# Error Analysis of Pure Rotation-Based Self-Calibration

Lei Wang, *Student Member, IEEE*, Sing Bing Kang, *Sr. Member, IEEE*, Heung-Yeung Shum, *Sr. Member, IEEE*, and Guangyou Xu, *Sr. Member, IEEE* 

Abstract—Self-calibration using pure rotation is a well-known technique and has been shown to be a reliable means for recovering intrinsic camera parameters. However, in practice, it is virtually impossible to ensure that the camera motion for this type of self-calibration is a pure rotation. In this paper, we present an error analysis of recovered intrinsic camera parameters due to the presence of translation. We derived closed-form error expressions for a single pair of images with nondegenerate motion; for multiple rotations for which there are no closed-form solutions, analysis was done through repeated experiments. Among others, we show that translation-independent solutions do exist under certain practical conditions. Our analysis can be used to help choose the least error-prone approach (if multiple approaches exist) for a given set of conditions.

- 🔶

Index Terms-Self-calibration, rotating cameras, error analysis.

# **1** INTRODUCTION

SELF-CALIBRATION refers to a collection of camera calibration techniques that rely only on scene features to extract camera parameters. As a result, it is a very desirable and practical means for calibrating the camera. There has been a significant body of work done in this area, such as that of Maybank and Faugeras' [10], in which the Kruppa equation was used. Of particular interest to us is Hartley's [5], [6] method of self-calibration from pure rotation based on recovered homographies. An alternative is to assume planar scenes [15], [16], [19]. More recently, others have extended the self-calibration problem to recovering variable intrinsic parameters. By making some assumptions on the camera intrinsic parameters, recovery is possible using a closed-form solution [2], [11], [12].

For the case of self-calibration with the assumption of purely rotating cameras, the effects of possible nonzero translation are mostly ignored. It is, in practice, impossible to enforce the purerotation assumption under nonlaboratory conditions. The case of translations is mentioned in [6], but no error analysis was provided. Stein [13] did consider nonzero translations in his work and his somewhat brief analysis indicated that translations are generally bad.

Our work is most similar to that of Hayman and Murray [7], while our work first appeared in [18]. They analyze the error caused by nonzero translations to only the focal length. While their analysis is for zooming cameras, they use the assumption that the camera is mounted on top of a tripod or on pan-tilt devices, which provides additional constraints to recover the focal length. Their

- L. Wang and G.Y. Xu are with the Department of Computer Science and Technology, Tsinghua University, Beijing 100084, P.R. China.
- E-mail: wanglei99@mails.tsinghua.edu.cn, dcs-xgy@tsinghua.edu.cn.
  S.B. Kang is with Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399. E-mail: sbkang@microsoft.com.
- H.-Y. Shum is with Microsoft Research Asia, No. 49, Zhichun Road, Haidian, Beijing 100080, P.R. China. E-mail: hshum@microsoft.com.

Recommended for acceptance by L. Quan.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 116578.

work also mentioned that nonlinear solutions provide better results, while our work mainly focused on linear solutions.

Our analysis shows that the effects of nonzero translations can be mitigated under certain conditions. It also shows why some methods are better than others in computing intrinsic camera parameters. In particular, we computed the error bounds for some practical special cases and showed the unexpected results that particular derivations are independent of certain translational components. In our analysis, we do not consider degenerate motions, as was done in [9], [2], [14]. In addition, we assume the intrinsic parameters remain the same throughout the calibration process.

The paper is organized as follows: We start by defining terms and describing the problem in Section 2. Our analysis of the 1, 2, and 3-parameter problem is described in Sections 3, 4, and 5, respectively. (The 4-parameter and 5-parameter analysis and comparison with perturbation analysis are given in [17]. It is omitted due to space limitations.) Section 6 discusses the results and limitations of our analysis, with concluding remarks provided in Section 7.

# 2 Some Preliminaries

In this section, we describe the different cases in our analysis, the problem of self-calibration with pure rotation, and we list our notations used in the rest of the paper.

# 2.1 Calibration Matrix

The calibration matrix we use is of the following form:

$$A = \begin{bmatrix} f & s & x_0 \\ 0 & af & y_0 \\ 0 & 0 & 1 \end{bmatrix},$$
 (1)

where *f* is the focal length, *a* is the aspect ratio, *s* is the image skew, and  $(x_0, y_0)$  is the principal point.

