christian-Albrechts-University
Kiel, Germany
{koeser,beder,rk}@mip.informatik.uni-kiel.de

Abstract

When rotating a pinhole camera, images are related by the infinite homography KRK^{-1}, which is algebraically a conjugate rotation. Although being a very common image transformation, e.g. important for self-calibration or panoramic image mosaicing, it is not completely understood yet. We show that a conjugate rotation has 7 degrees of freedom (as opposed to 8 for a general homography) and give a minimal parameterization. To estimate the conjugate rotation, authors traditionally made use of point correspondences, which can be seen as local zero order Taylor approximations to the image transformation. Recently however, affine feature correspondences have become increasingly popular. We observe that each such affine correspondence now provides a local first order Taylor approximation, which has not been exploited in the context of geometry estimation before. Using those two novel concepts above, we finally show that it is possible to estimate a conjugate rotation from a single affine feature correspondence under the assumption of square pixels and zero skew. As a byproduct, the proposed algorithm directly yields rotation, focal length and principal point.

1. Introduction

The *infinite homography* $\mathsf{H}_\infty = \mathsf{K}_1 R \mathsf{K}_2^{-1}$ is an image-to-image transformation, which relates points in one image with points in another image, if the camera has either only rotated or the corresponding 3D point is infinitely far away. It is a very important concept in self-calibration [22], projective geometry [9], or when dealing with purely rotating cameras, e.g. with pan-tilt-units [4], or for creating panoramic image mosaics[3].

In this work we concentrate on the case of constant intrinsic camera parameters, i.e. we assume $\mathsf{K}_1 = \mathsf{K}_2$. Then, algebraically a 3x3-matrix H acting in the projective image space $\mathbb{P}^2$ is such an infinite homography, if and only if it is proportional to a conjugate rotation, i.e. has the same eigenvalue structure as a scaled rotation matrix [9].

However, the dimension and structure of the set of conjugate rotations within the space of all possible homographies has not been fully understood yet, so that no algorithm for the direct computation of general conjugate rotations currently exists. So far a solution exists only for the special case when nearly all intrinsics (skew, aspect ratio, principal point) are known exactly [2]. In the general case, researchers typically estimate general homographies (e.g. using direct linear transformation [9] on $n \geq 4$ point correspondences) and state that - in the presence of little noise - the estimate should not be too far from a conjugate rotation [10, 4]. However, even in this case enforcing the "conjugate rotation constraints" afterwards is not as straightforward as for instance in the 8-point algorithm [9] for the fundamental matrix, because the eigenvalue decomposition of H_∞ will in general contain complex vectors. Simply projecting onto the allowed manifold has not been possible, because neither the dimension of this manifold has been known nor a suitable minimal parameterization is available.

In this work we will propose such a minimal parameterization for the conjugate rotation and show that the set of conjugate rotations is a 7-dimensional manifold in the space of all 3x3-matrices $\mathbb{R}^9$. As a second contribution we will show how to estimate a conjugate rotation from a single affine feature (cf. to [19]) correspondence, which provides already 6 constraints, using the additional assumption of zero skew and square pixels. This is in contrast to the algorithm presented in [2], which also requires the principal point to be known exactly, which is not always available.

This work is structured as follows: since the proposed parameterization and estimation is based upon a novel differential constraint from affine feature correspondences, we will start by explaining the concept of such a correspondence (a first order Taylor approximation of H_∞) in section 2 before we come to the conjugate rotation in section 3. In section 4, we will present a direct way of computing the conjugate rotation from a single affine correspondence and

978-1-4244-2243-2/08/$25.00 ©2008 IEEE

either another point or line correspondence or using the assumption of zero skew and square pixels. In the latter case rotation, focal length and principal point of the camera directly result from the proposed algorithm. Finally, in section 5 we will evaluate our novel algorithm, investigate its sensitivity to disturbances of the affine correspondence and show results in panoramic mosaicing with real images.

Notation: To improve the readability of the equations we use the following notation in this paper. Boldface italic serif letters $\boldsymbol{x}$ denote Euclidean vectors while boldface upright serif letters $\mathbf{x}$ denote homogeneous vectors. For matrices we do not use serifs, so that Euclidean matrices are denoted as A and homogeneous matrices are denoted as A.

2. Affine Correspondences

Progress in robust local features (cf. to [19, 18] for a thorough discussion) allows automatic matching of images in which appearance of local regions undergoes approximately affine changes of brightness and/or of shape, e.g. for automated panorama generation[3] or scene reconstruction through wide-baseline matching[24]. The idea is that interesting features are detected in each image and that the surrounding region of each feature is normalized with respect to the local image structure in this region, leading to about the same normalized regions for correspondences in different images, which can be exploited for matching. The concatenation of the normalizations provides affine correspondences between different views, i.e. not only a point-to-point relation but also a relative transformation of the local region (imagine scale, shear or rotation).

Although such correspondences carry more information than the traditional point correspondence used in estimation of multiple view geometry [9], this additional information is rarely used. Approaches not using point correspondences deal with conic correspondences [12, 11], which typically lead to systems of quadratic equations or require lots of matrix factorizations. Others require the identification of locally planar rectangle correspondences [13] for H_∞-computation. Schmid and Zisserman[25] investigated the behavior of local curvature under homography mapping. Chum et al. noted in [5] that an affine correspondence is somehow equivalent to three point correspondences: in addition to the center point two further points can be detected in the feature coordinate system (the *local affine frame*). This allowed the estimation of a fundamental matrix from 3 affine feature correspondences (from which 9 point correspondence were generated). The "local sampling" of the affine feature concept was recently also adopted by Riggi et al. [23] for fundamental matrix estimation and by Perdoch et al. for extended essential matrix estimation[20].

