# A Revisit of Methods for Determining the Fundamental Matrix with Planes

Yi Zhou[1,2], Laurent Kneip[1,2], and Hongdong Li[1,2,3]

[1]Research School of Engineering, Australian National University
[2]ARC Centre of Excellence for Robotic Vision
[3]NICTA Canberra Labs
{yi.zhou, laurent.kneip, hongdong.li}@anu.edu.au

***Abstract*—Determining the fundamental matrix from a collection of inter-frame homographies (more than two) is a classical problem. The compatibility relationship between the fundamental matrix and any of the ideally consistent homographies can be used to compute the fundamental matrix. Using the direct linear transformation (DLT), the compatibility equation can be translated into a least squares problem and can be easily solved via SVD decomposition. However, this solution is extremely susceptible to noise and motion inconsistencies, hence rarely used. Inspired by the normalized eight-point algorithm, we show that a relatively simple but non-trivial two-step normalization of the input homographies achieves the desired effect, and the results are at last comparable to the less attractive hallucinated points method. The algorithm is theoretically justified and verified by experiments on both synthetic and real data.**

## I. Introduction

The epipolar geometry of two perspective images can be described by a singular $3 \times 3$ matrix. When the camera is calibrated, the matrix is known as the essential matrix $\mathbf{E}$. For uncalibrated systems, it is known as the fundamental matrix $\mathbf{F}$. The estimation of the fundamental matrix is a classical and thoroughly studied topic which plays an essential role in many applications involving multiple-view geometry, such as visual odometry (VO), structure from motion (SfM), and visual SLAM, etc.

The most popular method for estimating the fundamental matrix is based on sparse correspondences between local invariant keypoints, for instance given by the popular SIFT algorithm [12]. Seven points constitute the minimal configuration because the fundamental matrix has 7 degrees of freedom (DoF). Compared to the eight-point algorithm, the seven point algorithm needs an additional step to calculate the linear combination factor of the obtained two-dimensional null-space. While seven point correspondences represent the minimum for estimating the fundamental matrix [18], the 8-point algorithm [11] is the most popular method because of its linear nature and thus simplicity to implement. However, it was only after Hartley published his seminal work [10] on using data normalization that the eight-point algorithm became truly useful in practice.

It is believed that the reconstruction performance can be improved by incorporating additional geometric constraints like coplanarity of certain points. Luong and Faugeras [13, 14] are the first who propose to estimate the fundamental matrix with multiple homographies in a linear way. They compared the linear solution with other non-linear ones concluding that none of the developed methods is stable under noise. In other words, though the direct linear method is quite simple and straightforward, it has limited practical usefulness. Zhang [24] gave a thorough review on the techniques of fundamental matrix estimation and its uncertainty. The bad performance of the Direct Linear Transformation (DLT) applied to the compatibility relation between the homography $\mathbf{H}$ and the fundamental matrix $\mathbf{F}$ was however not discussed in much detail. Szeliski and Torr [19] thoroughly discussed three methods used for solving structure from motion (SfM) with planes. They presented an analysis of the robustness of each method and then suggested to estimate the fundamental matrix with hallucinated points (HP) that lie on planes instead of the compatibility equation (and thus the homograhies directly). Anubhav *et al.* [1] demonstrates that the compatibility constraint is an implicit equation in $\mathbf{H}$ and $\mathbf{F}$. They also concluded that an explicit expression like $\mathbf{F} = [\mathbf{e}']_{\times}\mathbf{H}$ is more suitable for a computational algorithm. Vincent and Laganiere [21] proposed a detection algorithm for planar homographies working on a pair of uncalibrated images. They claimed that the estimation of the fundamental matrix from point correspondences derived from homographies allows to use data normalization techniques, and thus performs much better than using the homographies directly. A method was introduced to estimate the fundamental matrix with a homology in [15, 17, 9]. Theoretically, a homology has two identical eigenvalues and another unique one which is corresponding to the epipole $\mathbf{e}'$. However, in practical situations, the homographies are never perfect, which is why a double eigenvalue is never guaranteed. It is hard in practice to choose which eigenvalue corresponds to the unique one; the real parts of the eigenvalues are often equally spread and/or very close to each other.

The goal of this paper is to present a direct method for computing the fundamental matrix from a set of homographies estimated independently between two view-points of a rigid, piece-wise planar scene. Key to our method is a two-step normalization procedure leading to a rescaled linear solution of the compatibility equation. We finally achieve comparable results to the hallucinated points method, however without

introducing additional, virtual correspondences. The rest of the paper is organized as follows. Section 2 quickly reviews the three methods discussed in [19]. Our two-step linear (TSL) method with all theoretical derivations is described in Section 3. In Section 4, we compare DLT, HP and TSL by separately running them on synthetic and real data. An analysis of numerical stability and algorithmic complexity is also given.

## II. Quick Review Of the Three Methods

Szeliski and Torr discussed three methods that can be used for estimating the fundamental matrix given several ($\geqslant 2$) homographies in [19], which are reviewed in the following.