Our analysis covers the following cases:

2

- 1. 1-parameter estimation: f unknown, with a = 1, s = 0,  $x_0 = y_0 = 0$ ;
- 2. 2-parameter estimation: f and a unknown, with s = 0,  $x_0 = y_0 = 0$ ;
- 3. 3-parameter estimation:  $f, x_0$ , and  $y_0$  unknown, with a = 1, s = 0, and f, a, and s unknown, with  $x_0 = y_0 = 0$ ;

The 4-parameter case, with s known as zero, and the 5-parameter case, where all the intrinsics are unknown, are treated in [17].

# 2.2 Errors Due to Translation

The idea of self-calibration from pure rotation is to first register the rotated images to extract their homographies. Using these homographies, we can then extract the calibration matrix A. The homography associated with a pair of rotated images is of the form  $H = ARA^{-1}$ , where R is the rotation matrix. All these steps ignore errors due to image registration (e.g., resampling problems, photometric variation across images, bad local minimum). In our analysis, we ignore these errors as well. In theory, for a purely rotating camera, the estimated calibration matrix  $A_{est}$  is equal to the actual calibration matrix A.

Suppose that there is now some camera translation t. For the case of the single plane scenario, we have

$$H = A\left(R + \frac{\mathbf{tn}^T}{d}\right)A^{-1},\tag{2}$$

where  $\mathbf{n}$  is the vector of the plane and d is the distance of the camera center to the plane.

Now, if we were to self-calibrate with the assumption of a purely rotating camera, we will essentially be force-fitting

Manuscript received 20 May 2002; revised 13 Dec. 2002; accepted 17 July 2003.

$$A_{est}(R_{est})A_{est}^{-1} \approx H = A\left(R + \frac{\mathbf{tn}^T}{d}\right)A^{-1},\tag{3}$$

resulting in errors in the recovered intrinsic parameters, i.e.,  $\Delta A = A_{est} - A \neq 0_{3\times 3}.$ 

In the general case of multiple planes in the scene and with nonzero camera translation, we can greatly simplify our analysis by assuming the resulting homography from image registration is due to some effective "average" global plane in the scene with some parallax, i.e.,

$$H = A \left( R + \frac{\mathbf{tn}_e^T}{d_e} \right) A^{-1}.$$
 (4)

In our analysis, we use this effective one-plane assumption.

# 2.3 Additional Notations

The rotation matrix R and homography H are written as

$$R = \begin{bmatrix} r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 \\ r_7 & r_8 & r_9 \end{bmatrix} \text{ and } H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}.$$

We also represent *R* by its rotation axis, i.e.,  $R = \theta(r_x, r_y, r_z)$ .

In our analysis, we normalize the camera translation with respect to the distance of the scene plane from the camera, i.e.,  $d_e = 1$ , giving  $\frac{\mathbf{t}}{d_e} = \mathbf{t} = [t_x \ t_y \ t_z]^T$ . The effective plane normal is  $\mathbf{n}_e = [n_x \ n_y \ n_z]^T$ . We now start our analysis with the 1-parameter estimation and then work our way to the most general case.

# **3** 1-PARAMETER ESTIMATION

In this case, our calibration matrix is reduced to

$$A = \begin{bmatrix} f & 0 & 0\\ 0 & f & 0\\ 0 & 0 & 1 \end{bmatrix}.$$

For pure rotation, the homography matrix is

$$H = ARA^{-1} = \begin{bmatrix} r_1 & r_2 & fr_3\\ r_4 & r_5 & fr_6\\ \frac{r_7}{f} & \frac{r_8}{f} & r_9 \end{bmatrix}.$$

## 3.1 Closed-Form Solution

Since R is orthonormal, there are several ways in which we can extract f, as described in [12]. For example, either

$$h_1^2 + h_2^2 + h_3^2/f^2 = h_4^2 + h_5^2 + h_6^2/f^2$$
(5)

or

$$h_1h_4 + h_2h_5 + h_3h_6/f^2 = 0 (6)$$

can be used. The solution associated with (6) is

$$f_{est} = \sqrt{\frac{-h_3h_6}{h_1h_4 + h_2h_5}}.$$
 (7)

Note that this solution is applicable only if the camera is not rotated exactly about the x, y, or z-axis.

In the case of pure camera rotation, both (5) and (6) are accurate. However, with translation, we have

$$H = A \left( R + \frac{\mathbf{tn}_{e}^{T}}{d_{e}} \right) A^{-1} = A \left( R + \mathbf{tn}_{e}^{T} \right) A^{-1} = \begin{bmatrix} r_{1} + t_{x}n_{x} & r_{2} + t_{x}n_{y} & f(r_{3} + t_{x}n_{z}) \\ r_{4} + t_{y}n_{x} & r_{5} + t_{y}n_{y} & f(r_{6} + t_{y}n_{z}) \\ \frac{r_{7} + t_{e}n_{x}}{f} & r_{9} + t_{z}n_{z} \end{bmatrix}.$$
(8)

It is interesting to note how the translational error components are distributed according to rows in (7). This could influence our choice of rows to use to compute f, especially if we have a priori knowledge of the approximate relative magnitudes of the camera translation components.