In contrast to the latter we do not sample but use a compact analytic expression for the whole correspondence: We observe that the concatenation of the normalization transformations provides a good approximation to the first order Taylor expansion of the homography, i.e. that the affine transform is the linearization of the homography (see also figure 1):

Figure 1. An affine correspondence in two images related by an infinite homography H_∞: The linear transformation (e.g. shear, rotation, magnification) between the two magnified image regions approximates the local derivative of the global image-to-image mapping H_∞ in the center of the window. Considering also the center shift, the resulting affine transformation can be thought of being tangent to H_∞, which can be exploited for estimation.

$$\boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{H}(\boldsymbol{x}_0) + \left.\frac{\partial \boldsymbol{H}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_0} (\boldsymbol{x} - \boldsymbol{x}_0) + \dots \quad (1)$$

$$\mathit{A} \approx \left.\frac{\partial \boldsymbol{H}(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_0} \qquad \mathit{A} \in \mathbb{R}^{2\times 2} \quad (2)$$

Here $\boldsymbol{H} : \mathbb{R}^2 \to \mathbb{R}^2$ is the homography mapping between the two images in Euclidean coordinates and A represents local shear, scale and rotation between the two corresponding features.

This is actually already exploited for matching but has not been used for geometry estimation before. As will be seen in the next section, using this compact representation of a correspondence keeps local relations and allows for parameterizing the conjugate rotation, because it provides a differential constraint on the local image transformation.

2.1. Upgrade and Refinement

The considerations so far apply to affine covariant features (e.g. MSER[16]). However, if matches result from weaker features (e.g. DoG/SIFT[14]), the proposed method can also be applied. The main insight is that if a correct match has been established such that the local regions are approximately aligned, the affine transform based upon the relative parameters is already almost correct. The straightforward way is to compute the affine transformation directly from the local frames of these features. This is usually already quite close to correct, because the correct match is a result from a high image similarity, e.g. in small baseline matching. However, since we need an accurate estimate of the Jacobian of the image transformation, it is reasonable even for already affine features to apply a gradient-based

optimization of A using the Lucas-Kanade approach [15, 1]. The concept of A as the local derivative of the image transform now leads to a parameterization of the conjugate rotation, as is shown in the next section.

3. The Infinite Homography: A Conjugate Rotation

We will now derive a minimal parameterization for the conjugate rotation. Despite being a very important concept in multi view geometry, the number of degrees of freedom has not been investigated yet. Neither exists a parameterization with less than the 8 parameters (as the naive parameterization with 5 intrinsic parameters and 3 rotation parameters). Such an over-parameterization can cause trouble in optimization, e.g. degenerate covariance matrices in maximum-likelihood estimation. Some authors have simplified K for the conjugate rotation to pure diagonal shape with zero skew, known aspect ratio and principal point [2]. Consequently, in this simplified model only a subset of all possible conjugate rotations is allowed. Instead, we will now derive a minimal parameterization for general conjugate rotations and will then discuss estimation in the next section. A 2d homography mapping

$$\mathbf{x}' \simeq \mathsf{H}\mathbf{x} = \begin{pmatrix} \boldsymbol{h}_1^\mathsf{T} & \boldsymbol{d} \\ \boldsymbol{h}_2^\mathsf{T} & \\ \boldsymbol{h}_3^\mathsf{T} & 1 \end{pmatrix} \mathbf{x} \tag{3}$$

is expressed in Euclidean coordinates as

$$\boldsymbol{x}' = \frac{\begin{pmatrix} \boldsymbol{h}_1^\mathsf{T} \\ \boldsymbol{h}_2^\mathsf{T} \end{pmatrix} \boldsymbol{x} + \boldsymbol{d}}{\boldsymbol{h}_3^\mathsf{T}\boldsymbol{x} + 1} \tag{4}$$

Its derivative is

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \frac{\partial \boldsymbol{x}'}{\partial \boldsymbol{x}} = \tag{5}$$

$$\frac{(\boldsymbol{h}_3^\mathsf{T}\boldsymbol{x} + 1)\begin{pmatrix} \boldsymbol{h}_1^\mathsf{T} \\ \boldsymbol{h}_2^\mathsf{T} \end{pmatrix} - \begin{pmatrix} \boldsymbol{h}_1^\mathsf{T} \\ \boldsymbol{h}_2^\mathsf{T} \end{pmatrix} \boldsymbol{x}\boldsymbol{h}_3^\mathsf{T} - \boldsymbol{d}\boldsymbol{h}_3^\mathsf{T}}{(\boldsymbol{h}_3^\mathsf{T}\boldsymbol{x} + 1)^2}$$

We now change without loss of generality the coordinate systems of both images, such that $\boldsymbol{x} = (0,0)^\mathsf{T}$, then this simplifies to

$$\mathsf{A} = \begin{pmatrix} \boldsymbol{h}_1^\mathsf{T} \\ \boldsymbol{h}_2^\mathsf{T} \end{pmatrix} - \boldsymbol{d}\boldsymbol{h}_3^\mathsf{T} \tag{6}$$