- **Hallucinating additional correspondences:**
  Hallucinated points refer to augmented sample points on planes. Theses points are also called virtual control points. Hallucinated correspondences are generated by first creating several virtual 2D points $\mathbf{x}$ on image one which are assumed to be the projection of virtual points on the plane. Their corresponding points $\mathbf{x}'$ are then found by applying the corresponding homography to points $\mathbf{x}$. Then the fundamental matrix $\mathbf{F}$ is computed by applying normalized 8-point algorithm on the obtained hallucinated correspondences.
- **Direct linear method:**
  The implicit compatibility relationship between inter-frame homographies and the fundamental matrix can be directly used for computing the fundamental matrix. The compatibility equation $\mathbf{F^T H + H^T F = 0}$ gives six constraints [13] (for which only 5 are linearly independent). Therefore, at least 2 homographies are needed for computing the fundamental matrix. The question can be translated to a least squares problem by DLT and can be easily solved by SVD decomposition. However, this straightforward method is unstable for inaccurate homographies, sometimes leading to completely meaningless results. The reason given by Szeliski and Torr is that using the compatibility equation directly corresponds to sampling homographies at locations where their predictive power is very weak. The samples are far from having the normal distribution required for total least squares to work reasonably well.
- **Plane plus parallax:**
  Plane plus parallax techniques are always used to recover the depth (projective or Euclidean) of the scene. To compute the fundamental matrix, one of the homographies is choosen and used to unwarp all points in the current frame. The epipole $\mathbf{e}'$ is computed by minimizing the sum of the weighted distance between epipole and lines passing through corresponding points $\mathbf{x_i}$ and $\mathbf{x'_i}$. Then the fundamental matrix $\mathbf{F}$ can be computed by $\mathbf{F} = [\mathbf{e}']_\times \mathbf{H}$. This method cannot work well when points are evenly distributed over several planes. The computation is also more complicated and expensive compared to the former two methods.

## III. Robust Two-Step Linear Solution

The compatibility equation $\mathbf{F^T H + H^T F = 0}$ gives only 6 linear equations [13]. In fact, as shown later, only 5 of them are independent. Therefore, at least 2 homographies are needed for computing the fundamental matrix. Applying the DLT transformation to the compatibility equation leads to the least squares problem,

$$\mathbf{Af} = \begin{pmatrix} \mathbf{W_1} \\ \mathbf{W_2} \\ \vdots \\ \mathbf{W_n} \end{pmatrix} \mathbf{f} = \mathbf{0}, \tag{1}$$

where $\mathbf{f} = (f_{11}, f_{21}, f_{31}, f_{12}, f_{22}, f_{32}, f_{13}, f_{23}, f_{33})^T$ denotes a vector obtained by rearranging the entries of the fundamental matrix in a column vector. Matrix $\mathbf{A}$ is made up of several sub matrices $\mathbf{W_i}$ of same dimension which is defined as,

$$\mathbf{W_i} = \begin{pmatrix} 2h_{11}^{\pi_i} & 0 & 0 & 2h_{21}^{\pi_i} & 0 & 0 & 2h_{31}^{\pi_i} & 0 & 0 \\ h_{12}^{\pi_i} & h_{11}^{\pi_i} & 0 & h_{22}^{\pi_i} & h_{21}^{\pi_i} & 0 & h_{32}^{\pi_i} & h_{31}^{\pi_i} & 0 \\ h_{13}^{\pi_i} & 0 & h_{11}^{\pi_i} & h_{23}^{\pi_i} & 0 & h_{21}^{\pi_i} & h_{33}^{\pi_i} & 0 & h_{31}^{\pi_i} \\ 0 & 2h_{12}^{\pi_i} & 0 & 0 & 2h_{22}^{\pi_i} & 0 & 0 & 2h_{32}^{\pi_i} & 0 \\ 0 & h_{13}^{\pi_i} & h_{12}^{\pi_i} & 0 & h_{23}^{\pi_i} & h_{22}^{\pi_i} & 0 & h_{33}^{\pi_i} & h_{32}^{\pi_i} \\ 0 & 0 & 2h_{13} & 0 & 0 & 2h_{23}^{\pi_i} & 0 & 0 & 2h_{33}^{\pi_i} \end{pmatrix}. \tag{2}$$

The entries of the matrix $\mathbf{W_i}$ originate from the homography $\mathbf{H_i} = \begin{pmatrix} h_{11}^{\pi_i} & h_{12}^{\pi_i} & h_{13}^{\pi_i} \\ h_{21}^{\pi_i} & h_{22}^{\pi_i} & h_{23}^{\pi_i} \\ h_{31}^{\pi_i} & h_{32}^{\pi_i} & h_{33}^{\pi_i} \end{pmatrix}$ which is induced by plane $\pi_i$. As discussed in Section 2, the least squares problem described in Eq. (1) is seriously ill-conditioned, which means that even under a tiny perturbation of any entry of matrix $\mathbf{A}$, the solution quickly diverges from the groundtruth result. Thus, the matrix $\mathbf{A}$ should be re-conditioned in order to stabilize its null space.

We follow the idea of [10] and introduce normalization in order to stabilize the result. However, it is not trivial to directly normalize the matrix $\mathbf{A}$ as it has been done in prior work for estimating the fundamental matrix or even the homography from point correspondences. The reason is two-fold. First, the normalization includes two parts, translation and scaling. The translation operation can only be performed by a linear transformation when the normalized object is described in homogeneous form. Second, the normalization should be performed to data which has the same meaning.

The key to deal with above two issues comes from the special structure of the matrix $\mathbf{F^T H}$. The compatibility equation requires that $\mathbf{F^T H}$ is a skew-symmetric matrix, and thus is of the form

$$\mathbf{F}^T\mathbf{H} = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}. \tag{3}$$

The diagonal entries give three equations which describe an orthogonal relationship between corresponding column vectors of the fundamental matrix and a homography,

$$\mathbf{f_i^T h_i} = 0, \mathrm{i} = 1, 2, 3. \tag{4}$$

$\mathbf{f_i}$ and $\mathbf{h_i}$ denote the $\mathrm{i_{th}}$ column vector of the fundamental matrix $\mathbf{F} = \begin{pmatrix} \mathbf{f_1} & \mathbf{f_2} & \mathbf{f_3} \end{pmatrix}$ and the homography $\mathbf{H} =$

$\begin{pmatrix}\mathbf{h_1} & \mathbf{h_2} & \mathbf{h_3}\end{pmatrix}$. The other three equations enforce the skew symmetric property. However, only two of them are independent. This makes sense because a homography has 8 degrees of freedom (DoF). For the uncalibrated case, the intrinsic matrix is unknown which removes three constraints. Thus, only five independent constraints can be obtained from one homography, three from the orthogonal relationship described in Eq. (4) and the other two from the skew-symmetric property.

