Simultaneously Estimating the Fundamental Matrix and Homographies

Pei Chen and David Suter, *Senior Member, IEEE*

***Abstract*—The estimation of the fundamental matrix (FM) and/or one or more homographies between two views is of great interest for a number of computer vision and robotics tasks. We consider the joint estimation of the FM and one or more homographies. Given point matches between two views (and assuming rigid geometry of the camera-scene displacement), it is well known that all of the matched points satisfy the epipolar constraint that is usually characterized by the FM. Subsets of these point matches may also obey a constraint characterized by a homography (all matches in the subset coming from three-dimensional (3-D) points lying on a 3-D plane). The estimations of homographies and the FM are well-studied problems, and therefore, the (separate) estimation of the FM, or the homography matrices, can be considered as effectively solved problems with mature algorithms. However, the homographies and FM are not independent of each other: therefore, separate estimation of each is likely to be suboptimal. In this paper, we propose to simultaneously estimate the FM and homographies by employing the compatibility constraint between them. This is done by first concentrating on a set of parameters that (jointly) parameterize the entire set of homographies and FM (simultaneously) and that also implicitly enforce the compatibility between the estimates of each set. We then derive a reduced form with the purpose of improving the speed. We propose a solution method in which the Sampson error for the FM and homographies is minimized by the Levenberg–Marquardt (LM) algorithm. Experiments show that the gains can be compared with separate estimates (the FM and/or the homographies).**

***Index Terms*—Compatibility, fundamental matrix, homography.**

I. Introduction

Fundamental matrix (FM) estimation (or essential matrix estimation for a calibrated camera[1]) and homography estimation are fundamental problems in computer vision and robotics, so much so that one could cite many potential applications. Here, however, we simply point to those applications that particularly require high-quality homography estimation in robotics ([2], [16]).

Being important problems, they have been extensively investigated in [2], [5], [6], [9], [10], [11], [13]–[17], [22], [23], and [25]. Indeed, the (separate) estimation of the FM or homographies is a mature technology. However, they are not independent from each other, and there exists a compatibility constraint between the FM and homographies. A sufficient and necessary condition [10], [14], for a homography H to be compatible with the FM F, is that $H^T F$ is skew-symmetric, i.e.,

$$H^T F + F^T H = 0 \tag{1}$$

which is referred to as the compatibility constraint.

Manuscript received September 11, 2008; revised March 8, 2009. First published September 22, 2009; current version published December 8, 2009. This paper was recommended for publication by Associate Editor S. Hirai and Editor L. Parker upon evaluation of the reviewers' comments. The work of P. Chen was supported by the Ministry of Science and Technology, People's Republic of China, under the 973 Program 2006CB303104.

P. Chen is with the School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510275, China, and also with Shenzhen Institute of Advanced Integration Technology, Chinese Academy Sciences/Chinese University of Hong Kong, Shenzhen 518055, China (e-mail: chenpei@mail.sysu.edu.cn).

D. Suter is with the School of Computer Science, The University of Adelaide, Adelaide, S.A. 5005, Australia.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TRO.2009.2030224

[1]We work on uncalibrated cameras in this paper, and therefore, we focus on the FM.

If one separately estimates the FM and homographies from noisy data, their compatibility cannot be ensured. This leads to suboptimal estimation because the compatibility constraint can be utilized to potentially improve the *estimation accuracy*.

Methods have been proposed to estimate homographies given the FM, or *vice versa*, ensuring some degree of compatibility. For example, given F, a compatible homography is estimated from three correspondences [10]. Conversely, given H, the epipole $\mathbf{e}'$ is estimated from two correspondences that are not on the H-related plane, and then, $F = [\mathbf{e}']_\times H$. Similar techniques are employed in other structure from motion problems [12], [18], [19]. In [12], [18], and [19], the left 3×3 part of the projection matrices is first estimated from the homographies, and then, the rightmost 3×1 vector of the projection matrices is estimated. However, the approaches mentioned earlier are still suboptimal, as they are essentially sequential.

There are other possible approaches to compatible estimates. For instance, one can use the bundle adjustment (BA) algorithm [21] that tries to simultaneously estimate projection matrices (*principal parameters*) and three-dimensional (3-D) feature points (*nuisance parameters*). The BA algorithm has a heavier computational burden than other algorithms that only estimate the principal parameters, while hiding the nuisance parameters [4]. The issue of computational inefficiency will be worse when using the BA algorithm to estimate the FM and one or more homographies because additional constraints have to be added to enforce the planarity of the homography-related points, adding to the computational complexity.

Our approach falls into the category of the method that only estimates the principle parameters (i.e., the FM and homography matrices). We utilize a particular parameterization form of homography matrix and FM so that the compatibility condition (1) is *implicitly* satisfied. In addition, we prove that for points on planes, the related homography constraint implies the epipolar constraint when using such a parameterization form. Based on the above fact, we extend the Sampson error [10], [20], [23], [24] to the case of simultaneously estimating the FM and homographies. Thus, the parameters can be estimated by routine, unconstrained, optimization methods, for instance, the Levenberg–Marquardt (LM) algorithm.

Related paper includes jointly estimating several homographies [3] by exploiting the rank constraints. By employing the rank constraints, the estimation accuracy of the homographies is improved, for example, by approximately 20% in the cases of three or four planes. However, there are limitations in the method described in [3]. First, it can only be applied to at least three planes. Second, points on none of the planes are not exploited (potentially, accuracy is improved by using more data).

The rest of the paper is organized as follows. In Section II, we propose to simultaneously estimate the FM and homographies: First, we re-parameterize the FM and homographies so that their compatibility is implicitly satisfied and that the redundancy among them is reduced. Then, we use the LM algorithm to estimate the parameters, where the Sampson error is to be minimized. Simulations and the real experiment are presented in Section III.