Solving this for $\boldsymbol{h}_3$ and using $\boldsymbol{d} = \boldsymbol{x}' - \boldsymbol{x}$, the homography given $\boldsymbol{x}'$ and A is therefore

$$\mathsf{H} = \begin{pmatrix} \mathsf{A} + (\boldsymbol{x}' - \boldsymbol{x})\boldsymbol{h}_3^\mathsf{T} & \boldsymbol{x}' - \boldsymbol{x} \\ \boldsymbol{h}_3^\mathsf{T} & 1 \end{pmatrix} \tag{7}$$

$$= \begin{pmatrix} I_2 & \boldsymbol{x}' - \boldsymbol{x} \\ \mathbf{0}_2^\mathsf{T} & 1 \end{pmatrix} \begin{pmatrix} \mathsf{A} & \mathbf{0}_2 \\ \boldsymbol{h}_3^\mathsf{T} & 1 \end{pmatrix} \tag{8}$$

So far, H may be any homography and no special conjugate rotation assumptions have been made. We will now assume that H is proportional to a conjugate rotation, i.e.

$$\mathsf{H} = \lambda \mathsf{K}\mathsf{R}\mathsf{K}^{-1} \tag{9}$$

where R is the relative camera rotation and K is the camera calibration matrix holding focal length f, aspect ratio a, skew s and principal point $(c_x, c_y)^\mathsf{T}$:

$$\mathsf{K} = \begin{pmatrix} f & s & c_x \\ 0 & a\,f & c_y \\ 0 & 0 & 1 \end{pmatrix} \tag{10}$$

From the orthogonality of the rotation matrix R follows that its eigenvalues and therefore also the eigenvalues of $\frac{1}{\lambda}\mathsf{H}$ are $\{1, e^{i\phi}, e^{-i\phi}\}$. Exploiting that all eigenvalues have the same absolute value, Pollefeys et al. derived a fourth order polynomial constraint for self-calibration, called the modulus constraint [21], which is a *neccessary condition* for a conjugate rotation. In contrast to this the above parameterization now leads to a linear relation between h_{31} and h_{32}, which provides a *sufficient condition* for conjugate rotations. We therefore factorize its characteristic polynomial into its roots

$$\det\left(\frac{1}{\lambda}\mathsf{H} - \tau \mathsf{I}_3\right) = \alpha(\tau - 1)(\tau - e^{i\phi})(\tau - e^{-i\phi}) \tag{11}$$

Multiplying out both sides yields a 3rd order polynomial in τ on both sides of the equation.

$$c_3\tau^3 + c_2\tau^2 + c_1\tau + c_0 = \tag{12}$$

$$\alpha\tau^3 - \alpha(e^{i\phi} + e^{-i\phi} + 1)\tau^2 + \alpha(e^{i\phi} + e^{-i\phi} + 1)\tau - \alpha$$

where the coefficients c_i depend on H and λ. By comparison of the polynomial coefficients we eliminate the unknowns α and ϕ and obtain two constraints, which are equivalent to

$$\lambda^3 = \det \mathsf{A} \tag{13}$$

$$\lambda \mathrm{tr}(\mathsf{H}) = \frac{1}{2}((\mathrm{tr}(\mathsf{H}))^2 - \mathrm{tr}(\mathsf{H}^2)) \tag{14}$$

Observe that eq. (13) eliminates the scale factor λ from subsequent computations and that all homographies, which fulfill these constraints, must be conjugate rotations. We now insert equation (7) into those constraints and obtain the condition

$$\left((\lambda - \mathrm{tr}\mathsf{A})(\boldsymbol{x}' - \boldsymbol{x})^\mathsf{T} + (\boldsymbol{x}' - \boldsymbol{x})^\mathsf{T}\mathsf{A}^\mathsf{T}\right)\boldsymbol{h}_3 \tag{15}$$

$$= \frac{1}{2}(\mathrm{tr}\mathsf{A})^2 - \mathrm{tr}\mathsf{A} - \frac{1}{2}\mathrm{tr}\left(\mathsf{A}^2\right) - \lambda\left(\mathrm{tr}\mathsf{A} + 1\right)$$

which is linear in $\boldsymbol{h}_3$, so that, given an affine feature correspondence $(\boldsymbol{x} \leftrightarrow \boldsymbol{x}', \mathsf{A})$, only one unknown h_{32} is left, i.e. we can write

$$\boldsymbol{h}_3 = \begin{pmatrix} a h_{32} + b \\ h_{32} \end{pmatrix} \tag{16}$$

for some a and b derived from the linear constraint of equation (15). From a geometrical point of view the equation above enforces the fixpoint of the conjugate rotation: the fixpoint is the eigenvector corresponding to the eigenvalue 1 (the intersection of the rotation axis and the image plane).