Our two-step reconditioning method realizes the non-trivial normalization by fully using the special structure of matrix $\mathbf{F^T H}$. First, by utilizing the orthogonal relationship, we decompose the original least squares problem $\mathbf{Af} = \mathbf{0}$ into three sub least squares problems $\mathbf{A_i f_i} = \mathbf{0}$, where matrix $\mathbf{A_i} = \begin{pmatrix}\mathbf{h_i^{\pi_1}} & \mathbf{h_i^{\pi_2}} & \cdots & \mathbf{h_i^{\pi_n}}\end{pmatrix}^T$ and $\mathrm{i} = 1, 2, 3$. Each column $\mathbf{f_i}$ of $\mathbf{F}$ is estimated individually. The relative scale factor for each estimated solution $\mathbf{f_i}$ can then be recovered by using the skew-symmetric property of matrix $\mathbf{F^T H}$ in Eq. (3). With this formulation, every column of matrix $\mathbf{A_i}$ has the same meaning. Besides, in order to do the translation, the matrix $\mathbf{A_i}$ should be extended by an additional column $\mathbf{1}_{n\times 1} = (1 \quad 1 \quad \dots \quad 1)^T$ which leads to $\tilde{\mathbf{A}}_\mathbf{i} = [\mathbf{A_i}|\mathbf{1}_{n\times 1}]$. Accordingly, the extended solution vector $\tilde{\mathbf{f}}_\mathbf{i}$ is defined as $\tilde{\mathbf{f}}_\mathbf{i} = \begin{pmatrix}\lambda_i^{-1}\mathbf{f_i}\\ 0\end{pmatrix}$, where $\lambda$ denotes the relative scale factor of the individually estimated solution. The mathematical proof is given after the whole algorithm is introduced. This extension turns each row of matrix $\mathbf{A_i}$ into homogeneous form.

The normalization is then performed by inserting a $4 \times 4$ linear transformation matrix $\mathbf{Q_i}$ and its inverse in between $\tilde{\mathbf{A}}_\mathbf{i}$ and $\tilde{\mathbf{f}}_\mathbf{i}$, resulting in

$$\tilde{\mathbf{A}}_\mathbf{i}\mathbf{Q_i}\mathbf{Q_i^{-1}}\tilde{\mathbf{f}}_\mathbf{i} = \hat{\mathbf{A}}_\mathbf{i}\hat{\mathbf{f}}_\mathbf{i} = \mathbf{0}, \tag{5}$$

where $\hat{\mathbf{A}}_\mathbf{i} = \tilde{\mathbf{A}}_\mathbf{i}\mathbf{Q_i}$ and $\hat{\mathbf{f}}_\mathbf{i} = \mathbf{Q_i^{-1}}\tilde{\mathbf{f}}_\mathbf{i}$. The linear transformation $\mathbf{Q_i}$ includes a translation and a scaling. We regard each $\mathbf{h_i^{\pi_j}}$ as a 3d point. Following the idea of [10], the coordinates are translated such that the centroid $\mathbf{c}$ of the set of all such points becomes the origin.The coordinates are then scaled by applying an isotropic scaling factor $s$ to all three coordinates of each point. Finally, we choose to scale the coordinates such that the average distance of a point $\mathbf{h_i^{\pi_j}}$ from the origin is equal to $\sqrt{3}$. The linear transformation $\mathbf{Q_i}$ and scaling related variables are defined as below.

$$\mathbf{Q_i} = \begin{pmatrix} s & 0 & 0 & 0\\ 0 & s & 0 & 0\\ 0 & 0 & s & 0\\ -c_1 s & -c_2 s & -c_3 s & 1\end{pmatrix}. \tag{6}$$

$$\mathbf{c} = \begin{pmatrix}c_1 & c_2 & c_3\end{pmatrix}^T = \frac{\sum_{j=1}^{m}\mathbf{h_i^{\pi_j}}}{m} \tag{7}$$

$$s = \frac{\sqrt{3}}{\bar{d}}, \bar{d} = \frac{\sum_{j=1}^{m}\|\mathbf{h_i^{\pi_j}} - \mathbf{c}\|_F}{m} \tag{8}$$

The solution of the three sub least square problems $\hat{\mathbf{A}}_\mathbf{i}\hat{\mathbf{f}}_\mathbf{i} = \mathbf{0}$ can be easily obtained via SVD. Then $\tilde{\mathbf{f}}_\mathbf{i} = \mathbf{Q_i}\hat{\mathbf{f}}_\mathbf{i}$. The only remaining task is to find the scale factor $\lambda_i$.