Suppose we go ahead and use (7) to compute f. The solution now contains an error that is *independent* of  $t_{z_i}$  i.e.,

$$f_{est} = f \sqrt{\frac{-(r_3 + t_x n_z)(r_6 + t_y n_z)}{q}},$$

$$q = (r_1 + t_x n_x)(r_4 + t_y n_x) + (r_2 + t_x n_y)(r_5 + t_y n_y).$$

Noting that  $r_1r_4 + r_2r_5 = -r_3r_6$  and ignoring second and higher order terms involving  $t_x$  and  $t_y$ , we get

$$\frac{f_{est}}{f} \approx \sqrt{1 + t_x q_1 + t_y q_2},$$

$$q_1 = \left(\frac{n_z}{r_3} + \frac{r_4 n_x}{r_3 r_6} + \frac{r_5 n_y}{r_3 r_6}\right)$$
 and  $q_2 = \left(\frac{n_z}{r_6} + \frac{r_2 n_y}{r_3 r_6} + \frac{r_1 n_x}{r_3 r_6}\right)$ 

From the above expression, the absolute relative error is computed as

$$\left|\frac{f - f_{est}}{f}\right| \approx \frac{1}{2} \left| (t_x q_1 + t_y q_2) \right| \le \frac{1}{2} \left(\frac{|t_x| + |t_y|}{|r_3 r_6|}\right)$$

As mentioned before, this method is independent of  $t_z$  and should be used if  $t_z$  is known to dominate. We can use similar reasoning to choose different rows or columns to compute f under different conditions of t and  $n_e$ .

## 3.2 Special Configurations

where

Let us now consider two special cases of the 1-parameter estimation case that are either commonly assumed or used in practice, namely, the fronto-parallel plane and panning motion.

### 3.2.1 Fronto-Parallel Plane

If the scene plane is fronto-parallel, i.e.,  $\mathbf{n}_e = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ , the homography matrix reduces to

$$H = \begin{bmatrix} r_1 & r_2 & f(r_3 + t_x) \\ r_4 & r_5 & f(r_6 + t_y) \\ \frac{r_7}{f} & \frac{r_8}{f} & r_9 + t_z \end{bmatrix}.$$

Note that the first and second columns are independent of the translation! In theory, under this condition, we can compute f **exactly** from these vectors using

$$f = \sqrt{-\frac{h_1 h_2 + h_4 h_5}{h_7 h_8}} \tag{9}$$

or 
$$f = \sqrt{\frac{h_2^2 + h_5^2 - h_1^2 - h_4^2}{h_7^2 - h_8^2}},$$
 (10)

noting that, when the rotation angle is quite small,  $h_7$  and  $h_8$  are very close to zero, which will produce unstable results. It suggests using a large rotation angle for calibration.

We can also analyze the error bound of (5) and (6). For (6), by ignoring second and higher order terms for  $t_x$  and  $t_y$ , we get

$$f_{est} = f \sqrt{-\frac{(r_3 + t_x)(r_6 + t_y)}{r_1 r_4 + r_2 r_5}} \approx f \left(1 + \frac{1}{2} \left(\frac{t_x}{r_3} + \frac{t_y}{r_6}\right)\right)$$

For (5),

$$f_{est} = \sqrt{\frac{h_3^2 - h_6^2}{h_4^2 + h_5^2 - h_1^2 - h_2^2}} \approx f\left(1 + \frac{r_3 t_x - r_6 t_y}{r_3^2 - r_6^2}\right).$$
 (11)

If  $R = \theta(0, 1, 0)$ , (11) reduces to

$$\frac{f_{est} - f}{f} \approx \frac{t_x}{\sin \theta}.$$
(12)

This suggests that larger rotations are more favorable. However, we have ignored the registration error, which depends on the amount of overlap and overlapping texture. If  $\theta = 15^{\circ}$ , for example, the error is expected to be nearly  $4t_x$ .

In fact, it is not uncommon for the scene plane to be nearly frontoparallel or distant to the camera. Our analysis provides interesting insights in these cases.