We now have a family of homographies, which depends on the six parameters of the affine correspondence and one parameter of the equation above. In other words, the nine-dimensional H depends on seven parameters

$$\boldsymbol{p} = (a_{11}, a_{12}, a_{21}, a_{21}, d_1, d_2, h_{32})^\mathsf{T} \tag{17}$$

now. By construction, H must be a conjugate rotation and the manifold for conjugate rotation can have at most seven dimensions, since it depends on seven parameters only. We evaluate the Jacobian $\partial\mathsf{H}/\partial\boldsymbol{p}$, which is a 9 by 7 matrix with the partial derivatives of the 9 entries of H. Proving linear independence of seven of the columns is quite tedious and lengthy, but using a symbolic linear algebra processor[17] we obtained that

$$\text{rank}(\partial\mathsf{H}/\partial\boldsymbol{p}) = 7 \tag{18}$$

Intuitively this means, that if we vary $\boldsymbol{p}$ we can run in 7 orthogonal directions on the manifold and this defines the dimension of the manifold [8].

This may be surprising at first sight, since knowing the eigenvalue structure seems to be more information than a single constraint[21]. Note however, that the rotation angle ϕ is unknown and we therefore only know the absolute value of the second and the third eigenvalue. Also, since the characteristic polynomial is holomorphic (as all polynomials), complex conjugates of any root must also be a root and finally, in projective space a homography is equivalent to a scaled version, so we basically end up with the constraint "All eigenvalues have the same absolute value". In the next section we show how the conjugate rotation with its 7 degrees of freedom can be estimated based upon an affine correspondence, which already provides 6 constraints.

4. Estimation and Self-Calibration with Constraints

In the previous section we derived a homography of the form

$$\mathsf{H}(h_{32}) = \begin{pmatrix} \mathsf{I}_2 & \boldsymbol{x}' - \boldsymbol{x} \\ \mathbf{0}_2^\mathsf{T} & 1 \end{pmatrix} \begin{pmatrix} \mathsf{A} & \mathbf{0}_2 \\ \boldsymbol{h}_3^\mathsf{T} & 1 \end{pmatrix} \tag{19}$$

which, given an affine feature correspondence, depends only on one parameter h_{32}. Basically this means that the affine transform locally fixes the conjugate rotation, but the pre-image of the line at infinity $\boldsymbol{h}_3$ still depends on one unknown parameter: we do not know, what maps to infinity yet.

In order to determine this remaining parameter we need one additional constraint. This may be obtained from another point or line correspondence or from a constraint on the intrinsic camera parameters.

4.1. Additional Point or Line Correspondence

If an additional image point correspondence $(\boldsymbol{y}, \boldsymbol{y}')$ is given, it must fulfill the homography mapping (using the parameterization from equation (19))

$$\begin{aligned} \begin{pmatrix} \boldsymbol{y}' - \boldsymbol{x} \\ 1 \end{pmatrix} &\simeq \mathsf{H} \begin{pmatrix} \boldsymbol{y} - \boldsymbol{x} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \mathsf{I}_2 & \boldsymbol{x}' - \boldsymbol{x} \\ \mathbf{0}_2^\mathsf{T} & 1 \end{pmatrix} \begin{pmatrix} \mathsf{A} & \mathbf{0}_2 \\ \boldsymbol{h}_3^\mathsf{T} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{y} - \boldsymbol{x} \\ 1 \end{pmatrix} \end{aligned} \tag{20}$$

We bring the displacement matrix to the left hand side and require that the cross product of the left hand side and the right hand side is zero

$$\begin{aligned} &\begin{bmatrix} \boldsymbol{y}' - \boldsymbol{x}' \\ 1 \end{bmatrix}_\times \begin{pmatrix} \mathbf{0}_{2\times 2} \\ (\boldsymbol{y} - \boldsymbol{x})^\mathsf{T} \end{pmatrix} \boldsymbol{h}_3 \\ &= -\begin{bmatrix} \boldsymbol{y}' - \boldsymbol{x}' \\ 1 \end{bmatrix}_\times \begin{pmatrix} \mathsf{A}(\boldsymbol{y} - \boldsymbol{x}) \\ 1 \end{pmatrix} \end{aligned} \tag{21}$$

Selecting one of the first two rows yields a linear equation in $\boldsymbol{h}_3$, which in general[1] determines the last remaining degree of freedom and therefore the conjugate rotation without any restrictions on skew, aspect ratio, focal length or principal point. Alternatively, another line correspondence might be used, e.g. if the horizon can be found in both images. Lines are dual to points and backward-map with a transposed H, so basically the same linear algebra applies as in the point correspondence case.

Self calibration is now possible with the approach of Hartley [10]. Note however, that in contrast to the homography estimation method used in [10], our estimated homography will be a perfect conjugate rotation.

If on the other hand some intrinsics of the used camera are known beforehand, no additional correspondence is required for estimation of the infinite homography as will be shown next.

4.2. Constraints on the Intrinsics

If only a single affine feature correspondence is given, the remaining unknown h_{32} may be computed using constraints on the intrinsic camera parameters. We will assume zero skew and unit aspect ratio in the following, which is true for most consumer cameras on the market. The only other algorithm to estimate a conjugate rotation [2] additionally requires the exact principal point position (see figure 2 for the sensitivity of [2] to principal point deviations).

[1] In the case that the point is on the line between the fixpoint and the affine feature, equations (16) and (21) will not be linearly independent. In this case a different point must be used.