The skew-symmetric property of matrix $\mathbf{F^T H}$ can be translated into another least squares problem $\mathbf{A}_\lambda\boldsymbol{\lambda} = \mathbf{0}$ via DLT, where $\boldsymbol{\lambda} = \begin{pmatrix}\lambda_1 & \lambda_2 & \lambda_3\end{pmatrix}^T$ and $\mathbf{A}_\lambda$ is given by

$$\mathbf{A}_\lambda = \begin{pmatrix} \tilde{\mathbf{f}}_{\mathbf{1,1:3}}^\mathbf{T}\mathbf{h_2^{\pi_1}} & \tilde{\mathbf{f}}_{\mathbf{2,1:3}}^\mathbf{T}\mathbf{h_1^{\pi_1}} & 0\\ \tilde{\mathbf{f}}_{\mathbf{1,1:3}}^\mathbf{T}\mathbf{h_3^{\pi_1}} & 0 & \tilde{\mathbf{f}}_{\mathbf{3,1:3}}^\mathbf{T}\mathbf{h_1^{\pi_1}}\\ 0 & \tilde{\mathbf{f}}_{\mathbf{2,1:3}}^\mathbf{T}\mathbf{h_3^{\pi_1}} & \tilde{\mathbf{f}}_{\mathbf{3,1:3}}^\mathbf{T}\mathbf{h_2^{\pi_1}}\\ \vdots & \vdots & \vdots\\ \tilde{\mathbf{f}}_{\mathbf{1,1:3}}^\mathbf{T}\mathbf{h_2^{\pi_m}} & \tilde{\mathbf{f}}_{\mathbf{2,1:3}}^\mathbf{T}\mathbf{h_1^{\pi_m}} & 0\\ \tilde{\mathbf{f}}_{\mathbf{1,1:3}}^\mathbf{T}\mathbf{h_3^{\pi_m}} & 0 & \tilde{\mathbf{f}}_{\mathbf{3,1:3}}^\mathbf{T}\mathbf{h_1^{\pi_m}}\\ 0 & \tilde{\mathbf{f}}_{\mathbf{2,1:3}}^\mathbf{T}\mathbf{h_3^{\pi_m}} & \tilde{\mathbf{f}}_{\mathbf{3,1:3}}^\mathbf{T}\mathbf{h_2^{\pi_m}}\end{pmatrix}. \tag{9}$$

$\tilde{\mathbf{f}}_{\mathbf{i,1:3}}$ in $\mathbf{A}_\lambda$ is defined as the first three rows of vector $\tilde{\mathbf{f}}_\mathbf{i}$. $\mathbf{h_i^{\pi_j}}$ is defined same as before. The full two-step linear method (TSL) is described in Algorithm 1.

**Algorithm 1** Two-Step Linear Method (TSL)

1: Input: A collection of independently estimated homographies $\mathbf{H_s}$
2: **for** i = 1:3 **do**
3: $\quad\tilde{\mathbf{A}}_\mathbf{i} = [\mathbf{A_i}|\mathbf{1}_{n\times 1}]$
4: $\quad\hat{\mathbf{A}}_\mathbf{i} \leftarrow \tilde{\mathbf{A}}_\mathbf{i}\mathbf{Q_i}$
5: $\quad\hat{\mathbf{f}}_\mathbf{i} \leftarrow solve\hat{\mathbf{A}}_\mathbf{i}\hat{\mathbf{f}}_\mathbf{i} = \mathbf{0}$
6: $\quad\tilde{\mathbf{f}}_\mathbf{i} \leftarrow \mathbf{Q_i}\hat{\mathbf{f}}_\mathbf{i}$
7: **end for**
8: $\boldsymbol{\lambda} = \begin{pmatrix}\lambda_1 & \lambda_2 & \lambda_3\end{pmatrix}^T \leftarrow solve\mathbf{A}_\lambda\lambda = \mathbf{0}$
9: Output: $\mathbf{F} = \begin{pmatrix}\mathbf{f_1} & \mathbf{f_2} & \mathbf{f_3}\end{pmatrix}$

It should be noted that in order to apply the normalization, the original least squares problem is modified. However, we will see in the following that solving the modified problem $\tilde{\mathbf{A}}_\mathbf{i}\tilde{\mathbf{f}}_\mathbf{i} = \mathbf{0}$ is equivalent to solving the original problem $\mathbf{A_i f_i} = \mathbf{0}$. Therefore, two questions need to be answered in order to prove this claim:

1) After extending the matrix $\mathbf{A_i}$ by an additional column $\mathbf{1}_{n\times 1} = (1 \quad 1 \quad \dots \quad 1)^T$, what is the null-space configuration of $\tilde{\mathbf{A}}_\mathbf{i}$?
2) Why does the solution of problem $\tilde{\mathbf{A}}_\mathbf{i}\tilde{\mathbf{f}}_\mathbf{i} = \mathbf{0}$ have the structure as $\tilde{\mathbf{f}}_\mathbf{i} = \begin{pmatrix}\lambda_i^{-1}\mathbf{f_i}\\ 0\end{pmatrix}$?

The answers are given by proving the following two claims:

- **Property 1**
  $Rank(\mathbf{A_i}) = 2$, $1 \le dim(N(\tilde{\mathbf{A}}_\mathbf{i})) \le 2$ when the number of planes $m \ge 3$, , where $\mathbf{N}(\cdot)$ denotes the null space of $(\cdot)$
- **Property 2**
  $\mathbf{N}(\tilde{\mathbf{A}}_\mathbf{i}) = \begin{pmatrix}\mathbf{N}(\mathbf{A}_i)\\ 0\end{pmatrix}$.