## 3.2.2 Panning Motion

In many applications, the camera motion is that of controlled panning. Letting  $c_{\theta} = \cos \theta$ ,  $s_{\theta} = \sin \theta$ ,  $\mathbf{t} = [t_x \ 0 \ t_z]^T$ ,

$$R = \theta(0, 1, 0) = \begin{bmatrix} c_{\theta} & 0 & s_{\theta} \\ 0 & 1 & 0 \\ -s_{\theta} & 0 & c_{\theta} \end{bmatrix}, \text{ and}$$
$$H = \begin{bmatrix} c_{\theta} + t_x n_x & t_x n_y & f(s_{\theta} + t_x n_z) \\ 0 & 1 & 0 \\ \frac{-s_{\theta} + t_z n_x}{f} & c_{\theta} + t_z n_z \end{bmatrix}.$$

In this case, we have a closed-form exact solution. As we know,

so

$$h_7 = \frac{-s_{\theta} + t_z n_x}{f} = \frac{-s_{\theta}}{f} + \frac{h_8 t_x n_x}{h_2} = \frac{-s_{\theta}}{f} + \frac{h_8 (h_1 - c_{\theta})}{h_2}.$$

 $h_2/h_8 = ft_x/t_z,$ 

Similarly,

$$h_9 = c_\theta + \frac{h_8(h_3 - fs_\theta)}{h_2},$$

from which we can solve f using

$$\begin{cases} w_1 = h_3 h_8 - h_2 h_9, w_2 = h_2 h_7 - h_1 h_8\\ \theta = \cos^{-1} \left( \frac{-(w_1 w_2 + h_2 h_8)}{w_1 h_8 + w_2 h_2} \right), f = \frac{h_3 h_8 - h_2 h_9 + h_2 \cos \theta}{h_8 \sin \theta}. \end{cases}$$

However, for (5), ignoring second and higher order terms in  $t_x$ ,

$$f_{est} = \sqrt{\frac{h_3^2 - h_6^2}{h_4^2 + h_5^2 - h_1^2 - h_2^2}} \approx f\left(1 + \frac{1}{s_\theta} \left(t_x n_z + \frac{t_x n_x}{\tan \theta}\right)\right).$$

Note that this method is only dependent on  $t_x$  and that the error increases quickly when the rotation angle  $\theta$  decreases. If  $\mathbf{n} = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$ , the error reduces to (12).

# 4 2-PARAMETER ESTIMATION

For the 2-parameter case,

$$A = \begin{bmatrix} f & 0 & 0 \\ 0 & af & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } H = ARA^{-1} = \begin{bmatrix} r_1 & r_2/a & fr_3 \\ ar_4 & r_5 & afr_6 \\ r_7/f & r_8/af & r_9 \end{bmatrix}.$$

# 4.1 Closed-Form Solution

From the orthogonality constraints  $\begin{bmatrix} r_1 & r_2 & r_3 \end{bmatrix} \begin{bmatrix} r_4 & r_5 & r_6 \end{bmatrix}^T = 0$  and  $\begin{bmatrix} r_1 & r_4 & r_7 \end{bmatrix} \begin{bmatrix} r_2 & r_5 & r_8 \end{bmatrix}^T = 0$ , we have

$$\begin{cases} \frac{h_1h_4}{a} + ah_2h_5 + \frac{h_3h_6}{af^2} &= 0\\ ah_1h_2 + \frac{h_4h_5}{a} + af^2h_7h_8 &= 0 \end{cases}$$

Hence, we have a closed-form solution (choosing only real and positive solutions). Its error could be written as in the previous section.

## 4.2 Special Configuration

As before, we consider the two special cases.

4.2.1 Fronto-Parallel Plane If  $\mathbf{n}_e = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ , then

$$H = \begin{bmatrix} r_1 & r_2/a & f(r_3 + t_x) \\ ar_4 & r_5 & af(r_6 + t_y) \\ r_7/f & r_8/af & r_9 + t_z \end{bmatrix}$$

We still have first and second column vectors independent of the translation. Again, we get the **exact** solution.

# 4.2.2 Panning Motion

With camera panning, we have

$$H = \begin{bmatrix} c_{\theta} + t_x n_x & \frac{t_x n_y}{a} & f(s_{\theta} + t_x n_z) \\ 0 & 1 & 0 \\ \frac{-s_{\theta} + t_z n_x}{f} & \frac{t_z n_y}{af} & c_{\theta} + t_z n_z \end{bmatrix}$$

It is interesting that  $h_1, h_3, h_7, h_9, h_2/h_8$  are just the same as in 1-parameter estimation and that these parameters are sufficient to solve *f*. As a result, we have an **exact** solution for *f*, but not for *a*. This verifies the claim of [2].

