Homography Estimation from Correspondences of Local Elliptical Features

Ondřej Chum   and   Jiří Matas

CMP, Department of Cybernetics, Faculty of EE, CTU in Prague
chum@cmp.felk.cvut.cz

Abstract

We propose a novel unified approach for homography estimation from two or more correspondences of local elliptical features. The method finds a homography defined by first-order Taylor expansions at two (or more) points. The approximations are affine transformations that are constrained by the ellipse-to-ellipse correspondences. Unlike methods based on projective invariants of conics, the proposed method generates only a single homography model per pair of ellipse correspondences. We show experimentally, that the proposed method generates models of precision comparable or better than the state-of-the-art at lower computational costs.

1. Introduction

The geometric relation between two perspective images of a planar scene is captured by a planar homography. The transformation is typically estimated from noisy correspondences of local affine-covariant features, often by a robust estimator from the RANSAC [5] family. The speed of RANSAC is critically dependent on the size of the minimal sample from which model parameters are calculated. Typically, the homography is estimated from four point-to-point correspondences [7]. To reduce the complexity of the estimation process, more complex constraints have to be explored, which is not straightforward.

In the paper, we propose a novel method for homography estimation from two correspondences of local elliptical features. Moreover, the same method can be directly used for a least-squares linear estimation from three or more correspondences.

Recently, the need for very efficient robust estimation of geometric relations, especially on large datasets, has triggered interest in constraints imposed by image

feature-to-feature correspondences. Fast spatial verification based on the estimation of an affine transformation of a vertical plane from a single elliptical correspondence was used in [15]. In [1], a method for estimation of an affine epipolar geometry from two elliptical correspondence was used in the context of specific object retrieval.

Two-view geometry of planar curves was a popular topic of computer vision [6, 8, 16, 9]. There are three well known methods for homography estimation from two conic-to-conic correspondences [17, 4, 10]. These methods are described in more detail in section 2.

All the above-mentioned methods for homography estimation from curve or conic correspondences exploit projective properties of curves. However, the vast majority of conics used in image matching are ellipses that are output by feature detectors. Since the feature detectors are typically covariant with up to an affine transformation [19, 12, 13, 14], the projective invariants may be systematically biased by such a construction.

Our approach is more closely related to homography estimation from point correspondences, *i.e.* the four-point algorithm [7]. Instead of knowing the transformation at four points in a general position, we present an algorithm that estimates the homography from known *values and gradients* of the transformation at two distinct points. We show that two ellipse-to-ellipse correspondences provide enough constraints on affine transformations, which are the first order Taylor approximations to the homography at given points. This allows for the homography estimation using linear algebra only.

The paper is structured as follows. Related methods are reviewed in sec. 2. The proposed method is described in sec. 3, and comparison with the state-of-the-art is given in sec. 4. Conclusions are drawn in sec. 5.

2. The Background and Related Work

A conic is represented [7] by a symmetric 3×3 matrix C. Points $\mathbf{x}$ lying on a conic C satisfy $\mathbf{x}^\top C \mathbf{x} = 0$. A conic C_i is transformed by a homography transforma-

tion H as

$$\lambda_i C_i' = H^{-\top} C_i H^{-1}. \qquad (1)$$

In [18] a linear method to estimate the so called conic-based transformation from seven conics correspondences is introduced, and the relation to the point-based homography matrix is shown.

A straightforward approach to homography estimation from a pair of conic correspondences is given in [17]. Two different conics have four (real or complex) points of intersection, that can be used to estimate the homography. The intersections are found by solving a quadratic polynomial. The combinatorial problem of matching the intersection points between the images is reduced to at most eight possible solutions in [17].

Pole-polar relations [7] are exploited in [4]. The homography is estimated from two pairs of corresponding conics $C_1 \leftrightarrow C_1'$ and $C_2 \leftrightarrow C_2'$. The homography is computed from virtual point-to-point correspondences, such that the virtual points are originally obtained as poles, and that both the pole and the corresponding polar are common to the two conics. Such poles are obtained as eigenvectors of $C_1^{-1}C_2$. Further points are obtained by intersection polars defined by the poles from the previous step. This method requires an initial decision on the matching poles for which an affine ordering of the poles is proposed. However, we found out that the proposed pole ordering often does not provide correct correspondence of the poles, which leads to a potentially high number of possible homography models.