*Proof.* **Property 1**

Assuming that two camera matrices are given by $\mathbf{P_1} = [\mathbf{I_{3\times 3}}|\mathbf{0_{3\times 1}}]$ and $\mathbf{P_2} = [\mathbf{B}|\mathbf{b}]$, each homography induced by a plane $\pi_\mathbf{j} = [-\mathbf{v_j^T}, 1]$ observed by the two cameras can be denoted as

$$\mathbf{H}^{\pi_\mathbf{j}} \simeq \mathbf{B} + \mathbf{b}\mathbf{v_j^T}. \tag{10}$$

Each row of matrix $\mathbf{A_i}$ contains the $\mathbf{i_{th}}$ column of one homography, which gives

$$\mathbf{h_i^{\pi_j}} \simeq \mathbf{B_i} + \mathbf{v_{j,i}b}, \tag{11}$$

where $\mathbf{B_i}$ denotes the $\mathbf{i_{th}}$ column of the matrix $\mathbf{B}$ and $\mathbf{v_{j,i}}$ the $\mathbf{i_{th}}$ element of the vector $\mathbf{v_j}$. It is obvious to see that if we regard each row of matrix $\mathbf{A_i}$ as a general 3D point, all the points $\mathbf{h_i^{\pi_j}}$ are lying on the line with the direction of $\mathbf{v_{j,i}b}$ passing point $\mathbf{B_i}$. Thus $Rank(\mathbf{A_i}) = 2$.

Since matrix $\tilde{\mathbf{A}}_\mathbf{i}$ is obtained by adding an additional column $\mathbf{1}_{3\times1}$ to $\mathbf{A_i}$, it is also obvious to see that

$$Rank(\mathbf{A_i}) \leq Rank(\tilde{\mathbf{A}}_\mathbf{i}) \leq 3. \tag{12}$$

Because

$$Rank(\tilde{\mathbf{A}}_\mathbf{i}) + dim(N(\tilde{\mathbf{A}}_\mathbf{i})) = 4, \tag{13}$$

thus we finally have

$$1 \leq dim(N(\tilde{\mathbf{A}}_\mathbf{i})) \leq 2.^1 \tag{14}$$

□

*Proof.* **Property 2**

Assuming $\mathbf{x} \in \mathbf{N}(\mathbf{A_i}), and\ \tilde{\mathbf{x}} \in \mathbf{N}(\tilde{\mathbf{A}}_\mathbf{i})$, we have $\mathbf{A_i x} = \mathbf{0}, and\ \tilde{\mathbf{A}}_\mathbf{i}\tilde{\mathbf{x}} = \mathbf{0}$.

Obvioulsy, $\forall \mathbf{x} \in \mathbf{N}(\mathbf{A_i}), \tilde{\mathbf{A}}_\mathbf{i} \begin{pmatrix}\mathbf{x}\\0\end{pmatrix} = [\mathbf{A_i}|\mathbf{1}_{n\times1}] \begin{pmatrix}\mathbf{x}\\0\end{pmatrix} = \mathbf{0}$.

Thus, $\begin{pmatrix}\mathbf{N(A_i)}\\0\end{pmatrix} \in \mathbf{N}(\tilde{\mathbf{A}}_\mathbf{i})$.

Necessary condition Q.E.D.

On the other hand, assume $\forall \tilde{\mathbf{x}} = \begin{pmatrix}\mathbf{x}\\\omega\end{pmatrix}, \omega \neq 0$.

Since $\tilde{\mathbf{A}}_\mathbf{i}\tilde{\mathbf{x}} = \mathbf{0}$,

$\Rightarrow\ \mathbf{A_i x} + \omega\mathbf{1}_{n\times1} = \mathbf{0}$,

$\Rightarrow\ \omega\mathbf{1}_{n\times1} = \mathbf{0}$,

$\Rightarrow\ \omega = \mathbf{0}$,

$\Rightarrow$ contradiction,

Thus, $\mathbf{N}(\tilde{\mathbf{A}}_\mathbf{i}) \in \begin{pmatrix}\mathbf{N(A_i)}\\0\end{pmatrix}$.

Sufficient condition Q.E.D.

Summarizing, $\mathbf{N}(\tilde{\mathbf{A}}_\mathbf{i}) = \begin{pmatrix}\mathbf{N}(\mathbf{A}_i)\\0\end{pmatrix}$. □

Above mathematical proofs explain why we can get the solution to the original problem by solving the reconditioned least squares problems. One drawback of the proposed method is that at least 3 planes (homographies) are needed for computing the fundamental matrix. Otherwise additional information is needed for determining the exact null space.

[1]If $Rank(\tilde{\mathbf{A}}_\mathbf{i}) = 3$, $\tilde{\mathbf{A}}_\mathbf{i}$ has only a one dimensional null space which is the eigen vector corresponding to the smallest eigenvalue of matrix $\tilde{\mathbf{A}}_\mathbf{i}$. Otherwise, if $Rank(\tilde{\mathbf{A}}_\mathbf{i}) = 2$, the final solution of problem $\tilde{\mathbf{A}}_\mathbf{i}\tilde{\mathbf{f}}_\mathbf{i}$ resides in a two dimensional null space. However, during our experiment, we never observed the case of $Rank(\tilde{\mathbf{A}}_\mathbf{i}) = 2$.

## IV. Experimental Evaluation

In this section, we compare the performance of DLT, HP and TSL on both synthetic and real data. Numerical stability of DLT and TSL as well as algorithmic complexity of the three methods are also discussed.

The input homographies can be derived from either point or line features as they are dual geometric entities [6, 7, 4]. We use line features during the synthetic experiments, and point correspondences during the experiment on real data.