II. Simultaneously Estimating the Fundamental Matrix and Homographies

We use bold symbols ($\mathbf{x}$ and $\mathbf{x}'$, for the points on the first and second views, respectively) to denote the homogeneous representation of 2-D features. $d(\mathbf{x}, \mathbf{y})$ denotes for the Euclidean distance between inhomogeneous points represented by $\mathbf{x}$ and $\mathbf{y}$, respectively.

A. Problem

Given the matches $\{\mathbf{x}_i\}$ and $\{\mathbf{x}'_i\}$, the estimation of the FM F and one or more homographies $\{H_k\}$[2] is to minimize the following objective function:

$$\text{Obj}(F, \{H_k\}; \{\hat{\mathbf{x}}_i\}, \{\hat{\mathbf{x}}'_i\}) = \sum_i d^2(\mathbf{x}_i, \hat{\mathbf{x}}_i) + d^2(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)$$
$$\text{subject to} \quad c_i(\hat{\mathbf{x}}_i, \hat{\mathbf{x}}'_i; F, \{H_k\}) = 0 \tag{2}$$

where the constraint $c_i(\hat{\mathbf{x}}_i, \hat{\mathbf{x}}'_i; F, \{H_k\}) = 0$ is either an epipolar constraint $\hat{\mathbf{x}}_i'^T F \hat{\mathbf{x}}_i = 0$ when the point is on none of the planes or *jointly* an epipolar constraint $\hat{\mathbf{x}}_i'^T F \hat{\mathbf{x}}_i = 0$ and a homography constraint of $H_k \hat{\mathbf{x}}_i \sim \hat{\mathbf{x}}'_i$ when the ith point is on the kth plane (because a pair of matches on one 3-D plane still satisfies the epipolar constraint.)

The constrained minimizer of (2) satisfies the compatibility constraint (1), although we show (next section) how this can be done without its explicit inclusion as a formal constraint.

In (2), one has to find not only the homographies and the FM, but the nuisance parameters $\{\hat{\mathbf{x}}_i\}$ and $\{\hat{\mathbf{x}}'_i\}$ (in some sense, the optimally corrected points—so that they are exactly consistent with the found homographies and FM) as well. A direct solution for such a constrained optimization problem, with nuisance parameters $\{\hat{\mathbf{x}}_i\}$ and $\{\hat{\mathbf{x}}'_i\}$, has a very heavy computational burden [24].

B. Compatible Parameterization for FM and Homographies

We know from [10] that given the FM F, the three-parameter family of homographies induced by a world plane is

$$H = A - \mathbf{e}'\bar{\mathbf{v}}^T \tag{3}$$

where

$$[\mathbf{e}']_\times A = F \tag{4}$$

is any decomposition of the FM.

The skew symmetric matrix associated with a three-vector $\mathbf{a}$ is defined as

$$[\mathbf{a}]_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}.$$

Note that later, we will use the notation $[\mathbf{a}]_{\times,2} \in R^{3,2}$ formed by taking two columns of $[\mathbf{a}]_\times$, for example, the first two columns assuming $a_3 \neq 0$ [3]. It holds that

$$[\mathbf{a}]_\times = [\mathbf{a}]_{\times,2} \begin{bmatrix} 1 & 0 & -a_1/a_3 \\ 0 & 1 & -a_2/a_3 \end{bmatrix} \tag{5}$$

By substituting, we see that H in (3) and F in (4) satisfy the compatibility condition (1).

Equations (3) and (4) suggest a BA-type approach, where these parameters (A, $\mathbf{e}'$, and $\bar{\mathbf{v}}$) are optimized (by LM, for example), using any of the objective functions related to reprojection error. To the best of our knowledge, no one has directly used such an approach to estimate the FM and homographies (jointly). (The above parameterization of (3) and (4) was employed only to estimate a compatible homography given the FM, or *vice versa* [10].) In any case, there are two significant issues that we address here: optimally estimating the FM and homographies in such a parameterization needs an elegant treatment of the objective function (see Section II-E); moreover, one can work with a reduced parameterization (see the next section).

C. Reduced Form

We now derive a reduced parameterization that leads to the improved speed.

There is overparameterization in using A, $\mathbf{e}'$, and $\bar{\mathbf{v}}$ to represent H in (3) and F in (4). For example, A can take values given by

$$A_\eta = A + \mathbf{e}'\eta^T \tag{6}$$

with any choice of $\eta \in R^3$. Note that $F = [\mathbf{e}']_\times A_\eta$ still holds when A_η is used in place of A in (4) and that (3) should be then replaced by $H = A_\eta - \mathbf{e}'(\bar{\mathbf{v}} + \eta)^T$. For a particular choice of A, $\bar{\mathbf{v}}$ is uniquely determined in (3).

For reducing (potentially) the computational complexity, the redundancy in parameters should be reduced as much as possible.

Since F has a rank of two, it can be expressed as

$$F = [\mathbf{e}']_{\times,2} R \tag{7}$$

where $R \in R^{2,3}$. From (4), (5), (7), and the fact that $[\mathbf{e}']_{\times,2}$ is of full-column rank, it holds that

$$R = \begin{bmatrix} 1 & 0 & -e'_1/e'_3 \\ 0 & 1 & -e'_2/e'_3 \end{bmatrix} A. \tag{8}$$

By (8), R is determined from A. Equation (7) can be expressed as $F = [\mathbf{e}']_\times A_F$, where

$$A_F = \begin{bmatrix} R \\ \mathbf{0} \end{bmatrix}.$$

Consequently, the homography is determined by a three-vector $\mathbf{v}$

$$H = \begin{bmatrix} R \\ \mathbf{0} \end{bmatrix} - \mathbf{e}'\mathbf{v}^T. \tag{9}$$

H in (9) and F in (7) still satisfy (1).