The most relevant method to ours was proposed by Kannala et al. [10]. The homogeneous equation (1) is first de-homogenized by considering the determinants of the matrices representing the conics. Then from each pair of conic correspondences, a set of linear equations

$$C_i'^{-\top} C_j' H = H C_i^{-1} C_j \qquad (2)$$

is constructed. For three or more conic correspondences, these equations provide enough constraints for linear least-squares estimation of the homography H. The solution for the minimal case of a pair of conic correspondences is more difficult and the exact derivation can be found in [10]. We will only state that the homography H can be expressed as

$$H = F'^{-\top} Q P U^\top F^T,$$

where F and F' are complex matrices, Q and U are complex orthogonal matrices, and $P = \mathrm{diag}(\pm 1, \pm 1, \pm 1)$. This leads to a four-fold ambiguity in the solutions for H. All these solutions are complex in practice due to the measurement noise in the input conics matrices.
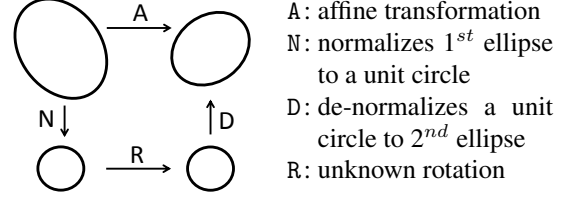

A: affine transformation
N: normalizes 1st ellipse to a unit circle
D: de-normalizes a unit circle to 2nd ellipse
R: unknown rotation

Figure 1. An ellipse-to-ellipse correspondence defines local affine transformation A up to an unknown rotation R.

3. The Method

To estimate the homography, we will study the constraints imposed by ellipse-to-ellipse correspondences on the homography through the first order Taylor approximations at two or more different points.

Let $\mathbf{x} = (x, y, 1)^\top$ be an image point in the first image, $\mathbf{x}' = (x', y', 1)^\top$ be a point in the second image, H be a regular matrix representing a planar homography

$$H = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix}, \text{ and}$$

$\lambda \mathbf{x}' = H\mathbf{x}$ for $\lambda = h_7 x + h_8 y + h_9 \neq 0$. A first order Taylor expansion $\bar{A}(H, \mathbf{x})$ of a homography H at a point $\mathbf{x}$ is an affine transformation [3] defined as

$$\begin{pmatrix} h_1 - x'h_7 & h_2 - x'h_8 & h_3 + x'(xh_7 + yh_8) \\ h_4 - y'h_7 & h_5 - y'h_8 & h_6 + y'(xh_7 + yh_8) \\ 0 & 0 & h_9 + h_7 x + h_8 y \end{pmatrix}. \qquad (3)$$

Let A be an affine transformation so that $\mathbf{x}' = A\mathbf{x}$. All homographies that are approximated (first order Taylor approximation) by an affine transformation A at a point $\mathbf{x}$ are expressed as

$$H = h_7 H_7 + h_8 H_8 + \lambda A, \qquad (4)$$

where

$$H_7 = \begin{pmatrix} x' & 0 & -xx' \\ y' & 0 & -xy' \\ 1 & 0 & -x \end{pmatrix}, \quad H_8 = \begin{pmatrix} 0 & x' & -yx' \\ 0 & y' & -yy' \\ 0 & 1 & -y \end{pmatrix}.$$

The equation (4) can be easily verified by substituting $\bar{A}(H, \mathbf{x})$ from (3) in place of λA. If two affine transformations locally approximating the homography at different points are known, equation (4) provides enough linear constraints to estimate the homography H.

A correspondence of two elliptical regions defines an affine transformation up to an unknown rotation (see Fig. 1)

$$A = DRN, \qquad (5)$$

where N is an affine transformation normalizing the ellipse in the first image to a unit circle, R is a rotation by an angle α, and D is an affine transformation denormalizing the unit circle into the ellipse in the second image. We will use notation of $c = \cos\alpha$ and $s = \sin\alpha$. Let $\mathbf{n_1}^\top$, $\mathbf{n_2}^\top$, and $\mathbf{n_3}^\top = (0,0,1)$ be rows of N. The equation (5) is linear in unknowns c and s, and can be rewritten as

$$ \mathtt{A} = c\mathtt{D} \begin{pmatrix} -\mathbf{n_2}^\top \\ \mathbf{n_1}^\top \\ \mathbf{0}^\top \end{pmatrix} + s\mathtt{D} \begin{pmatrix} \mathbf{n_1}^\top \\ \mathbf{n_2}^\top \\ \mathbf{0}^\top \end{pmatrix} + \mathtt{D} \begin{pmatrix} \mathbf{0}^\top \\ \mathbf{0}^\top \\ \mathbf{n_3}^\top \end{pmatrix} . \quad (6) $$