### A. Synthetic experiment

For each single experiment, we construct two artificial views observing planes in a 3D environment. Groundtruth motion and structure (planes) is generated in the same way as in [20]. Without loss of generality, the camera pose of the first view is assumed to be identical with the world frame. The absolute pose of the second view is defined by motion parameters lying within a certain range. The rotation angles along each axis (roll, pitch, yaw) lie within $(-5°, 5°)$ and the translation in each direction (X, Y, Z) is within $(-100, 100)$. The structure is randomly generated by creating $N = 5$ planes with known homographies. 4 groups of Gaussian noise ($\mu = 0, \sigma \in [0, 0.5]$) corrupted points[2] are created on each plane, which are used for fitting the line features. The image size is $640\times480$ and the focal length is $f_x = f_y = 250$. The relative motion parameters are extracted from the estimated fundamental matrix $\mathbf{F}$ (in fact from essential matrix $\mathbf{E} = \mathbf{K^T F K}$).

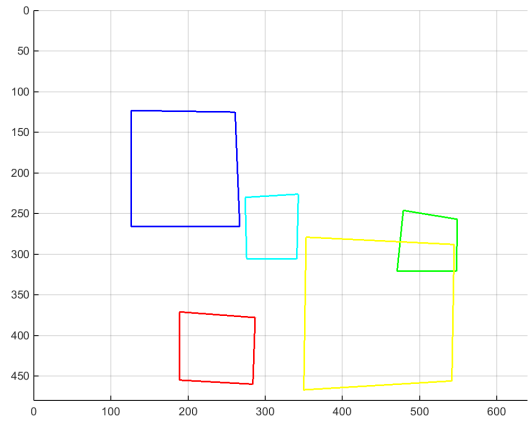
As shown in Fig 1, both HP and TSL outperform DLT in the accuracy of the estimated fundamental matrix and the motion parameters. We use max norm of the difference between $\mathbf{F_{groundtruth}}$ and $\mathbf{F_{estimated}}$ as a criterion for assessing the accuracy of the estimated $\mathbf{F}$. The estimated rotation matrix is compared to groundtruth by computing the anlge $\Theta = arccos(\frac{trace(\mathbf{R^T_{groundtruth}R_{estimated}})-1}{2})$. And the estimated translation is compared against groundtruth by computing the angle between two translation vectors $\mathbf{t_{groundtruth}}$ and $\mathbf{t_{estimated}}$. TSL is more noise resilient in terms of fundamental matrix estimation in comparison to HP. Concerning the accuracy of the extracted motion parameters, TSL and HP perform equally well.

### B. Experiment on real images

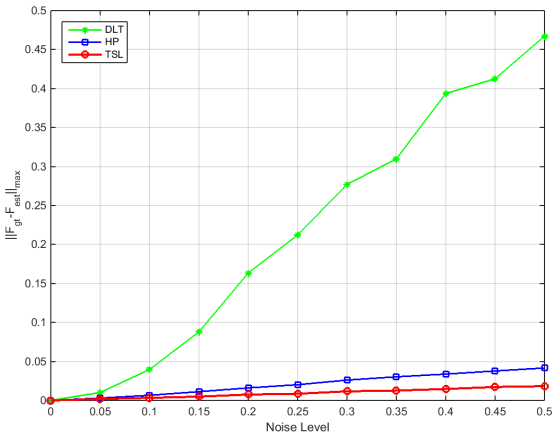
The algorithm is tested on the famous Oxford Corridor sequence. Homographies are estimated from Harris corner correspondences [8]. Points on each plane are grouped manually and outliers are rejected by applying the Random sample consensus (RANSAC) technique [5].

As shown in Fig. 2, the epipole estimated by TSL (i.e. the intersection point of blue lines) is closest to groundtuth. The epipole $\mathbf{e}$ is extracted from the null space of the fundamental

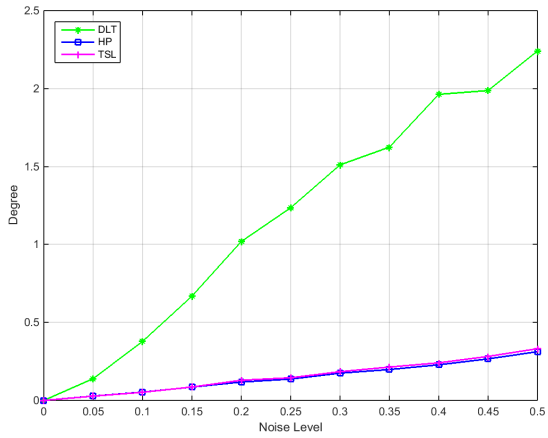
[2]As shown in [23], when the line is close to or passing through the origin of the coordinate frame, the quality of the estimated homographies decreases dramatically. This problem can be solved by proceeding a prior normalization of the line parameters. For the sake of simplicity and without losing generality, the lines generated in our experiment are forced to be away from the origin of the coordinate frame by at least 10 pixels.



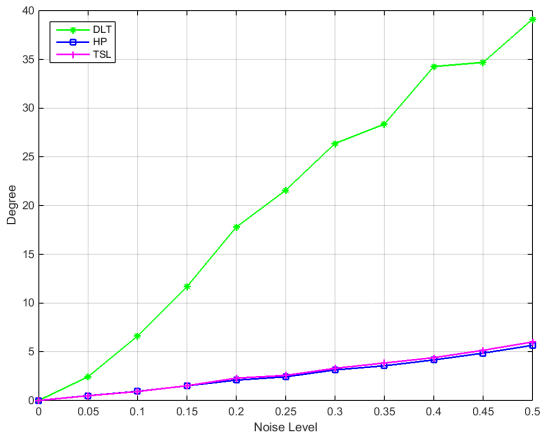
(a) Synthetic experiment configuration.




(b) Max norm of difference between groundtruth and estimated F.




(c) Error in rotation.