There is still some redundancy in using (7) and (9) to represent F and H. The FM has seven degrees of freedom (DOF) [10], and each homography is determined by specification of an (additional) three-vector. Thus, there are $(7 + 3n)$-DOF for the FM and n homographies. However, there are $9 + 3n$ parameters in (7) and (9). The difference is due to the fact that the size of both $\mathbf{e}'$ and R is arbitrary. However, we do not wish to impose explicit constraints (such as unit magnitude) as such constraints complicate the optimization.

D. Calculation of R, $\mathbf{e}'$, and $\{\mathbf{v}_k\}$ from FM and Homographies

We calculate F and $\{H_k\}$ from observed feature points and then calculate R, $\mathbf{e}'$, and $\{\mathbf{v}_k\}$ from the coefficients of F and $\{H_k\}$.

[2] In the following, matrices, vectors, and scalars are represented by italic uppercase, bold lowercase, and italic letters, respectively. Subscripts are used for three purposes. A subscript of an italic uppercase (or bold lowercase) letter denotes the indice, for instance, of homography (the kth homography as H_k) or that of $\mathbf{x}_i$ (or $\mathbf{v}_k$ in Sections II-D and E). An italic letter with a subscript denotes one component of the associated vector, for instance, a_1 for the first component of the vector $\mathbf{a}$. I_k denotes the $k \times k$ identity matrix. No explicit explanation is added when the meaning of the subscript can be interpreted from the context.

[3] If $a_3 = 0$, we can select the first and third (or the second and third) columns of $[\mathbf{a}]_\times$, and (5) and (8)) should be modified accordingly.

1) The epipole $\mathbf{e}'$ is calculated as the left singular vector of F, associated with the least singular value.
2) Jointly calculate A from F and $\{H_k\}$:
   a) Calculate R by (7) and set $A_F = \begin{bmatrix} R \\ \mathbf{0} \end{bmatrix}$.
   b) Calculate A'_F and A_k by (12) and (10), respectively.
   c) Make a matrix by normalizing and vectorizing each of $\{A'_F, \{A_k\}\}$ and calculate its left singular vector associated with the largest singular value. Reorder this singular vector as the 3×3 matrix A.
3) Calculate R by (8).
4) Calculate $\{\mathbf{v}_k\}$ by (9).

Fig. 1. Algorithm 1. Initial estimates of R, $\mathbf{e}'$, and $\{\mathbf{v}_k\}$ from F and $\{H_k\}$.

1) First, $\mathbf{e}'$ is set to the left null vector of F (the left singular vector with the least singular value, in practice).

2) Then, A can be determined from the FM or a homography. From F, one choice of A is

$$A_F = \begin{bmatrix} R \\ \mathbf{0} \end{bmatrix}$$

where R can be calculated from (7). Alternatively, from H_k, one choice of A is

$$A_k = (I_3 - \mathbf{e}'\mathbf{e}'^T)H_k. \tag{10}$$

We compute A jointly from the FM and homographies. From (3) and (6), the following quality is independent of the vector $\bar{\mathbf{v}}$ or η and, consequently, is invariant, up to an unknown scale:

$$(I_3 - \mathbf{e}'\mathbf{e}'^T)A_\eta = (I_3 - \mathbf{e}'\mathbf{e}'^T)H = (I_3 - \mathbf{e}'\mathbf{e}'^T)A. \tag{11}$$

Define

$$A'_F = (I_3 - \mathbf{e}'\mathbf{e}'^T)A_F. \tag{12}$$

The vectorization forms of $\{A'_F, A_k\}$ lie in a one-dimension subspace. In practice, A is calculated in the following way: Normalize each of $\{A'_F, A_k\}$, and form a matrix by vectorizing $\{A'_F, A_k\}$; calculate its left singular vector associated with the largest singular value; and (finally) reorder this singular vector as the 3×3 matrix A.

3) R is calculated from (8), with A calculated above.

4) $\mathbf{v}_k$ is calculated by (9).

These steps are summarized in Fig. 1.

E. Optimal Estimate of R, $\mathbf{e}'$, and $\{\mathbf{v}_k\}$

The estimates of R, $\mathbf{e}'$, and $\{\mathbf{v}_k\}$, by Algorithm 1 in Fig. 1, usually have poor accuracy because the error in the FM and homography parameters is implicitly assumed to be independent identically distributed (i.i.d.) Gaussian. In this section, starting from an initial estimate, we use the LM algorithm to iteratively estimate (i.e., improve) R, $\mathbf{e}'$, and $\{\mathbf{v}_k\}$.

From (7) and (9), F and $\{H_k\}$ can be calculated from R, $\mathbf{e}'$, and $\{\mathbf{v}_k\}$. Thus, in accordance with (2), the optimal solution of estimating the FM and one or more homographies is to estimate the optimal R, $\mathbf{e}'$, and $\{\mathbf{v}_k\}$, as well as $\{\hat{\mathbf{x}}_i\}$ and $\{\hat{\mathbf{x}}'_i\}$, by minimizing the following objective function:

$$\text{Obj}(R, \mathbf{e}', \{\mathbf{v}_k\}; \{\hat{\mathbf{x}}_i\}, \{\hat{\mathbf{x}}'_i\}) = \sum_i d^2(\mathbf{x}_i, \hat{\mathbf{x}}_i) + d^2(\mathbf{x}'_i, \hat{\mathbf{x}}'_i)$$

$$\text{subject to} \quad c_i(\hat{\mathbf{x}}_i, \hat{\mathbf{x}}'_i; R, \mathbf{e}', \{\mathbf{v}_k\}) = 0 \tag{13}$$

where the constraint $c_i(\hat{\mathbf{x}}_i, \hat{\mathbf{x}}'_i; R, \mathbf{e}', \{\mathbf{v}_k\}) = 0$ is either an epipolar constraint or *jointly* an epipolar constraint and a homography constraint, as described in (2).

In order to eliminate the nuisance parameters $\{\hat{\mathbf{x}}_i\}$ and $\{\hat{\mathbf{x}}'_i\}$, the *Sampson error* [10], [20], [23], and [24] is used to replace the Euclidean distance square in a 2-D image. It is a first-order approximation of the Euclidean distance square.