# **5 3-PARAMETER ESTIMATION**

For 3-parameter estimation,

$$\begin{split} A &= \begin{bmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and} \\ H &= \begin{bmatrix} r_1 + \frac{r_7 x_0}{f} & r_2 + \frac{r_8 x_0}{f} & fr_3 + r_9 x_0 - r_1 x_0 - r_2 y_0 - \frac{r_7 x_0^2 + r_8 x_0 y_0}{f} \\ r_4 + \frac{r_7 y_0}{f} & r_5 + \frac{r_8 y_0}{f} & fr_6 + r_9 y_0 - r_4 x_0 - r_5 y_0 - \frac{r_7 x_0 y_0 + r_8 y_0^2}{f} \\ \frac{r_7}{f} & \frac{r_8}{f} & r_9 - \frac{r_7 x_0 r_7 x_0 + r_8 y_0}{f} \end{bmatrix}$$

# 5.1 Closed-Form Solution

First, we impose the constraint det(H) = 1. Since HA = AR, we get  $HAA^{T}H^{T} = ARR^{T}A^{T} = AA^{T}$ . Taking the inverse [2], we have  $H^{-T}A^{-T}A^{-1}H^{-1} = A^{-T}A^{-1}$ . Let

$$B = f^2 A^{-T} A^{-1} = \begin{bmatrix} 1 & 0 & b_1 \\ 0 & 1 & b_2 \\ b_1 & b_2 & b_3 \end{bmatrix}, \quad G = H^{-T} = \begin{bmatrix} g_1 & g_2 & g_3 \\ g_4 & g_5 & g_6 \\ g_7 & g_8 & g_9 \end{bmatrix},$$

where  $b_1 = -x_0, b_2 = -y_0, b_3 = f^2 + x_0^2 + y_0^2$ . Linearizing  $GBG^T = B$ , we have  $G'\mathbf{b} = 0$ , where

$$G' = \begin{bmatrix} 2g_1g_3 & 2g_2g_3 & g_3^2 & g_1^2 + g_2^2 - 1 \\ g_3g_4 + g_1g_6 & g_3g_5 + g_2g_6 & g_3g_6 & g_1g_4 + g_2g_5 \\ g_3g_7 + g_1g_9 - 1 & g_3g_8 + g_2g_9 & g_3g_9 & g_1g_7 + g_2g_8 \\ 2g_4g_6 & 2g_5g_6 & g_6^2 & g_4^2 + g_5^2 - 1 \\ g_6g_7 + g_4g_9 & g_6g_8 + g_5g_9 - 1 & g_6g_9 & g_4g_7 + g_5g_8 \\ 2g_7g_9 & 2g_8g_9 & g_9^2 - 1 & g_7^2 + g_8^2 \end{bmatrix} \text{ and }$$

 $\mathbf{b} = \begin{bmatrix} b_1 & b_2 & b_3 & 1 \end{bmatrix}^T.$ 



Fig. 1. Three-parameter estimation results (t: translation,  $\theta$ : rotation,  $\sigma_{err}$ : error standard deviation). Effect of: (a) |t| on  $\sigma_{err}$  of f, (b) number of homographies on  $\sigma_{err}^{-2}$  of f, (c) |t|, and  $\theta$  on  $\sigma_{err}$  of f, (d) |t| and  $\theta$  on mean error of f, (e) |t| on  $\sigma_{err}$  of  $x_0$ , and (f) |t| on  $\sigma_{err}$  of  $y_0$ . Each data point is based on 200 randomized trials in all cases, except (b) on 1,000 randomized trials.

From three independent rows, we can get a closed-form solution, which is too long to write out in full here. Once  $b_1, b_2$ , and  $b_3$  are computed, we can obtain  $f, x_0$ , and  $y_0$  with a closed-form solution as well as the error function. The degenerate configurations, which are described in [2], are beyond the scope of this paper.

When we have more homographies, we just stack them up to yield (similar to [6])

$$G'_M \mathbf{b} = 0$$

from which we could solve B and then solve the intrinsic f and  $x_0$  can then be found from the following equations: parameter matrix A.

#### 5.2 Special Configurations

As the number of unknowns increases, even the special configurations become significantly more complicated.

## 5.2.1 Fronto-Parallel Plane

It is difficult to find an exact solution in this case. However, since  $\mathbf{n}_e = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ , the terms suggest that it is desirable to put more weight on the first two columns in parameter estimation, since they are theoretically independent of translation.