(d) Error in translation.

Fig. 1. Figure (a) shows the configuration of the experiment. The accuracy of fundamental matrix estimation is shown in Fig (b) with max norm as assessing criteria. Figures (c) (d) separately depicts rotation and translation error of DLT,HP,TSL.

matrix. A small error in any entry of the fundamental matrix can easily cause the resulting epipole to severly deviate from the groundtruth location.

We can easily see that our conclusions from the synthetic experiment are verified, namely that the proposed method clearly outperforms DLT and shows advantages over HP as well.

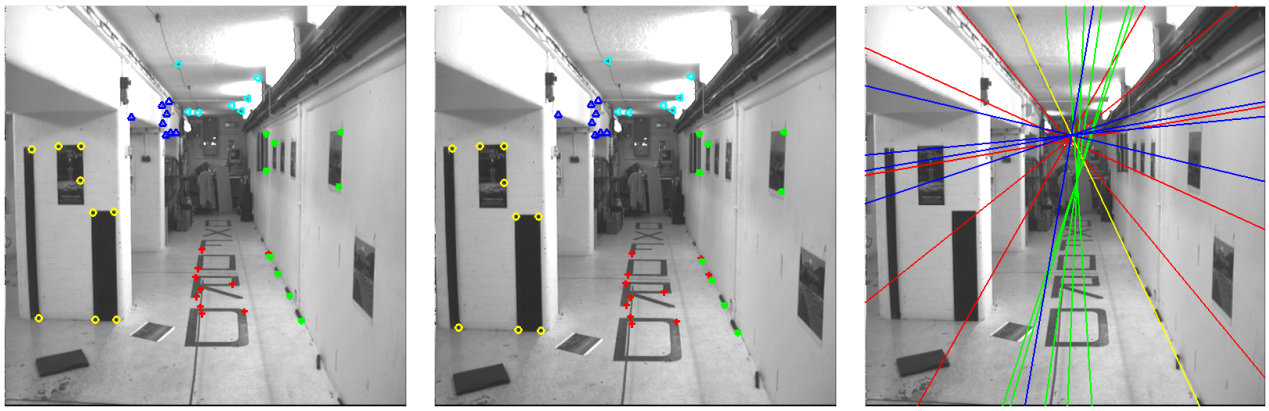
### C. Numerical stability and algorithmic complexity

It is easy to understand why the performance of DLT can be dramatically improved by including normalization. Without the normalization, as shown in Eq. (1) and Eq. (2), some of the entries are smaller than the others by several magnitude which directly causes the serious ill-conditioning of the original least squares problem. We record the numerical stability of DLT and TSL. As can be seen in Fig 3, the condition numbers of the three normalized sub least squares problem are far smaller than the one of the DLT solution. The average variance of the condition number also demonstrates that TSL is numerically more stable. A simple complexity is given in Tab. I. In our experiment, $N = 80$ and $M = 5$. TSL and HP lead to similar performance under these conditions, while TSL however needs less computational resources than HP.

TABLE I
ALGORITHM COMPLEXITY COMPARISON

| Method | Input | Matrix size to be solved |
|---|---|---|
| HP | $N$ points (not coplanar, $N \geqslant 8$) | $\mathbf{A}_{N\times 9}$ |
| TSL | $M$ planes ($M \geqslant 3$) | $3 \times \mathbf{A}_{M\times 4} + \mathbf{A}_{3M\times 3}$ |

It is worth pointing out that, during the experiment, we discovered that if the consistency among the inter-frame homographies is guaranteed, the estimated fundamental matrix is always accurate and robust no matter which method is used. Usually, perfect consistency constraints are available only in implicit form which can only be achieved by iterative non-linear methods, e.g. Joint Bundle Adjustment (BA-Joint) and

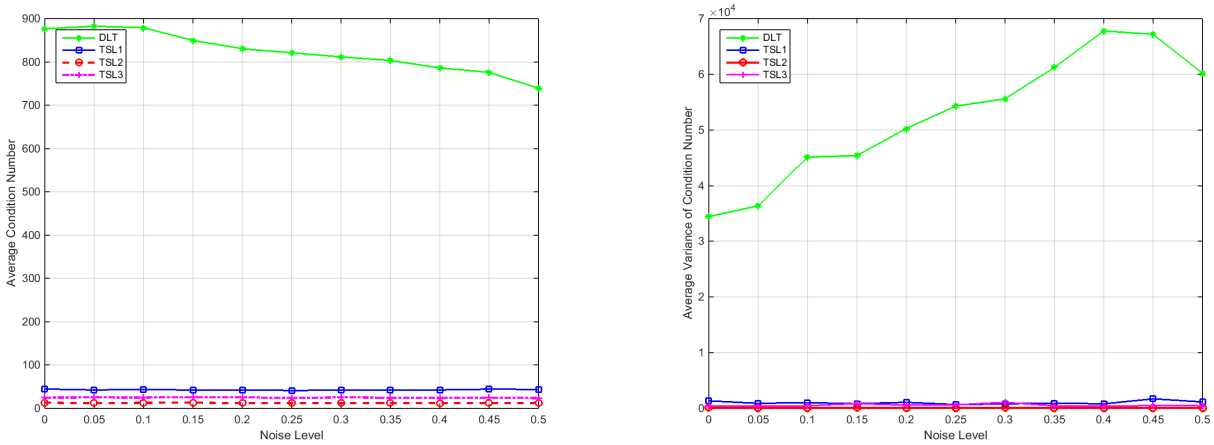


(a) Grouped point features in image one

(b) Grouped point features in image two

(c) Epipolar lines of groundtruth and all three methods

Fig. 2. Grouped point features which are used for estimating the homographies are shown in Fig (a) and (b). Epipolar lines obtained by DLT(yellow), HP(green), TSL(blue) and groundtruth (red) are shown in Fig (c).