The Sampson error is recommended for FM estimation in [24] because the computational burden is about 1/50th of that value when the Euclidean distance square is minimized, while the performance degradation is negligible when the Sampson error is minimized.

The Sampson error for the epipolar constraint [4], [10], and [24] $\hat{\mathbf{x}}'^T_i F \hat{\mathbf{x}}_i = 0$ is

$$S_{F,i} = \frac{(\mathbf{x}'^T_i F \mathbf{x}_i)^2}{(F\mathbf{x}_i)_1^2 + (F\mathbf{x}_i)_2^2 + (F^T\mathbf{x}'_i)_1^2 + (F^T\mathbf{x}'_i)_2^2} \tag{14}$$

where $(F\mathbf{x}_i)_j^2$ denotes the square of the jth entry of the vector $F\mathbf{x}_i$.

The Sampson error for the homography constraint [10] $H_k\hat{\mathbf{x}}_i \sim \hat{\mathbf{x}}'_i$ is

$$S_{H,k,i} = \epsilon^T_{k,i}(J_{k,i}J^T_{k,i})^{-1}\epsilon_{k,i} \tag{15}$$

where $\epsilon_{k,i}$ is a two-vector, i.e., the *algebraic error* of the ith match, and $J_{k,i}$ is a 2×4 matrix, i.e., the partial derivative of $\epsilon_{k,i}$, with respect to $\mathbf{x}_{k,i}$ and $\mathbf{x}'_{k,i}$. For more details (the derivation of Sampson error and the definition of $\epsilon_{k,i}$ and $J_{k,i}$ for homography estimation), see [10] (also Appendix A for completeness.)

Here, care should be taken to minimize the Sampson-error approximation described in (13), particularly for points on 3-D planes, i.e., points satisfying both the associated homography constraint and the epipolar constraint. For such points [in the form of Sampson error in (15)], the algebraic error ϵ is a three-vector, and J is a 3×4 matrix (there are three constraints: one with the epipolar constraint and two with the homography constraint). However, J has a rank of two. Consequently, the 3×3 matrix JJ^T has a rank of two, and its inverse in (15) does not make sense. This is caused by the fact that the homography constraint implies the epipolar constraint when using the parameterizations described in (3, 4) or (7, 9).[4] Particularly, the row of J, associated with the epipolar constraint, is not independent of the other two rows with the homography constraint. Thus, due to this dependence, ϵ and J can be reduced, by ignoring the epipolar constraint, to a two-vector and a 2×4 matrix, respectively (for points on 3-D planes), and the Sampson error still takes the form in (15). For a point not on any plane, the Sampson error takes the form in (14).

Thus, the optimization of (13) is reduced to an unconstrained optimization problem

$$\{\hat{R}, \hat{\mathbf{e}}', \{\hat{\mathbf{v}}_k\}\} = \arg\min_{R, \mathbf{e}', \{\mathbf{v}_k\}} \left(\sum_{i \notin \cup I_k} S_{F,i} + \sum_k \sum_{i \in I_k} S_{H,k,i} \right) \tag{16}$$

where I_k is the index set of points that are on the kth plane. It should be stressed that only the points *not on any plane* are included when calculating the objective function of $S_{F,i}$ in (16). (The reason for this is proved in Appendix B.) Note that, in (16), only the principal parameters of the FM and homography matrices (that is, R, $\mathbf{e}'$, and $\{\mathbf{v}_k\}$) are to be estimated.

Then, F and $\{H_k\}$ are calculated from $\hat{R}$, $\hat{\mathbf{e}}'$, and $\{\hat{\mathbf{v}}_k\}$ by (7) and (9). Such estimates of the FM and homography matrices satisfy the compatibility condition (1).

[4] See the proof in Appendix B. Note that the points satisfying a homography constraint will satisfy the related epipolar constraint if both the homography and FM are noise-free. This also holds when the homography and FM are compatible, as can be seen in Appendix B. However, it is not true, generally, if there is error in their coefficients, for instance, the homography and FM are separately estimated from noisy data.

1) Calculate F and $\{H_k\}$, by the normalized eight-point algorithm or normalized DLT.
2) Use algorithm 1 in figure 1 to calculate the initial value of R, $\mathbf{e}'$ and $\{\mathbf{v}_k\}$.
3) Use the LM algorithm to iteratively estimate the final R, $\mathbf{e}'$ and $\{\mathbf{v}_k\}$, by minimizing the objective function in (16).
4) Calculate F and $\{H_k\}$ from R, $\mathbf{e}'$ and $\{\mathbf{v}_k\}$, by (7) and (9).

Fig. 2. Algorithm 2. Optimal estimates by the LM algorithm.

TABLE I
COMPUTATIONAL EFFICIENCY OBTAINED BY REPARAMETERIZING THE FM AND HOMOGRAPHIES AS IN (7) AND (9), INSTEAD OF USING (3) AND (4). NOTE THAT THE COMPUTATIONAL TIME OF USING (3) AND (4) IS REGARDED AS A UNIT

No. planes (n)	1	2	3	4
Computational time of algorithm 2 when using (7) and (9)	0.7937	0.8146	0.8323	0.8413

The initial R^0, $\mathbf{e}'^0$, and $\{\mathbf{v}_k^0\}$ are calculated, by Algorithm 1 in Fig. 1, from F and $\{H_k\}$, which are estimated by the normalized eight-point algorithm or normalized direct linear transformation (DLT) algorithm [10]. The approach is summarized in Fig. 2.

F. DOF and Complexity

$3n + 12$ parameters are used to represent the FM and n homographies in (3) and (4) and $3n + 9$ parameters in (7) and (9), which is difference (improvement) of 3 DOF. Since only a few planes (of significant size) are observed in both views in common situations, the implications of this (at first sight small change) are not negligible.

