A Simple Solution to the Six-point Two-view Focal-length Problem

Hongdong Li

RSISE,
The Australian National University.
National ICT Australia.

Abstract. This paper presents a simple and practical solution to the 6-point 2-view focal-length estimation problem. Based on the *hidden-variable* technique we have derived a 15th degree polynomial in the unknown focal-length. During this course, a simple and constructive algorithm is established. To make use of multiple redundant measurements and then select the best solution, we suggest a kernel-voting scheme. The algorithm has been tested on both synthetic data and real images. Satisfactory results are obtained for both cases. For reference purpose we include our Matlab implementation in the paper, which is quite concise, consisting of 20 lines of code only. The result of this paper will make a small but useful module in many computer vision systems.

1 Introduction

This paper considers the problem of estimating a constant unknown focal-length from six corresponding points of two semi-calibrated views. By *semi-calibration* we mean that all camera intrinsic parameters but a fixed focal-length are known. This scenario is quite common (not restrictive) in daily camera use. For example, except for the case where the camera lens is allowed to zoom continuously, it is often practical to assume that its focal-length is constant across multiple views. In fact, all *other* camera intrinsic parameters (such as principal point and aspect ratio) can be considered fixed and known for a certain camera. In other words, the only user-adjustable (therefore variable) camera intrinsic parameter is the focal-length. Yet still, the focal-length is often kept constant over two successive image shoots [7][8].

It is well known that five points of two fully-calibrated views are possible to recover the essential matrix E between the two views. Since an essential matrix is a faithful representation of the camera motion (up to an unknown scale), namely, $\mathsf{E} = [t]_\times \mathsf{R}$, it therefore has five degrees of freedom. So, from five points it is possible to estimate the camera motion—this is exactly what the five-point algorithm does [11].

Now consider a semi-calibrated case where only a fixed focal-length f is unknown. For this case, it is shown that six points (in general position) are enough to estimate the camera motion as well as the unknown focal-length. This can be easily seen by the following reasoning. Compared to the fully-calibrated five-point case, the one extra point correspondence will provide one more constraint on the camera intrinsic matrix. Consequently, a single unknown focal-length, as well as the relative camera motion, can be computed from it.

The above conclusion can also be approached from the other direction. If the two camera views are uncalibrated, then seven points are the minimal requirement to compute a fundamental matrix F. Since a fundamental matrix has seven degrees of freedom, it provides two more constraints on the camera intrinsics, besides the camera motion encoded by an essential matrix. These two extra constraints are essentially the two Kruppa equations. Therefore, if the two views are only partially calibrated in that all camera intrinsics but *two* possibly different focal-lengths f and f' are known, then seven points are enough to estimate the relative camera motion and the two unknown focal-lengths [2]. Now, if we have only seven less one points, and the two focal-lengths are assumed identical, then it is possible to recover the single unknown focal-length as well as the camera motion from six corresponding points.

For the first time, Stewénius et al have proposed a concrete algorithm to solve the 6-point focal-length problem [1]. They have utilized a special mathematical tool—the Gröbner basis technique. The idea behind the Gröbner basis is to construct a complete and algebraically-closed polynomial system (an *ideal*) by adding in some newly *generated* compatible equations. By this tool they show that there are at most 15 solutions to the six-point algorithm. The Gröbner basis is a mathematically elegant technique for handling polynomial system. However, since it originates from a special mathematical field (i.e. *computational commutative algebra and algebraic geometry*), some readers may find it not fully-comfortable to follow, let alone to actually implement it and use it.

Why Six Points? Traditionally, the focal-length problem is solved through the fundamental matrix which itself can be computed from seven points. Moving from seven points to six points provides some benefits. The first benefit lies in its theoretic value. Compared with its non-minimal counterpart, the minimal algorithm offers a deeper theoretical understanding to the problem itself. For example, both the five-point algorithms [11] and the six point algorithm all better exploit the constraints provided by the epipolar equations and the Kruppa equations (cf. [14][15]); Secondly, effective techniques developed during the course of deriving the six-point algorithm are very useful for other similar vision problems too (e.g. [11]); Thirdly, for the task of focal-length estimation itself, it is demonstrated by experiments that the six-point algorithm sometimes offers even better performance than the seven-point algorithm; In addition, as shown in [11], six-point algorithm has less degenerate configurations than the seven-point algorithm; Moreover, when combine a minimal solver with the RANSAC scheme using six points (rather than seven) allows significant reduction in computation [5].