## 5.2.2 Panning Motion

With camera panning,

$$\begin{split} H = & \\ \begin{bmatrix} c_{\theta} + t_x n_x + \frac{t_z n_x - s_{\theta}}{f} x_0 & t_x n_y + \frac{t_z n_y}{f} x_0 & f(s_{\theta} + t_x n_z) + t_z n_z x_0 - t_x n_{xy0} + \frac{s_{\theta} x_0 - t_z n_{xy0}}{f} x_0 \\ \frac{-s_{\theta} + t_z n_x}{f} y_0 & 1 + \frac{t_z n_y}{f} y_0 & y_0 (-1 + c_{\theta} + t_z n_z) + \frac{s_{\theta} x_0 - t_z n_{xy0}}{f} y_0 \\ \frac{-s_{\theta} + t_z n_x}{f} & t_z n_y & c_{\theta} + t_z n_z + \frac{s_{\theta} x_0 - t_z n_{xy0}}{f} \end{bmatrix}, \end{split}$$

where  $n_{xy0} = n_x x_0 + n_y y_0$ . By inspection, we can easily see that  $y_0 = h_4/h_7$ . In addition, if the angle  $\theta$  is known,

$$\begin{aligned} t_x n_x &= h_1 - c_\theta - h_7 x_0, t_z n_x = f h_7 + s_\theta, \\ t_x n_y &= h_2 - h_8 x_0, t_z n_y = f h_8, \\ t_x n_z &= (h_3 - h_7 x_0^2 + h_1 x_0 - h_9 x_0 + h_2 y_0 - h_8 x_0 y_0 - f s_\theta) / f, \end{aligned}$$

and  $t_{z}n_{z} = h_{9} + h_{7}x_{0} + h_{8}y_{0} - c_{\theta}$ , which lead to

$$\begin{aligned} \frac{t_x}{t_z} &= \frac{h_1 - c_\theta - h_7 x_0}{f h_7 + s_\theta} = \frac{h_2 - h_8 x_0}{f h_8} \\ &= \frac{(h_3 - h_7 x_0^2 + h_1 x_0 - h_9 x_0 + h_2 y_0 - h_8 x_0 y_0 - f s_\theta)/f}{h_9 + h_7 x_0 + h_8 y_0 - c_\theta}. \end{aligned}$$

$$\begin{bmatrix} h_1h_8 - h_8c_\theta - h_2h_7 & h_8s_\theta \\ h_8s_\theta & h_2h_7 + h_8c_\theta - h_1h_8 \end{bmatrix} \begin{bmatrix} f \\ x_0 \end{bmatrix}$$
$$= \begin{bmatrix} h_2s_\theta \\ h_3h_8 - h_2h_9 + h_2c_\theta \end{bmatrix}.$$

### Experiments 5.3

If  $\mathbf{n}_e$  is randomized and rotation angle is 15°, we obtain the error graph shown in Figs. 1a, 1e, and 1f for  $f, x_0$ , and  $y_0$  separately. As indicated in Section 2.3, the magnitude of the translation t is normalized with respect to the distance of the scene plane to the first camera.

These error graphs show that camera translation can have a dramatic effect on the accuracy of the intrinsic parameters. For example, in Figs. 1e and 1f, the error standard deviation is more than 100 percent even though |t| is only 0.1. However, we expect that using different homographies will improve the stability of this method, even though it does not provide closed-form solutions [6]. The error graph in Figs. 1a, 1e, and 1f shows the effect of using different numbers of homographies on the standard deviation of the error in f,  $x_0$ , and  $y_0$  under independently random rotations and random translations. As we can see, in general, the results

Unknown Parameters	f	f,a	$f, x_0, y_0$	f, a, s [17]
Closed-form solution	Yes	Yes	Yes	Yes
Fronto-parallel	Exact for $f$	Exact for $f, a$	Indeterminate	Indeterminate
Panning	Exact for $f$	Exact for $f$ ,	Exact for $y_0$ only	Exact for $f$ ,
		not for $a$	$(\theta \text{ unknown});$	not for <i>a</i> , <i>s</i>
			Exact for $f, x_0, y_0$	( $\theta$ is known)
			$(\theta \text{ is known})$	

TABLE 1 Comparison of Different Parameter Estimation Cases

become more stable with a higher number of homographies. It is interesting to note that, for estimating f, the error standard deviation for the case of two homographies is smaller than that of five homographies. The reason of this behavior would be an interesting topic for future research.

For a translation t with a fixed magnitude, we would expect the error variance  $\sigma_{err}$  to vary with the number of homographies  $N_h$  according to the statistical sampling relation

$$\sigma_{err} \propto \frac{1}{\sqrt{N_h}} \text{ or } \frac{1}{\sigma_{err}^2} \propto N_h.$$

This relationship is verified in Fig. 1b.