The main computational burden of the LM algorithm lies in calculating the Jacobian matrix and solving the augmented normal equations, whose size is the number of parameters. The complexity of solving a linear equation $M\mathbf{x} = \mathbf{b}$, with $M \in R^{m,m}$ is of $O(m^3)$ [8]. The calculation of the Jacobian matrix mainly consists of matrix–vector multiplications and matrix additions. It has an complexity of $O(m^2)$. However, the improvement of computational efficiency, in practice, is not exactly as expected above because the algorithm shares some part of the computations whether it uses (7, 9) or (3, 4). See, for example, in Table I, the computational time of Algorithm 2 when using (7) and (9), where that of using (3) and (4) is regarded as a *unit*.

III. EXPERIMENTS

We compare the proposed Algorithm 2, in Fig. 2 (referred to as SIMU[5]), with the two other methods: LM and normalized DLT/normalized eight-point. The LM method refers to separately estimating the FM or homographies by the LM algorithm. The normalized DLT/normalized eight point refers to the normalized DLT (or the normalized eight-point algorithm) of estimating the homography (or the FM, respectively).

A. Simulation

We conduct simulations on cases where one, two, three, or four planes are observed.

Experimental setting

1) *Feature points:* The camera matrices and 3-D points are randomly generated. Each plane has 20 3-D points, and another extra 20 3-D points are randomly generated. When generating 3-D points, the process is carried out in such a way that ensures that these points are visible (not obscured by the planes). The image points are scaled to be in the region of $(-256 \sim 256, -256 \sim 256)$.
2) *Noise:* Zero-mean i.i.d. Gaussian noise is added to 2-D points: σ varies from 0.5 to 2.5.
3) *Performance evaluation of $\hat{H}$:* We use the following reprojection error as the index for evaluating the estimated homography $\hat{H}$:

$$\sqrt{\frac{\sum_{i=1}^{p} d^2(\bar{\mathbf{x}}_i', \hat{H}\bar{\mathbf{x}}_i)}{p}} \tag{17}$$

where $\bar{\mathbf{x}}$ denotes for the ground truth of $\mathbf{x}$. Note that the ground truth homography has an error of zero.
4) *Performance evaluation of $\hat{F}$*: In evaluating the estimated $\hat{F}$, we use the following index by measuring its difference from the ground truth:

$$\sqrt{\frac{\sum_{i=1}^{p} \frac{(\bar{\mathbf{x}}_i'^T \hat{F}\bar{\mathbf{x}}_i)^2}{(\hat{F}\bar{\mathbf{x}}_i)_1^2 + (\hat{F}\bar{\mathbf{x}}_i)_2^2} + \frac{(\bar{\mathbf{x}}_i'^T \hat{F}\bar{\mathbf{x}}_i)^2}{(\hat{F}^T\bar{\mathbf{x}}_i')_1^2 + (\hat{F}^T\bar{\mathbf{x}}_i')_2^2}}{2p}}. \tag{18}$$

Note that this is approximately the distance (in pixels) between the two FMs, which has been introduced in [23]. Also note that the ground truth FM has an error of zero.
5) *Statistics:* We repeat this 100 times for each noise level σ.

From Fig. 3, two main conclusions can be drawn. First, the SIMU performs best. Second, as far as the homography is concerned, the quality of the estimates by the SIMU improves as the number of planes increases, compared with the other two methods. This phenomenon should be useful in homography-related robot tasks [2], [16].

B. Real Example

In this section, we show a real example on the corridor sequence [1]. The corridor sequence consists of 11 frames and 737 feature points. RANdom SAmple Consensus (RANSAC) [7] is used to detect three planes, each with 109/70/42 features. Initially, the fourth and fifth views are used in our experiment, with the purpose of illustration. Then, all pairs ($C_{11}^2 = 55$) between 11 frames are used to get statistics for comparison.

On this corridor example, we evaluate algorithms in terms of the accuracy of the FM. The calibrated projection matrices are available at [1], and consequently, the calibrated FM (referred to as CALI) can be calculated directly from these projection matrices [10]. The normalized eight-point algorithm is referred to as NORM.

The distance (in pixels) between two FMs, introduced in [23], is used to evaluate three estimates, which measures the difference between two pencils of epipolar lines defined by the two FMs. It is zero, if two matrices are equal (up to a scale). See [23] for more details.

In the case between the fourth and fifth views, the distances between four FMs are given in Table II. The most meaningful index for comparison is the distance between the estimated FM and the calibrated one (assumed to be the ground truth), which is 0.7270, 1.3982, and 3.5602 pixels, for the SIMU, LM, and NORM methods, respectively.

The estimated epipoles by three methods, with two epipolar lines for each epipole, are shown in Fig. 4 (with the CALI as the ground truth). The Euclidean distances of the estimated epipoles (by the SIMU, LM, and NORM) from the calibrated epipole (CALI) are 0.5232, 1.3065, and 2.4146 pixels, respectively. (Note that the distance (in pixels) between FMs is a more reliable index for comparison than the Euclidean distance between epipoles, because the epipole is the null vector of the FM, and the other information in the FM is discarded.)

The statistics calculated from all the 55 frame pairs are then calculated. Table III shows the mean values (with the variance in the bracket) of the distances between the four FMs. The Euclidean distances of the

[5]The Matlab code is available at http://sist.sysu.edu.cn/∼chenpei.