By substituting (6) into (4) we get seven linearly independent equations, linear in 12 unknowns (9 elements of H, λ, λc, and λs). By adding another pair of matching ellipses, seven equations and three new unknowns (λ_i, $\lambda_i c_i$, and $\lambda_i s_i$) are introduced. Therefore, with two pairs of matching ellipses, there are 14 linear equations and 15 unknowns, leading to one-dimensional space of solutions. Due to the homogeneous nature of the problem (note that by multiplying equation (4) by a non-zero scalar β multiplies all unknowns by β), the one dimensional space of solutions for H uniquely determines the homography transformation. For more than two pairs of ellipses, the method leads to a least squares problem.

Discussion. In the estimation, the quadratic constraint arising from the rotation R, $\cos^2\alpha + \sin^2\alpha = 1$ is ignored. In robust RANSAC-like [5] estimators, the constraint can be used to verify the validity of the estimated model. Hypotheses of invalid models are quickly rejected without the need to evaluate the consensus set.

The construction of the ellipse-to-ellipse mapping (eqn. 5 and Fig. 1) maps the center of gravity of one ellipse to the center of gravity of the other ellipse. This is not in general a constraint under a projective transformation of a plane. However, many approaches to wide-baseline stereo matching perform reduction of feature-to-feature matches to point-to-point matches [Tuytelarrs-BMVC00, Matas-IVC04] through the center of gravity. The error induced by such an approximation is negligible compared to detection errors. As a consequence, even in a noise-less case (two ellipse-to-ellipse correspondences exactly related by a homography) the algorithm returns only approximation to the true homography.

4. Experimental evaluation

We have experimentally compared the proposed method[1] with the method of Kannala et al. [10], for
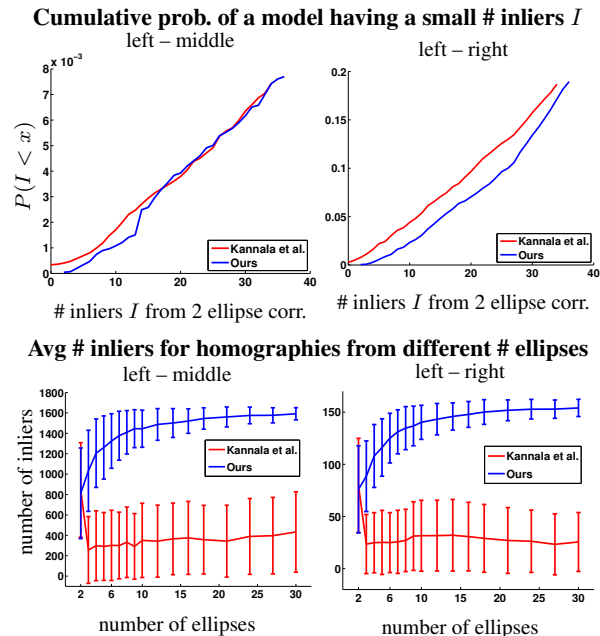

Figure 2. Three Graffiti images (top) and the experimental results.

which we used the publicly available implementation[2]. In the experiment, we present results on the standard Graffiti dataset [14] images. Two different settings were selected: smaller (left and middle image in Fig. 2 top) and large (left and right image in Fig. 2 top) change of the viewpoint.

In the images MSER [12] regions were detected and transformed to elliptical features. A commonly used affine covariant construction was applied: the elliptical features and the MSER regions have common moments up to the second order [19, 12]. The correspondence between the features were established based on SIFT [11] descriptors. The precision of the homography estimation from two pairs of matching ellipses was measured by the number of inliers to the models. The ground truth pairs of matching ellipses were selected by a RANSAC procedure and all mismatched features were discarded.