Camera rotation appears to be a significant factor in accuracy and stability. Fig. 1c shows the graph of error and translation for different rotation angles. The error standard deviation decreases with increasing angle. The graph in Fig. 1d shows how the mean error of *f* changes with different translations and rotation angles. While the mean error is more significant with  $|\mathbf{t}|$ , its degradation is more dramatic with the decrease in rotation angle. It is interesting to see that the average value of results is not the right answer for  $|\mathbf{t}| \neq 0$ . This is because overestimation and underestimation of f are not symmetric with respect to the sign of t. For example, if  $\mathbf{n} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ ,  $R = 15^{\circ}(0.1, 0.9747, 0.2)$ ,  $\mathbf{t} = \begin{bmatrix} 0.01 & 0.02 & -0.01 \end{bmatrix}^T$ ,  $\frac{\Delta f}{f}=0.0157;$  with the same plane vector and rotation,  $\mathbf{t} = [-0.01 \ -0.02 \ 0.01]^T$ ,  $\frac{\Delta f}{f} = 0.0004$ , i.e., they are both overestimated. In fact, experiments show that the mean of f is usually underestimated. We obtained similar results for the other two parameters.

## 5.4 Another Configuration

For another 3-parameter camera configuration,

$$A = \begin{bmatrix} f & s & 0\\ 0 & af & 0\\ 0 & 0 & 1 \end{bmatrix}, \text{ and}$$
$$H = \begin{bmatrix} r_1 + \frac{r_4s}{f} & \frac{r_2}{a} + \frac{r_5 - r_1}{af}s - \frac{r_4s^2}{af^2} & fr_3 + r_6s\\ ar_4 & r_5 - \frac{r_4s}{f} & afr_6\\ \frac{r_7}{f} & \frac{r_8}{af} - \frac{r_7s}{af^2} & r_9 \end{bmatrix}.$$

In this case, we still have a closed-form solution to solve for *A*. For special configurations like fronto-parallel plane and panning motion, it is difficult to find an exact solution, but if  $\theta$  is known (such as [1], [3]), we can find an **exact** closed-form solution. Due to space limitation, we left the details in a longer version of this paper [17].

# 6 DISCUSSION

Our analysis on the different multiple parameter estimation cases has yielded rather interesting results. For example, under certain conditions (see Table 1), exact solutions can be obtained despite the presence of translation. Also, in the 1, 2, and 3-parameter estimation cases with the fronto-parallel plane condition, the first two columns of the homography are theoretically independent of the camera translation. This can be used to influence the manner in which the unknowns are computed, especially if multiple approaches exist.

We have verified that using more homographies produces more stable and accurate results. In addition, using a bigger rotation angle is usually better. We should caution, however, that we have ignored registration effects such as reduction of overlapped image areas, local texture distribution, and errors in resampling. The inclusion of these effects to the analysis could be an area for future research.

The complexity of analysis increases dramatically with the number of parameters to be estimated. Our experiments also indicated that parameter estimation is more sensitive to translational errors with more parameters.

In addition to ignoring registration errors, we also made the assumption of the existence of an effective plane  $(\mathbf{n}_e, d_e)$ . We think this is a reasonable assumption as the resulting homography computed from image registration (despite the presence of residual parallax) can always be interpreted as a 3D plane.

An alternative means for investigating the effect of translation on the intrinsic parameters is through component-wise perturbation analysis [8]. However, it is unable to provide the same level of indepth and detailed insights on the problem as our analysis (which gives exact solutions in certain cases). In addition, as we have shown in a longer version of this paper [17], the perturbation analysis results in high error bounds, which are not very useful in practice.

# 7 SUMMARY AND CONCLUSIONS

We have analyzed the error for pure rotation-based self-calibration in the presence of camera translation. In particular, we considered three cases, namely, the 1, 2, and 3-parameter estimation problems. The algorithm and analysis of 4-parameter case is an extension of the  $(f, x_0, y_0)$  case, and the 5-parameter estimation problems are similar to the (f, a, s) case. We have a closed-form solution for the 4parameter case, while we have no closed-form solution for 5parameter case. They are considered in [17]. A summary of results of our analysis of these cases is given in Table 1. Note that having a closed-form solution does not necessarily mean that there is an exact solution. The closed-form solution is a direct function of the homography matrix entries, which may include error terms due to camera translation. The exact solution is an error-free solution that is independent of translation.