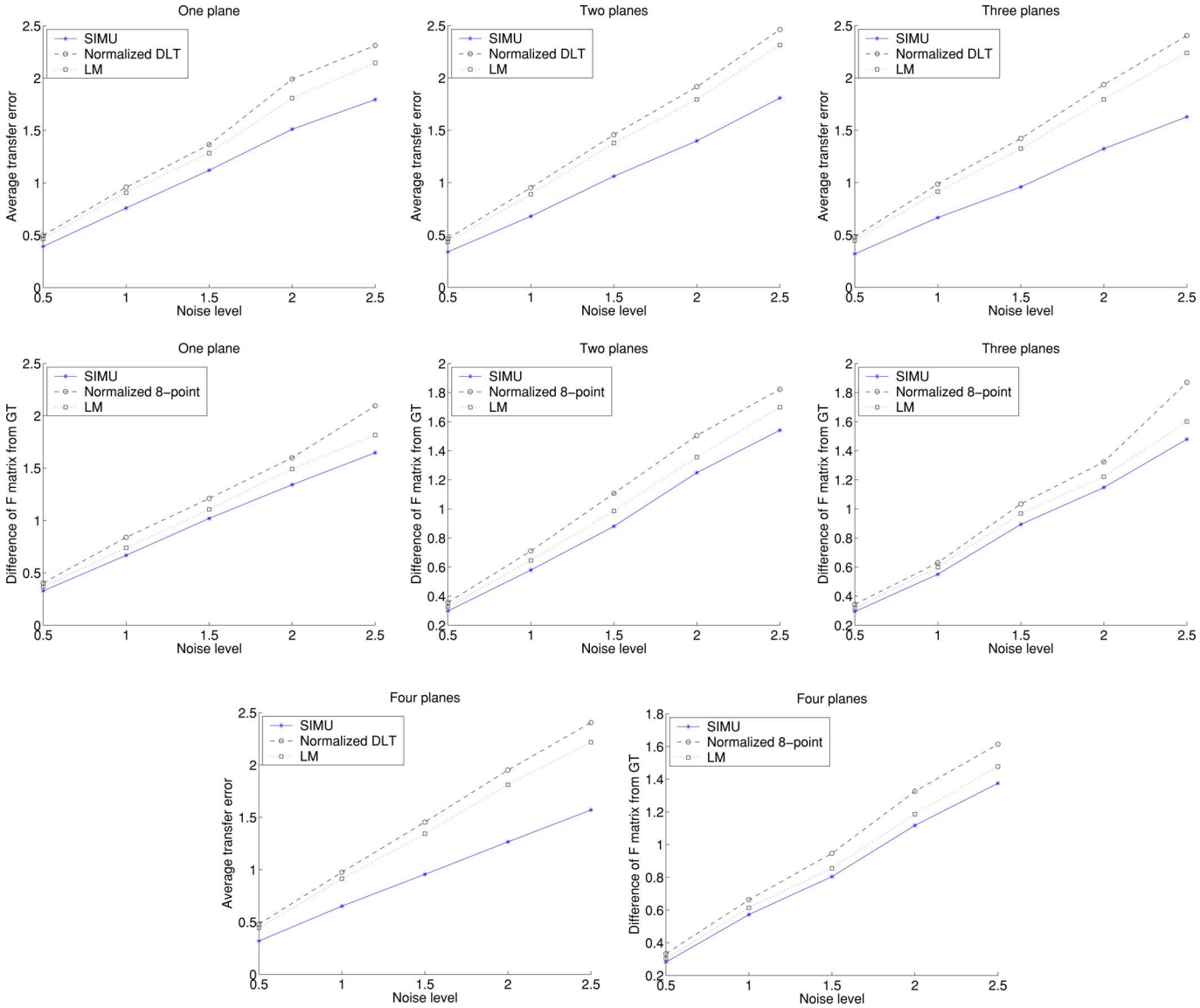

Fig. 3. (Top row) Errors of the estimated homographies and (middle) FM, where the first/second/third columns correspond to the cases of one/two/three planes, respectively. (Bottom) The four-plane case is listed as two subfigures. (Left) Homography. (Right) FM.

TABLE II
DISTANCES IN PIXELS BETWEEN THE ESTIMATED FMS AND THE CALIBRATED ("GROUND TRUTH") ONE FOR THE CASE BETWEEN THE FOURTH AND FIFTH VIEWS

	F_{SIMU}	F_{LM}	F_{NORM}
F_{CALI}	**0.7270**	**1.3982**	**3.5602**
F_{SIMU}		0.8251	4.1752
F_{LM}			4.3356

estimated epipoles (by the SIMU, LM, and NORM) from the calibrated epipole are 2.7301(4.8439), 3.0088(7.9934), and 3.9017(8.5121) pixels, respectively.

Thus, it can be seen that the SIMU performs the best on this corridor example, either evaluated by the distance (in pixels) between the estimated FMs and the calibrated FM or evaluated by the Euclidean distances between the estimated epipoles and the calibrated epipole.

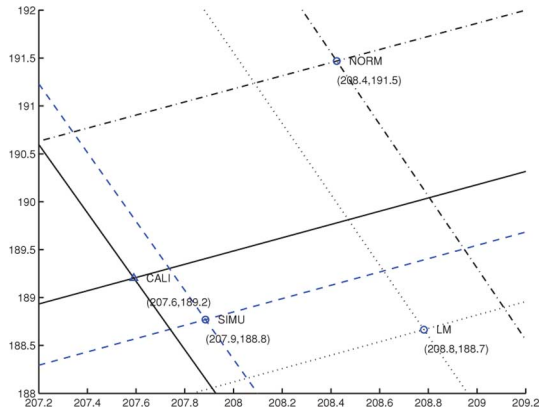

Fig. 4. Epipoles estimated by the three methods and the calibrated epipole for the case between the fourth and fifth views. (Note the different scales between the abscissa and ordinate axes.)

TABLE III
MEAN VALUES OF THE DISTANCES IN PIXELS BETWEEN THE ESTIMATED FMS AND THE CALIBRATED ("GROUND TRUTH") ONE FOR ALL 55 PAIRS. THEIR VARIANCES ARE IN THE BRACKETS

	F_{SIMU}	F_{LM}	F_{NORM}
F_{CALI}	**2.9246(4.8752)**	**3.5024(9.8456)**	**5.0040(14.4759)**
F_{SIMU}		1.2501(1.5879)	3.7667(5.9013)
F_{LM}			3.5436(5.2110)

IV. CONCLUSION

In this paper, we propose a method for simultaneously estimating the FM and homographies that are compatible with each other. Simulations and a real example show that the proposed algorithm produces better estimates, compared with the method of separately estimating them. The proposed method will be useful in robot tasks, particularly such as the homography-related tasks [2], [16].