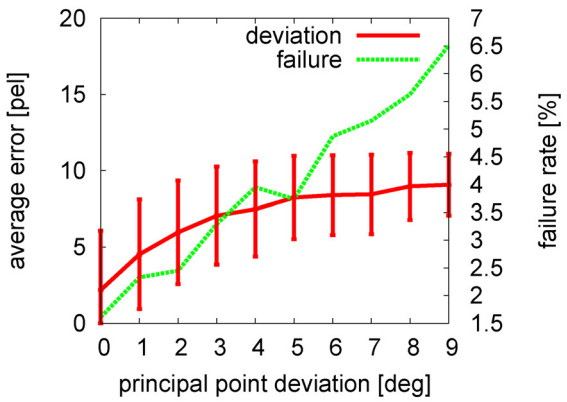

Figure 2. Synthetic evaluation of the sensitivity of the 2-point-algorithm [2] to principal point position (10.000 point pairs on a 50° field of view camera with width 1024 pixels), where we shifted the principal point several degrees away from the assumed position (the image center). The solid red curve shows the robust average error as evaluated by Brown et al. [2], while the dotted green curve shows the fraction of cases in which the algorithm did not come up with a solution at all. Already at 3° (5% image width) principal point error, the average error is above 6 pixels. Note that this is not a numerical or an implementation issue but caused by the resulting rays when the principal point varies.

Since often the principal point is only roughly known, e.g. close to the image center, our algorithm does not assume anything about the principal point.

In order to compute the remaining parameter h_{32} we use the image of the absolute conic (IAC, cf. [9])

$$\boldsymbol{\omega} = \left(\mathsf{K}\mathsf{K}^\mathsf{T}\right)^{-1} = \begin{pmatrix} \boldsymbol{\omega}_{11} & \boldsymbol{\omega}_{12} \\ \boldsymbol{\omega}_{12}^\mathsf{T} & t \end{pmatrix} \tag{22}$$

which is transformed by the conjugate rotation as follows

$$\mathsf{H}^\mathsf{T}\boldsymbol{\omega}\mathsf{H} = \lambda^2\mathsf{K}^{-\mathsf{T}}\boldsymbol{R}^\mathsf{T}\mathsf{K}^\mathsf{T}\mathsf{K}^{-\mathsf{T}}\mathsf{K}^{-1}\mathsf{K}\boldsymbol{R}\mathsf{K}^{-1} = \lambda^2\boldsymbol{\omega} \tag{23}$$

Collecting the entries of the upper triangular part of $\boldsymbol{\omega}_{11}$ in the vector $\boldsymbol{\kappa} = \text{vech}(\boldsymbol{\omega}_{11})$ (cf. [7]), it will be shown in appendix A, how a linear constraint $\boldsymbol{n}^\mathsf{T}\boldsymbol{\kappa} = 0$ can be derived from the above equation and that, given the affine feature correspondence, the vector $\boldsymbol{n}(h_{32})$ is a rational function of h_{32}.

The next step is to impose auto-calibration constraints on the IAC. We will assume zero skew, hence $\kappa_2 = 0$, and unit aspect ratio, hence $\kappa_1 = \kappa_3$. Together with the linear constraint $\boldsymbol{n}^\mathsf{T}\boldsymbol{\kappa} = 0$ from equation (43) (see appendix A for the derivation), we obtain the following homogeneous equation system

$$\begin{pmatrix} \boldsymbol{n}^\mathsf{T}(h_{32}) \\ 0 \quad 1 \quad 0 \\ 1 \quad 0 \quad -1 \end{pmatrix} \boldsymbol{\kappa} = \mathbf{0}_3 \tag{24}$$

Figure 3. Qualitative distribution of homography mapping error (length of error vector) in a sample image, black means low error while white means large error. Left: The homography mapping error is low near the affine feature and increases outwards (result from extrapolation). Center: In the 2-point algorithm the error can have two local minima near the two feature positions. Right: Error for DLT on 4 point correspondences. σ was 0.5 pixels and principal point distortion for 2-point was 1 degree.

This equation system can only have a non-trivial solution, if the determinant of the matrix is zero. Because $\boldsymbol{n}^\mathsf{T}$ is a rational function of h_{32}, also this determinant is a rational function of h_{32}. Its numerator is a quadratic polynomial in the unknown h_{32}, so that the desired solution can finally be computed. If the unknown intrinsic parameters are needed it is straightforward to substitute h_{32} backward to obtain the IAC.

5. Evaluation

So far we showed that a conjugate rotation has seven degrees of freedom and derived ways to estimate it from as few data as possible, which is interesting e.g. for RANSAC-like algorithms [6] or in scenarios where user initialization or interaction is required. In RANSAC-like algorithms the performance decreases exponentially with the number of correspondences needed to estimate a solution (see also [5]), e.g. 4 point correspondences with DLT or 2 (SIFT[14]) correspondences as proposed in [2]. Our method pushes this concept to the extreme such that we need only one affine correspondence, while we do not require the principal point to be known exactly. However, it is clear that in such a situation, where one local measurement determines a global transformation, small disturbances of the measurement can have severe effects on the extrapolated transformation. Figure 3 qualitatively shows for an example that the error is small at the correspondence and slowly grows in the vicinity while it becomes larger far away from the feature. This suggests the application of a growing strategy, which first incorporates nearby correspondences for estimation of the global homography before iterating and increasing the neighborhood radius. In the two-point algorithm [2] there are two local minima, because both features are forced to fit well.