Minimal case - two ellipse correspondence. All pairs of corresponding ellipses were used to estimate the homography. The number of true inliers to those models were recorded. In the case of [10], the highest number of inliers from the four hypothesized homographies was

---

[1]Matlab code available http://cmp.felk.cvut.cz/~chum/code/ell2h.html

[2]http://www.ee.oulu.fi/~jkannala/bmvc.html

selected.

From the bottom row of Fig. 2 it can be seen that the method of [10] produces slightly more precise models from a pair of ellipse correspondences (the left most point on the plots). However, for a method estimating a model from minimal samples in a RANSAC-like robust estimation, it is important to produce a low fraction of models from an all-inlier sample with small support. Models with sufficient support are upgraded by standard tools, such as local optimization in LO-RANSAC [2], while models with small support are simply ignored. The plots in Fig. 2 (middle row) show the cumulative distribution of a probability that an estimated homography will have a certain number of inliers. It can be observed that the proposed method has lower probability, compared to [10], of returning a model with low number of inliers.

A further advantage of the proposed method is that it generates a single model per pair of corresponding ellipses. Compared to four hypotheses of [10], it represents a four fold speed-up in the verification phase. Moreover, our method is about 20% faster[3] than [10].

Least squares solution. In this experiment, we have generated non-minimal samples of 3 to 30 random ground-truth ellipse correspondences. The plots in Fig. 2 (bottom row) show the average number of inliers (over 500 repetitions) with standard deviation. It can be seen that the precision of our method increases with the number of ellipses, the method of [10] produces the best results with only two ellipse correspondences. We explain such a behaviour by an instability of the linear algebraic constrains (2) used in the least squares method of [10].

5. Conclusions

We have proposed a novel method for homography estimation from correspondences of local elliptical features. The proposed method generates models with precision comparable or higher than the state-of-the-art. Unlike methods based on projective invariants of conics, the proposed method generates only a single homography model per a pair of ellipse correspondences. The method is also very efficient and thus suitable as a hypothesis generator in RANSAC-like robust estimators.

References

[1] R. Arandjelović and A. Zisserman. Efficient image retrieval for 3D structures. In *Proc. BMVC.*, 2010.

[2] O. Chum, J. Matas, and J. Kittler. Locally optimized RANSAC. In *Proc. DAGM*. Springer-Verlag, 2003.

[3] O. Chum, T. Pajdla, and P. Sturm. The geometric error for homographies. *CVIU*, 97(1):86–102, 2005.

[4] C. Conomis. Conics-based homography estimation from invariant points and pole-polar relationships. In *3DPVT*, pages 908–915, 2006.

[5] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, June 1981.

[6] D. A. Forsyth, J. L. Mundy, A. Zisserman, C. Coelho, A. Heller, and C. A. Rothwell. Invariant descriptors for 3-D object recognition and pose. *IEEE PAMI*, 13(10):971–991, 1991.

[7] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.

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

[9] J. Y. Kaminski and A. Shashua. Multiple view geometry of general algebraic curves. *IJCV*, 56(3):195–219, 2004.

[10] J. Kannala, M. Salo, and J. Heikkilä. Algorithms for computing a planar homography from conics in correspondence. In *Proc. BMVC.*, pages 77–86, 2006.

[11] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.

[12] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide-baseline stereo from maximally stable extremal regions. *IVC*, 22(10):761–767, Sep 2004.

[13] K. Mikolajczyk and C. Schmid. Scale & affine invariant interest point detectors. *IJCV*, 1(60):63–86, 2004.

[14] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. A comparison of affine region detectors. *IJCV*, 65(1/2):43–72, 2005.

[15] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatial matching. In *Proc. CVPR*, 2007.

[16] L. Quan. Conic reconstruction and correspondence from two views. *IEEE PAMI*, 18(2):151–160, Feb 1996.

[17] C. Rothwell, A. Zisserman, C. Marinos, D. A. Forsyth, and J. L. Mundy. Relative motion and pose from arbitrary plane curves. *IVC*, 10(4):250–262, 1992.

[18] A. Sugimoto. A linear algorithm for computing the homography from conics in correspondence. *JMIV*, 13(2):115–130, 2000.

[19] T. Tuytelaars and L. Van Gool. Wide baseline stereo matching based on local, affinely invariant regions. In *Proc. BMVC.*, 2000.

---

[3]Comparing MATLAB implementations that are available online.