(a) Average condition number of DLT and TSL

(b) Average variance of condition number

Fig. 3. The average condition number under each noise level is shown in Figure (a). TSL1, TSL2 and TSL3 are the three sub least squares problems of TSL. Figure (b) shows the corresponding average variance of the condition number.

AML [20, 3]. Explicit methods like [16, 22, 2] use a low-rank approximation under the Frobenius norm or the Mahalanobis norm to enforce the rank-four constraint. However, the explicit form is derived from a relaxed consistency constraint which means the consistency cannot be perfectly guaranteed. This discovery in fact gives an alternative explanation to why the direct estimation of the fundamental matrix by the compatibility equation is not stable.

## V. Conclusion

In this paper, we revisited an old topic: accurately and robustly estimating the fundamental matrix given a collection of independently estimated homograhies. We first review three classical methods and then show that a simple but non-trivial two-step normalization within the direct linear method achieves similar performance than the less attractive and more computationally intensive hallucinated points based method. We verify the correctness and robustness of our method by giving a mathematical proof and an experimental evaluation on both synthetic and real data. The numerical stability analysis and algorithm complexity discussion finally demonstrates our improvement and further advantages of the proposed technique.

## Acknowledgment

The research is supported by the ARC Centre of Excellence for Robotic Vision. The work is furthermore supported by ARC grants DP120103896 and DE150101365. Yi Zhou is funded by the Chinese Scholarship Council to be a PhD student at the Australian National University. We also thank Dr. Yuchao Dai for his comments and suggestions to improve this work.

## REFERENCES


[1] Anubhav Agarwal, CV Jawahar, and PJ Narayanan. A survey of planar homography estimation techniques. *Centre for Visual Information Technology, Tech. Rep. IIIT/TR/2005/12*, 2005.

[2] Pei Chen and David Suter. Rank constraints for homographies over two views: revisiting the rank four constraint. *International journal of computer vision*, 81(2):205–225, 2009.

[3] Wojciech Chojnacki, Zygmunt L Szpak, Michael J Brooks, and Anton van den Hengel. Enforcing consistency constraints in uncalibrated multiple homography estimation using latent variables. *Machine Vision and Applications*, 26(2-3):401–422, 2015.

[4] Elan Dubrofsky. *Homography estimation*. PhD thesis, UNIVERSITY OF BRITISH COLUMBIA (Vancouver, 2009.

[5] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[6] JJ Guerrero and C Sagues. From lines to homographies between uncalibrated images. In *IX Symposium on Pattern Recognition and Image Analysis, VO4*, pages 233–240, 2001.

[7] José J Guerrero and Carlos Sagüés. Robust line matching and estimate of homographies simultaneously. In *Pattern Recognition and Image Analysis*, pages 297–307. Springer, 2003.

[8] Chris Harris and Mike Stephens. A combined corner and edge detector. In *Alvey vision conference*, volume 15, page 50. Citeseer, 1988.

[9] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[10] Richard Hartley et al. In defense of the eight-point algorithm. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(6):580–593, 1997.

[11] H Christopher Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Readings in Computer Vision: Issues, Problems, Principles, and Paradigms, MA Fischler and O. Firschein, eds*, pages 61–62, 1987.

[12] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[13] QT Luong and O Faugeras. Determining the fundamental matrix with planes. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 489–494, 1993.

[14] Quan-Tuan Luong and Olivier D Faugeras. The fundamental matrix: Theory, algorithms, and stability analysis. *International journal of computer vision*, 17(1):43–75, 1996.

[15] Philip Pritchett and Andrew Zisserman. Matching and reconstruction from widely separated views. In *3D Structure from Multiple Images of Large-Scale Environments*, pages 78–92. Springer, 1998.

[16] Amnon Shashua and Shai Avidan. The rank 4 constraint in multiple ($\geq$3) view geometry. In *Computer VisionECCV'96*, pages 196–206. Springer, 1996.

[17] David Sinclair, H Christensen, and C Rothwell. Using the relation between a plane projectivity and the fundamental matrix. In *Proc. SCIA*, pages 181–188, 1995.

[18] Charles V Stewart. Robust parameter estimation in computer vision. *SIAM review*, 41(3):513–537, 1999.

[19] Richard Szeliski and Philip HS Torr. Geometrically constrained structure from motion: Points on planes. In *3D Structure from Multiple Images of Large-Scale Environments*, pages 171–186. Springer, 1998.

[20] Zygmunt L Szpak, Wojciech Chojnacki, Anders Eriksson, and Anton van den Hengel. Sampson distance based joint estimation of multiple homographies with uncalibrated cameras. *Computer Vision and Image Understanding*, 125:200–213, 2014.

[21] Etienne Vincent and Robert Laganiére. Detecting planar homographies in an image pair. In *Image and Signal Processing and Analysis, 2001. ISPA 2001. Proceedings of the 2nd International Symposium on*, pages 182–187. IEEE, 2001.

[22] L Zeinik-Manor and Michal Irani. Multiview constraints on homographies. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(2):214–223, 2002.

[23] Hui Zeng, Xiaoming Deng, and Zhanyi Hu. A new normalized method on line-based homography estimation. *Pattern Recognition Letters*, 29(9):1236–1244, 2008.

[24] Zhengyou Zhang. Determining the epipolar geometry and its uncertainty: A review. *International journal of computer vision*, 27(2):161–195, 1998.