To evaluate the sensitivity of our algorithm with respect to noise, we used the quality measure proposed in [2], where the average reprojection error across the overlap image re-

Figure 4. Conjugate rotation estimation from only one correspondence in each image pair. The first image is warped to the second and vice versa using the estimated homography.

gion is measured and clipped at 10 pixels to ensure robustness against gross errors in the homography estimation.

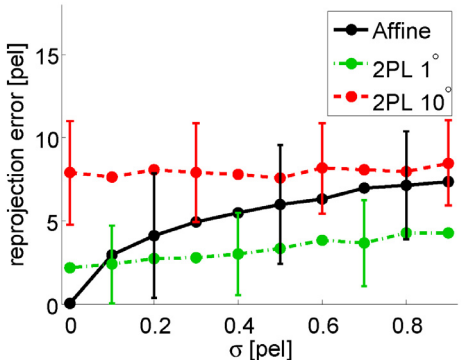

Figure 5. Sensitivity of the affine and the local 2-point algorithm against noise of the affine correspondence. The derivative is disturbed with 1% of the position uncertainty, which approximately evolves from the assumption that the corners of a patch for gradient-based minimization are found with the same uncertainty as the patch center. The images are of size 640x480 with 50° FOV and related by random rotations. The principal point prediction for the 2-point algorithm is disturbed with Gaussian noise of 1° (lower green curve) and 10° (upper red curve), which simulates that the principal point is only close to the image center for real cameras.

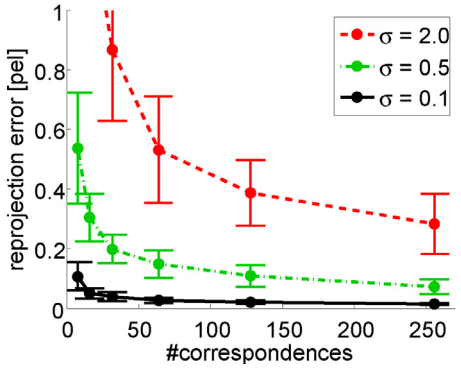

Figure 6. Quality of least-squares optimized conjugate rotation plotted against the number of correspondences used. The three curves represent three different standard deviations of the position noise, where we again used 1% noise in the homography's derivative. The setting is the same as in figure 5.

Figure 5 shows this quality measure plotted against the standard deviation of the noise on synthetic data. We compare our method to the two point algorithm [2] for different principal point distributions. For a fair comparison we generate the two correspondences by obtaining two points of distance 2 pixels from each affine feature as described in [23] (called *local two-point algorithm* in the following).

In this evaluation we observe that our algorithm requires a good affine correspondence. It is particularly sensitive to errors in the estimate of the homography's derivative. The error of the local two-point algorithm on the other hand is

dominated by the principal point distortion. However, if many correspondences are available the novel minimal parameterization allows for a simple nonlinear least squares maximum likelihood optimization of conjugate rotations given affine feature correspondences. Figure 6 shows the reprojection error plotted against the number of correspondences for different noise levels. The reprojection error decreases significantly if more affine feature correspondences are used.

Next we obtained SIFT-feature correspondences between the images pairs depicted in figure 4 according to the method of [3] and computed focal length and principal point from each of the correspondences, which we refined beforehand to enhance the accuracy of the homography's derivative estimate. While some of the estimates were reasonable, we observe that the variation in the resulting intrinsic parameters was too large, so that the one-feature approach for auto-calibration seems rather of theoretical interest.

Finally, we demonstrate some results on real images taken with different cameras as depicted in figure 4. From only one local correspondence we estimated a conjugate rotation using the proposed algorithm under the assumption of zero skew and square pixels. The images were stitched together and the results are shown in figure 4. Although not being subpixel correct, particularly in regions far away from the correspondence, we find the results quite appealing given the minimalistic data they are based upon.

6. Conclusion

We have shown that a general conjugate rotation has seven degrees of freedom and proposed a minimal parameterization. This parameterization arises from the insight that an affine feature correspondence provides a first order Taylor approximation to the image transformation, allowing for a differential constraint onto the homography.

The second major contribution of this work is an algorithm for estimating a conjugate rotation from a single affine feature correspondence under the assumption of zero skew and known aspect ratio involving nothing more expensive than the solution of a quadratic equation.

Also, our method does not require the principal point to be exactly at the image center, a crucial assumption to which previous methods are sensitive to, but which might not exactly be fulfilled in real cameras. Although not being suitable for auto-calibration, we have demonstrated that panoramic stitching is possible using only one single affine feature correspondence.

A. Constraints on the IAC

We will now show, how to derive a linear constraint $\boldsymbol{n}^\mathsf{T}\boldsymbol{\kappa} = 0$ on $\boldsymbol{\kappa} = \mathrm{vech}(\boldsymbol{\omega}_{11})$ (cf. equation (22) and [7]) from equation (23). Therefore we first compute the left hand side of equation (23) stepwise. Starting with the inner part, the translated conic is

$$\begin{pmatrix} I_2 & \mathbf{0}_2 \\ (\boldsymbol{x}'-\boldsymbol{x})^\mathsf{T} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\omega}_{11} & \boldsymbol{\omega}_{12} \\ \boldsymbol{\omega}_{12}^\mathsf{T} & t \end{pmatrix} \begin{pmatrix} I_2 & \boldsymbol{x}'-\boldsymbol{x} \\ \mathbf{0}_2^\mathsf{T} & 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\omega}_{11} & \boldsymbol{\phi}_{12} \\ \boldsymbol{\phi}_{12}^\mathsf{T} & u \end{pmatrix} \tag{25}$$