APPENDIX A

SAMPSON ERROR FOR HOMOGRAPHY ESTIMATION

For a homography H, a pair of match $\mathbf{x}_i$ and $\mathbf{x}'_i$ on the associated plane satisfies the homography constraint $H\mathbf{x}_i \sim \mathbf{x}'_i$. In the form of a cross product, the homography constraint is

$$[\mathbf{x}'_i]_\times H\mathbf{x}_i = \mathbf{0}. \tag{19}$$

Furthermore, the homography constraint (19) can be expressed as, in Kronecker product

$$(\mathbf{x}_i^T \otimes [\mathbf{x}'_i]_\times)\ \mathrm{vec}(H) = \mathbf{0} \tag{20}$$

where $\otimes$ denotes the Kronecker product, and $\mathrm{vec}(H)$ denotes the column-first vectorization of the matrix H. Because the 3×9 matrix $\mathbf{x}_i^T \otimes [\mathbf{x}'_i]_\times$ is row dependent, the third equation is usually omitted. The homography constraint (19) becomes

$$(\mathbf{x}_i^T \otimes [\mathbf{x}'_i]_{2,\times})\mathrm{vec}(H) = \mathbf{0} \tag{21}$$

where $[\mathbf{x}]_{2,\times}$ denotes a 2×3 matrix, taking the first two rows of $[\mathbf{x}]_\times$.

When noisy points are observed, the homography constraint (19) or (21) does not exactly hold. The *algebraic error* is defined as a two-vector

$$\epsilon_i = (\mathbf{x}_i^T \otimes [\mathbf{x}'_i]_{2,\times})\ \mathrm{vec}(H) \tag{22}$$

J_i is a 2×4 matrix, which is defined as the partial derivative of ϵ_i, with respect to $\mathbf{x}_i$ and $\mathbf{x}'_i$ (exactly, the x and y coordinates of the 2-D inhomogeneous points, which are represented by $\mathbf{x}_i$ and $\mathbf{x}'_i$.)

The *Sampson error* is defined as

$$\epsilon_i^T (J_i J_i^T)^{-1} \epsilon_i \tag{23}$$

APPENDIX B

PROOF THAT THE HOMOGRAPHY CONSTRAINT IMPLIES THE EPIPOLAR CONSTRAINT WHEN USING THE PARAMETERIZATIONS IN (3, 4) OR (7, 9)

Here, we only prove the case of using the parameterizations in (3, 4) because the case of using (7, 9) is a special case of those in terms of (3, 4).

Suppose that F and H are represented as (3) and (4). From the homography constraint, we have

$$H\mathbf{x} \sim \mathbf{x}' \Rightarrow [\mathbf{x}']_\times H\mathbf{x} = 0 \Rightarrow [\mathbf{x}']_\times A\mathbf{x} = [\mathbf{x}']_\times \mathbf{e}'\bar{\mathbf{v}}^T\mathbf{x} \tag{24}$$

Substituting the last equality above into $\mathbf{x}'^T F\mathbf{x}$, we have

$$\begin{aligned}\mathbf{x}'^T F\mathbf{x} &= \mathbf{x}'^T [\mathbf{e}']_\times A\mathbf{x} = -\mathbf{e}'^T [\mathbf{x}']_\times A\mathbf{x} \\ &= -\mathbf{e}'^T [\mathbf{x}']_\times \mathbf{e}'\bar{\mathbf{v}}^T\mathbf{x} = \mathbf{e}'^T [\mathbf{e}']_\times \mathbf{x}'\bar{\mathbf{v}}^T\mathbf{x} \\ &= 0.\end{aligned} \tag{25}$$

This is the epipolar constraint. Note that in (25), we use the following properties: $\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_\times \mathbf{b} = -[\mathbf{b}]_\times \mathbf{a}$, $\mathbf{a} \times \mathbf{a} = 0$, and $[\mathbf{a}]_\times$ is a skew-symmetric matrix, where $\mathbf{a}$ and $\mathbf{b}$ are three vectors.

ACKNOWLEDGMENT

The authors would like to thank the Visual Geometry Group, Oxford University, Oxford, U.K., for providing the images and data of the Corridor sequence.