1.1 Main Contributions

This paper provides an alternative yet much simpler and practical solution to the 6-point focal-length problem, compared to the one originally proposed in [1].

We will show that to solve the 6-point problem there is *no* need to generate new equations. The original equations system, which includes the six epipolar conditions, one singularity condition and two Kruppa equations, already provides sufficient and algebraically-closed constraints to the problem. As a result, in the real domain $\mathbb{R}$ it is already enough to solve the six-point problem using 10 *rigidity equations*—equivalent to the above equations—without resorting to the Gröbner basis technique. For reference purposes we provide our implementation in the appendix of the paper, which is very concise and consists of 20 lines of general Matlab code only.

Paper [1] tested its algorithm mainly on noise-free simulation data. In this paper, we go beyond such an idealized scenario. We have tested the performance of our algorithm on both synthetic and real images (with different levels of noise). We demonstrate our results by the accuracy of focal-length estimation *per se*, rather than by the errors in the reprojected fundamental matrix.

In the root-selection stage (whose purpose is to single-out the best root from multiple solutions), we propose a *kernel-voting* scheme, as an alternative to the conventionally adopted RANSAC. We show by experiments that our scheme is suitable for the particular problem context, and there is no need to wait until the reprojected fundamental matrix error is obtained.

2 Theoretic Backgrounds

Consider a camera, with constant intrinsic parameters denoted by a matrix $\mathsf{K} \in \mathbb{R}^{3\times 3}$, observing a static scene. Two corresponding image points $\mathbf{m}$ and $\mathbf{m}'$ are related by a fundamental matrix $\mathsf{F} \in \mathbb{R}^{3\times 3}$:

$$\mathbf{m}'^{\mathbf{T}}\mathsf{F}\mathbf{m} = 0. \tag{1}$$

A valid fundamental matrix must satisfy the following singularity condition:

$$\det(\mathsf{F}) = 0. \tag{2}$$

This is a cubic equation. Remember that the 3×3 fundamental matrix is only defined up to a scale, it therefore has 7 degrees of freedom in total. Consequently, seven corresponding points are sufficient to estimate the F.

If the camera is fully-calibrated, then the fundamental matrix is reduced to an *essential matrix*, denoted by E, and the relationship between them reads as:

$$\mathsf{K}^{-\mathbf{T}}\mathsf{E}\mathsf{K}^{-1} = \mathsf{F}. \tag{3}$$

Since an essential matrix E is a faithful representation of the relative camera motion (translation and rotation, up to a scale), it has only five degrees of freedom. Consequently, to be a valid essential matrix E, it must further satisfy two more constraints, which are characterized by the following theorem.

Theorem-1: A real 3×3 *matrix* E is an essential matrix *if and only if* it satisfies the following condition

$$2\mathsf{E}\mathsf{E}^{\mathbf{T}}\mathsf{E} - \mathbf{tr}(\mathsf{E}\mathsf{E}^{\mathbf{T}})\mathsf{E} = 0. \tag{4}$$

This gives 9 equations in the elements of E, but only two of them are algebraically independent. The above theorem, owing to many researchers (e.g, Kruppa, Demazure, Maybank, Huang, Trivedi, Faugeras, etc, just name a few, cf. [6][5][15]), is an important result in geometric vision.

For the semi-calibrated case considered here, since only one focal-length is unknown, without loss of generality we can assume the intrinsic camera matrix is: $\mathsf{K} =$

$\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}$, where f is the focal-length. Define a matrix $\mathsf{Q} = w^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & w \end{bmatrix}$, where $w = f^{-2}$. Write down the epipolar relations Eq. (1) for six points $\mathbf{m}_i$ and $\mathbf{m}'_i$,

$$\mathbf{m}'^{\mathbf{T}}_i \mathsf{F} \mathbf{m}_i = 0, \tag{5}$$

for $i = 1, \cdots, 6$. Using the six points we get a linear representation of the fundamental matrix:

$$\mathsf{F} = x\mathbf{F}_0 + y\mathbf{F}_1 + z\mathbf{F}_2, \tag{6}$$

where x, y, z are three unknown scalars to be estimated, and $\mathbf{F}_0, \mathbf{F}_1, \mathbf{F}_2$ are the bases of the null-space of the *epipolar design matrix*, which can be readily computed from the six points (cf. [5]).