with the substitution

$$\boldsymbol{\phi}_{12} = \boldsymbol{\omega}_{12} + \boldsymbol{\omega}_{11}(\boldsymbol{x}'-\boldsymbol{x}) \tag{26}$$

$$u = t - (\boldsymbol{x}'-\boldsymbol{x})^\mathsf{T}\boldsymbol{\omega}_{11}(\boldsymbol{x}'-\boldsymbol{x}) + 2\boldsymbol{\phi}_{12}^\mathsf{T}(\boldsymbol{x}'-\boldsymbol{x}) \tag{27}$$

Next we apply the affine and projective part of the transformation yielding

$$\begin{pmatrix} A^\mathsf{T} & \boldsymbol{h}_3 \\ \mathbf{0}_2^\mathsf{T} & 1 \end{pmatrix} \begin{pmatrix} \boldsymbol{\omega}_{11} & \boldsymbol{\phi}_{12} \\ \boldsymbol{\phi}_{12}^\mathsf{T} & u \end{pmatrix} \begin{pmatrix} A & \mathbf{0}_2 \\ \boldsymbol{h}_3^\mathsf{T} & 1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\rho}_{11} & \boldsymbol{\rho}_{12} \\ \boldsymbol{\rho}_{12}^\mathsf{T} & u \end{pmatrix} \tag{28}$$

with the substitutions

$$\boldsymbol{\rho}_{11} = A^\mathsf{T}\boldsymbol{\omega}_{11}A + \boldsymbol{h}_3\boldsymbol{\phi}_{12}^\mathsf{T}A + A^\mathsf{T}\boldsymbol{\phi}_{12}\boldsymbol{h}_3^\mathsf{T} + \boldsymbol{h}_3 u \boldsymbol{h}_3^\mathsf{T} \tag{29}$$

$$\boldsymbol{\rho}_{12} = A^\mathsf{T}\boldsymbol{\phi}_{12} + \boldsymbol{h}_3 u \tag{30}$$

From equation (23) now follows $u = \lambda^2 t$, $\boldsymbol{\rho}_{12} = \lambda^2\boldsymbol{\omega}_{12}$ and $\boldsymbol{\rho}_{11} = \lambda^2\boldsymbol{\omega}_{11}$. We start by substituting $u = \lambda^2 t$ into equation (27) yielding

$$u = \frac{\lambda^2}{1-\lambda^2}((\boldsymbol{x}'-\boldsymbol{x})^\mathsf{T}\boldsymbol{\omega}_{11}(\boldsymbol{x}'-\boldsymbol{x}) - 2\boldsymbol{\phi}_{12}^\mathsf{T}(\boldsymbol{x}'-\boldsymbol{x})) \tag{31}$$

Next we solve equation (26) for $\boldsymbol{\omega}_{12}$ and obtain from the condition $\boldsymbol{\rho}_{12} = \lambda^2\boldsymbol{\omega}_{12}$ and equation (30) the equation

$$A^\mathsf{T}\boldsymbol{\phi}_{12} + \boldsymbol{h}_3 u = \lambda^2(\boldsymbol{\phi}_{12} - \boldsymbol{\omega}_{11}(\boldsymbol{x}'-\boldsymbol{x})) \tag{32}$$

Substituting equation (31) and solving this for $\boldsymbol{\phi}_{12}$ yields

$$\boldsymbol{\phi}_{12} = M\boldsymbol{\kappa} \tag{33}$$

with

$$M = P^{-1}Q((\boldsymbol{x}'-\boldsymbol{x})^\mathsf{T} \otimes I_2)D_2 \tag{34}$$

using the substitutions

$$P = \lambda^2\left(I_2 + \frac{2}{1-\lambda^2}\boldsymbol{h}_3(\boldsymbol{x}'-\boldsymbol{x})^\mathsf{T}\right) - A^\mathsf{T} \tag{35}$$

$$Q = \lambda^2\left(I_2 + \frac{1}{1-\lambda^2}\boldsymbol{h}_3(\boldsymbol{x}'-\boldsymbol{x})^\mathsf{T}\right) \tag{36}$$

and the duplication matrix

$$D_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{37}$$

Substituting this back into equation (31) yields

$$u = \boldsymbol{m}^\mathsf{T}\boldsymbol{\kappa} \tag{38}$$

with

$$\boldsymbol{m}^\mathsf{T} = \frac{\lambda^2}{1-\lambda^2}(\boldsymbol{x}'-\boldsymbol{x})^\mathsf{T}(I_2 - 2P^{-1}Q)((\boldsymbol{x}'-\boldsymbol{x})^\mathsf{T}\otimes I_2)D_2 \tag{39}$$

Finally we use the last remaining condition $\boldsymbol{\rho}_{11} = \lambda^2\boldsymbol{\omega}_{11}$ and obtain from equation (29) the condition

$$A^\mathsf{T}\boldsymbol{\omega}_{11}A + \boldsymbol{h}_3\boldsymbol{\phi}_{12}^\mathsf{T}A + A^\mathsf{T}\boldsymbol{\phi}_{12}\boldsymbol{h}_3^\mathsf{T} + \boldsymbol{h}_3 u \boldsymbol{h}_3^\mathsf{T} - \lambda^2\boldsymbol{\omega}_{11} = 0_2 \tag{40}$$