REFERENCES

[1] [Online]. Available: http://www.robots.ox.ac.uk/∼vgg/data.html
[2] J. Chen, W. E. Dixon, M. Dawson, and M. McIntyre, "Homography-based visual servo tracking control of a wheeled mobile robot," *IEEE Trans. Robot.*, vol. 22, no. 2, pp. 406–415, Apr. 2006.
[3] P. Chen and D. Suter, "Rank constraints for homographies over two views: Revisiting the rank four constraint," *Int. J. Comput. Vis.*, vol. 81, no. 2, pp. 201–225, 2009.
[4] W. Chojnacki, M. J. Brooks, A. van den Hengel, and D. Gawley, "On the fitting of surfaces to data with covariances," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1294–1303, Nov. 2000.
[5] O. D. Faugeras and Q. T. Luong, *The Geometry of Multiple Images: The Laws that Govern the Formation of Multiple Images of a Scene and Some of Their Applications*. Cambridge, MA: MIT Press, 2001.
[6] O. D. Faugeras, Q.-T. Luong, and S. J. Maybank, "Camera self-calibration: Theory and experiments," in *Proc. Eur. Conf. Comput. Vis.*, 1992, pp. 321–334.
[7] M. A. Fischler and R. C. Rolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[8] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
[9] R. I. Hartley, "In defense of the eight-point algorithm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 6, pp. 580–593, Jun 1997.
[10] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision,* 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[11] T. S. Huang and O. D. Faugeras, "Some properties of the e matrix in two-view motion estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 12, pp. 1310–1312, Dec. 1989.
[12] R. Kaucic, R. I. Hartley, and N. Y. Dano, "Plane-based projective reconstruction," in *Proc. Int. Conf. Comput. Vis.*, 2001, pp. 420–427.
[13] H. C. Longuet Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, vol. 293, pp. 133–135, 1981.
[14] Q. T. Luong and T. Vieville, "Canonical representations for the geometries of multiple projective views," *Comput. Vis. Image Understand.*, vol. 64, no. 2, pp. 193–229, 1996.
[15] Y. Ma, S. Soatto, J. Kosecka, and S. Sastry, "An invitation to 3-D vision: From images to geometric models," in *Interdisciplinary Applied Mathematics*. New York: Springer-Verlag, 2003.
[16] H. Malm and A. Heyden, "Extensions of plane-based calibration to the case of translational motion in a robot vision setting," *IEEE Trans. Robot.*, vol. 22, no. 2, pp. 322–333, Apr. 2006.
[17] J. I. Mulero-Martinez, "Uniform bounds of the coriolis/centripetal matrix of serial robot manipulators," *IEEE Trans. Robot.*, vol. 23, no. 5, pp. 1083–1089, Oct. 2007.
[18] C. Rother and S. Carlsson, "Linear multi view reconstruction and camera recovery," in *Proc. Int. Conf. Comput. Vis.*, 2001, pp. 42–51.
[19] C. Rother and S. Carlsson, "Linear multi view reconstruction and camera recovery using a reference plane," *Int. J. Comput. Vis.*, vol. 49, no. 2, pp. 117–141, 2002.
[20] P. D. Sampson, "Fitting conic sections to 'very scattered' data: An iterative refinement of the Bookstein algorithm," *Comput. Graph. Image Process.*, vol. 18, no. 1, pp. 97–108, 1982.

[21] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon, "Bundle adjustment—A modern synthesis," in *Vision Algorithms: Theory and Practice*, W. Triggs, A. Zisserman, and R. Szeliski, Eds. New York: Springer-Verlag, Lecture Notes in Computer Science, 2000, pp. 298–375.
[22] J. Weng, T. S. Huang, and N. Ahuja, "Motion and structure from two perspective views: Algorithms, error analysis, and error estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 5, pp. 451–476, May 1989.
[23] Z. Zhang, "Determining the epipolar geometry and its uncertainty: A review," *Int. J. Comput. Vis.*, vol. 27, no. 2, pp. 161–195, 1998.
[24] Z. Zhang, "On the optimization criteria used in two-view motion analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 7, pp. 717–729, Jul. 1998.
[25] Z. Zhang, R. Deriche, O. Faugeras, and Q. T. Luong, "A robust technique for matching two uncalibrated image through the recovery of the unknown epipolar geometry," *Artif. Intell.*, vol. 78, pp. 87–119, 1995.

Vision-Based Localization for Leader–Follower Formation Control

Gian Luca Mariottini, *Member, IEEE*,
Fabio Morbidi, *Student Member, IEEE*,
Domenico Prattichizzo, *Member, IEEE*,
Nicholas Vander Valk, *Student Member, IEEE*,
Nathan Michael, *Student Member, IEEE*,
George Pappas, *Fellow, IEEE*,
and Kostas Daniilidis, *Senior Member, IEEE*

***Abstract*—This paper deals with vision-based localization for leader–follower formation control. Each unicycle robot is equipped with a panoramic camera that only provides the view angle to the other robots. The localization problem is studied using a new observability condition valid for general nonlinear systems and based on the extended output Jacobian. This allows us to identify those robot motions that preserve the system observability and those that render it nonobservable. The state of the leader–follower system is estimated via the extended Kalman filter, and an input-state feedback control law is designed to stabilize the formation. Simulations and real-data experiments confirm the theoretical results and show the effectiveness of the proposed formation control.**

***Index Terms*—Feedback linearization, formation control, mobile robots, nonlinear observability, panoramic cameras.**

I. Introduction

A. Motivation and Related Work

A growing interest in coordination and control of multiple autonomous agents has matured over the past few years [7], [12], [22], [27], [28]. The formation control problem has been playing an important role in this research area, giving rise to a rich literature [2], [15], [17], [30], [31]. By *formation control*, we simply mean the problem of controlling the relative position and orientation of group of robots while allowing the group to move as a whole. In the *leader–follower* formation control approach, a robot, i.e., the leader, moves along a predefined trajectory, while the other robots, i.e., the followers, are supposed to

Manuscript received August 28, 2008; revised January 19, 2009, May 28, 2009, and August 22, 2009. First published October 30, 2009; current version published December 8, 2009. This paper was recommended for publication by Associate Editor D. Song and Editor L. Parker upon evaluation of the reviewers' comments. This work was supported by the National Science Foundation-Environmental Industry Associations under Grant 0324977, the National Science Foundation-Information and Intelligent System under Grant 0713260, the National Science Foundation-Industrial Innovation and Partnerships under Grant 0742304, the Department of Defense Multidisciplinary University Research Initiative under Grant 19-02-1-0383, and by the Fondazione Monte dei Paschi di Siena 2005.

G. L. Mariottini is with the Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: gianluca@cs.umn.edu).

F. Morbidi and D. Prattichizzo are with the Dipartimento di Ingegneria dell'Informazione, University of Siena, Siena 53100, Italy.

N. V. Valk, N. Michael, G. Pappas, and K. Daniilidis are with the General Robotics, Automation, Sensing and Perception Laboratory, University of Pennsylvania, Philadelphia, PA 19104 USA.

This paper has supplementary downloadable multimedia material available at http://ieeexplore.ieee.org provided by the author.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TRO.2009.2032975