Substituting this F into Eq.(3) and Eq.(4), we get the following equations in the unknown set $\{w, x, y, z\}$.

$$2\mathsf{F}\mathsf{Q}\mathsf{F}^{\mathbf{T}}\mathsf{Q}\mathsf{F} - \mathbf{tr}(\mathsf{F}\mathsf{Q}\mathsf{F}^{\mathbf{T}}\mathsf{Q})\mathsf{F} = 0. \tag{7}$$

This is a group of nine equations, and they provide sufficient conditions to find the unknown $\{w, x, y, z\}$ (up to an unknown scalar). If we somehow solve these equations, then the task of estimating the focal-length is accomplished. The above reasoning basically follows [1].

3 Review the Previous Algorithm

Stewénius et al proposed a clever algorithm based on the Gröbner basis technique [1]. More precisely, it is a variant of the classical Gröbner technique [4]. The key steps of their algorithm are briefly reviewed below.

Given six corresponding image points in general position, write down Eq.(7) and Eq.(2). Rearrange them in such a way that a 10×33 matrix equation $\mathsf{A}\mathbf{X} = 0$ is obtained, where A is a 10×33 coefficient matrix, and $\mathbf{X}$ a vector containing 33 terms of monomials of the unknowns. Now we have a polynomial system of 10 equations. This system is then ported into a finite field $\mathbb{Z}_p$ (p is a large *prime number*), and is solved using the Gröbner basis elimination procedure. This procedure is stopped when the whole system becomes an algebraically-closed *ideal generator set* of the original system. So far, a minimal solver (for $\mathbb{Z}_p$) has been built up.

The next step is to apply the same solver (i.e, the same sequence of elimination) to the original problem. One then obtains an enlarged polynomial system containing $n \times 33$ ($n > 10$) monomial terms. Finally, a generalized eigen-decomposition is employed to solve the polynomial system, for which there are 15 solutions. In order to improve numerical stability, a pivoted Gauss-Jordan elimination is used.

An important detail of the algorithm is that the arbitrary scale factor of the fundamental matrix is parameterized by setting one unknown to an arbitrary scalar. Thereby the number of unknowns is reduced by one, which simplifies the later derivation. By contrast, in this paper we avoid such scale parametrization in order to keep the homogeneity of some unknowns of the equation system. The reason will be explained later.

Limitations The main mathematical device adopted by [1] is the Gröbner basis technique. The Gröbner basis is an elegant and powerful technique[3] [4]. Many commercial or free mathematical software packages include it as a standard module (for instance, in Maple and Mathematica etc). In many cases, to use it the user is not assumed to have specialized knowledge of it, and thus can simply apply it in a black-box manner, as also claimed by [1]. However, using a tool in a black-box manner is not always a safe way. Whenever a program runs into trouble, it would be nicer if the user could understand its internal mechanism. Moreover, due to its special origin (*computational commutative algebra and algebraic geometry*), not every reader finds it easy to follow. Furthermore, paper [1] did not test its algorithm extensively on more realistic case. It experimented on perfect simulated data only. No result on real images was given there.

Finally, the *root-selection* procedure (i.e., single-out the best root from the possibly multiple solutions) is not addressed by paper [1], because it deals with simulated cases only, and thereby assumes the ground-truth data is available. However, in a real-world problem an efficient root-selection mechanism is necessary. It is in fact a common requirement for various minimal solvers (see for example [11] and [8]), where one often obtains multiple and maybe complex roots. The RANSAC is a good scheme to find the best solution from multiple candidates. In this paper, we propose an alternative *kernel-voting* scheme which is suitable for the particular context.

4 Our New Six-Point Algorithm

In this paper, we propose a new method for solving the six-point focal-length problem, using the *hidden variable* technique which is probably the best known technique for algebra elimination.

We claim that the recommended *hidden variable* technique is *not yet-another* specialized mathematical technique (which otherwise would be equally unfamiliar and uncomfortable to readers), but it follows very straightforward principle and procedures. It is so transparent and simple to the end-user that is almost self-explained. As will be described later, to better apply this technique to the problem, we introduce a small trick that is to keep the homogeneity of some unknowns of the equation system.

Hidden Variable Technique The Hidden-Variable technique (also known as the *Dialytic Elimination*) is possibly one of the best known *resultant* techniques in algebraic geometry [4]. It is used to eliminate variables from a multivariate polynomial equation system. The basic idea is as follows.