Substituting equations (33) and (38) yields the homogeneous linear equation system

$$N\boldsymbol{\kappa} = \mathbf{0}_4 \tag{41}$$

with

$$\begin{aligned} N \;=\; & (A^\mathsf{T}\otimes A^\mathsf{T})D_2 + (A^\mathsf{T}M)\otimes \boldsymbol{h}_3 \\ & + \boldsymbol{h}_3\otimes(A^\mathsf{T}M) + \boldsymbol{h}_3\otimes(\boldsymbol{h}_3\boldsymbol{m}^\mathsf{T}) - \lambda^2 D_2 \end{aligned} \tag{42}$$

This equation system contains only rational functions of h_{32} (the denominator is $\det P$, as $P^{-1} = P^*/\det P$) and turns out to have only rank 1. Hence we may select one single row $\boldsymbol{n}^\mathsf{T}$ from N

$$\boldsymbol{n}^\mathsf{T}\boldsymbol{\kappa} = 0 \tag{43}$$

which is the desired result.

Acknowledgments

This work has been supported by the German Research Foundation (DFG) in the project KO-2044/3-1.

References

[1] S. Baker and I. Matthews. Lucas-Kanade 20 Years On: A Unifying Framework. *International Journal of Computer Vision*, 56(3):221–255, 2004. 3

[2] M. Brown, R. Hartley, and D. Nistér. Minimal solutions for panoramic stitching. In *Proceedings of CVPR 2007*, 2007. 1, 3, 4, 5, 6

[3] M. Brown and D. G. Lowe. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, 74(1):59–73, 2007. 1, 2, 7

[4] D. Capel and A. Zisserman. Automatic mosaicing with super-resolution zoom. In *Proceedings of CVPR 1998*, page 885, Washington, DC, USA, 1998. IEEE Computer Society. 1

[5] O. Chum, J. Matas, and S. Obdrzalek. Epipolar geometry from three correspondences. In *Proc. Computer Vision Winter Workshop 2003, Prague*, pages 83–88, 2003. 2, 5

[6] M. Fischler and R. Bolles. RANdom SAmpling Consensus: a paradigm for model fitting with application to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. 5

[7] A. Fusiello. A matter of notation: Several uses of the kronecker product in 3d computer vision. *Pattern Recognition Letters*, 28:2127–2132, 2007. 5, 7

[8] A. Gray. *Modern Differential Geometry of Curves and Surfaces*. CRC Press, Boca Raton, Florida, 1994. 4

[9] R. Hartley and A. Zissermann. *Multiple View Geometry in Computer Vision*. Cambridge university press, second edition, 2004. 1, 2, 5

[10] R. I. Hartley. Self-calibration from multiple views with a rotating camera. In *LNCS 800 (ECCV 94)*, pages 471–478. Springer-Verlag, 1994. 1, 4

[11] F. Kahl and A. Heyden. Using conic correspondence in two images to estimate the epipolar geometry. In *Proceedings of ICCV*, pages 761–766, 1998. 2

[12] Kannala, Salo, and Heikkila. Algorithms for computing a planar homography from conics in correspondence. In *Proceedings of BMVC 2006*, 2006. 2

[13] J.-S. Kim and I. S. Kweon. Infinite homography estimation using two arbitrary planar rectangles. In *Proceedings of ACCV 2006*, pages 1–10, 2006. 2

[14] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004. 2, 5

[15] B. Lucas and T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. In *IJCAI81*, pages 674–679, 1981. 3

[16] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust Wide baseline Stereo from Maximally Stable Extremal Regions. In *Proceedings of BMVC02*, 2002. 2

[17] Matlab Symbolic Math Toolbox V.3.2(R2007a). 4

[18] K. Mikolajczyk and C. Schmid. A Performance Evaluation of local Descriptors. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 27(10):1615–1630, 2005. 2

[19] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. A Comparison of Affine Region Detectors. *International Journal of Computer Vision*, 65(1-2):43–72, 2005. 1, 2

[20] M. Perdoch, J. Matas, and O. Chum. Epipolar geometry from two correspondences. In *Proceedings of ICPR 2006*, pages 215–220, 2006. 2

[21] M. Pollefeys and L. V. Gool. Stratified self-calibration with the modulus constraint. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):707–724, 1999. 3, 4

[22] R.Hartley. Self-calibration of stationary cameras. *International Journal of Computer Vision*, 22(1):5–23, 1997. 1

[23] F. Riggi, M. Toews, and T. Arbel. Fundamental matrix estimation via TIP - transfer of invariant parameters. In *Proceedings of the 18th International Conference on Pattern Recognition*, pages 21–24, Hong Kong, August 2006. 2, 6

[24] F. Rothganger, S. Lazebnik, C. Schmid, and J. Ponce. 3d object modeling and recognition using local affine-invariant image descriptors and multi-view spatial constraints. *Int. J. Comput. Vision*, 66(3):231–259, 2006. 2

[25] C. Schmid and A. Zisserman. The geometry and matching of lines and curves over multiple views. *International Journal of Computer Vision*, 40(3):199–234, 2000. 2