Consider a system of M homogeneous polynomial equations in N variables, say, $p_i(x_1, x_2, ..., x_N) = 0$, for $i = 1, 2, ..., M$. If we treat one of the unknowns (for example, x_1) as a *parameter* (in the conventional terms, we *hide* the variable x_1), then by some simple algebra we can re-write the equation system as a matrix equation

$$\mathtt{C}(x_1)\mathbf{X} = 0,$$

where the coefficient matrix $\mathtt{C}$ will depend on the *hidden variable* x_1, and the $\mathbf{X}$ is a vector space consisting of the homogeneous monomial terms of all other N-1 variables (say, $x_2, x_3, \cdots, x_N$).

If the number of equations equals the number of monomial terms in the vector $\mathbf{X}$ (i.e. the matrix C is square), then one will have a *resultant equation* defined on x_1, say, $\det(C(x_1)) = 0$ *if and only if* the equation system has non-trivial solutions. By such procedure, one thus successfully eliminates N-1 variables from the equation system all at once. Solving the resulting resultant equation for x_1 and back-substituting it, one thus eventually solves the whole system.

4.1 Algorithm Derivation

Remember that Eq.(2) and Eq.(7) are the main equations we are to use. Notice that they are ten cubic equations in the four unknowns $\{w, x, y, z\}$. A careful analysis will show that within the real domain $\mathbb{R}$, Eq.(7) already implies Eq.(2). However, we would keep all these ten equations together in our derivation, and the reason will become clear soon.

Now we treat the unknown w as the hidden variable, and collect a coefficient matrix (denoted by $\mathsf{C}(w)$) with respect to the other three variables $\{x, y, z\}$. Here we do not replace one variable with an arbitrary scalar. Rather, we keep the homogeneous forms in the monomials formed by $[x, y, z]$. These are all cubic monomials which actually span a vector space:

$$\mathbf{X} = [xyz, x^2z, xy^2, xz^2, y^2z, yz^2, x^3, y^3, x^2y, z^3]^{\mathbf{T}} \tag{8}$$

To give a more close examination of the coefficient matrix C, we list it element-wise:

	0 xyz	1 x^2z	2 xy^2	3 xz^2	4 y^2z	5 yz^2	6 x^3	7 y^3	8 x^2y	9 z^3
0	s_0	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9
1	$[w]_{10}$	$[w]_{11}$	$[w]_{12}$	$[w]_{13}$	$[w]_{14}$	$[w]_{15}$	$[w]_{16}$	$[w]_{17}$	$[w]_{18}$	$[w]_{19}$
2	$[w]_{20}$	$[w]_{21}$	$[w]_{22}$	$[w]_{23}$	$[w]_{24}$	$[w]_{25}$	$[w]_{26}$	$[w]_{27}$	$[w]_{28}$	$[w]_{29}$
3	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$[w]_{i-1,j}$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\cdots$	$\cdots$	$[w]_{i,j-1}$	$[w]_{i,j}$	$[w]_{i,j+1}$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$[w]_{i+1,j}$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
8	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
9	$[w]_{90}$	$[w]_{91}$	$[w]_{92}$	$[w]_{93}$	$[w]_{94}$	$[w]_{95}$	$[w]_{96}$	$[w]_{97}$	$[w]_{98}$	$[w]_{99}$

Here, elements in the first row are some scalars, $\mathsf{C}(i, j) = s_j$, for $i = 0$, computed from the singularity constraint Eq.(2). Elements of all other rows are quadratic in w, computed from the nine rigidity constraints Eq.(4). More precisely, it is in the form of $\mathsf{C}(i, j) = [w]_{ij} \doteq a_{ij}w^2 + b_{ij}w + c_{ij}$, for $1 \leqslant i \leqslant 9$.

As the monomial vector has been kept homogenous, the equation $\mathsf{C}(w)\mathbf{X} = 0$ will have non-trivial solutions of $\{x, y, z\}$ if and only if the determinant of the coefficient matrix vanishes. That is:

$$\boxed{\det(\mathsf{C}(w)) = 0.} \tag{9}$$

This determinant is better known as a *hidden-variable-resultant*, which is an univariate polynomial of the hidden variable, w.

By observing the elements of C, one would expect that its determinant is an 18th degree polynomial. However, a more close inspection reveals that: it is actually a 15th