The special conditions of a fronto-parallel plane and camera panning is especially interesting, primarily because of the convenience of camera panning. Our analysis shows that, in some cases, it is theoretically possible to recover correct solutions in spite of translation! As a result, our analysis can help the user choose the least error-prone approach (if multiple approaches exist) for a given set of conditions.

The other conclusions from our analysis are:

- Camera translation can have a dramatic effect on the accuracy of the extracted intrinsic parameters. From a practical point of view, it is desirable to self-calibrate using distant scenes.
- The larger the rotation, the more accurate and stable the solution is (ignoring registration issues).
- The greater number of different homographies used the better.
- Because the overestimation and underestimation of f is not symmetric with respect to the sign of t, simply taking the average value of several trials will not improve the result.

# ACKNOWLEDGMENTS

The authors would like to thank Andrew Zisserman for recommending the comparison with perturbation analysis. They are also grateful to Zhoucheng Lin for fruitful discussions on matrix analysis techniques, as well as to Steve Lin and Tao Feng for their suggestions on improving the paper. A shorter version of this submission appeared in the Proceedings of the International Conference on Computer Vision, vol. 1, pp. 464-471, July 2001.

## REFERENCES

- A. Basu, "Active Calibration: Alternative Strategy and Analysis," Proc. [1] Conf. Computer Vision and Pattern Recognition, pp. 495-500, June 1993.
- [2] L. de Agapito, E. Hayman, and I.D. Reid, " Self-Calibration of Rotating and Zooming Cameras, *Int'l J. Computer Vision*, vol. 45, no. 2, pp. 107-127, 2001. F. Du and M. Brady, "Self-Calibration of the Intrinsic Parameters of
- [3] Cameras for Active Vision Systems," Proc. Conf. Computer Vision and Pattern Recognition, pp. 477-482, June 1993. G. Golub and L.C. Van, Matrix Computations, third ed. Baltimore: John
- [4] Hopkins Univ. Press, 1996.
- R.I. Hartley, "Self-Calibration from Multiple Views of a Rotating Camera," [5] Proc. European Conf. Computer Vision, vol. 1, pp. 471-478, May 1994. R.I. Hartley, "Self-Calibration of Stationary Cameras," Int'l J. Computer
- [6] *Vision*, vol. 22, no. 1, pp. 5-23, Feb. 1997. E. Hayman and D.W. Murray, "The Effect of Translational Misalignment in
- [7] the Self-Calibration of Rotating and Zooming Cameras," Technical Report OUEL 2250/02, Oxford Univ. Eng. Library, Sept. 2002.
- N.J. Higham, "A Survey of Componentwise Perturbation Theory in Numerical Linear Algebra," Math. of Computation 1943-1993: A Half Century [8] of Computational Mathematics, Proc. Symp. Applied Math., W. Gautschi, ed., vol. 48, pp. 49-77, 1994.
- Y. Ma, S. Soatto, J. Kosecka, and S. Sastry, "Euclidean Reconstruction and Reprojection up to Subgroups," *Proc. Int'l Conf. Computer Vision*, vol. 1, [9]
- pp. 773-780, Sept. 1999. [10] S. Maybank and O. Faugeras, "A Theory of Self-Calibration of a Moving
- Camera," *Int'l J. Computer Vision*, vol. 8, no. 2, pp. 123-151, 1992. Y. Seo and K.S. Hong, "About the Self-Calibration of a Rotating and [11] Zooming Camera: Theory and Practice," Proc. Int'l Conf. Computer Vision, vol. 1, pp. 183-188, Sept. 1999.
- [12] H.-Y. Shum and R. Szeliski, "Systems and Experiment Paper: Construction of Panoramic Image Mosaics with Global and Local Alignment," Int'l J. Computer Vision, vol. 36, no. 2, pp. 101-130, Feb. 2000. [13] G. Stein, "Accurate Internal Camera Calibration Using Rotation, with
- Analysis of Sources of Error," Proc. Int'l Conf. Computer Vision, pp. 230-236, Iune 1995.
- P. Sturm, " Critical Motion Sequences for Monocular Self-Calibration and [14] Uncalibrated Euclidean Reconstruction," Proc. Conf. Computer Vision and Pattern Recognition, pp. 1100-1105, June 1997.
- P. Sturm and S. Maybank, "On Plane-Based Camera Calibration: A General [15] Algorithm, Singularities, Applications," Proc. Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 432-437, June 1999. [16] B. Triggs, "Autocalibration from Planar Scenes," Proc. European Conf.
- Computer Vision, vol. 1, pp. 89-105, June 1998.

- [17] L. Wang, S. B. Kang, H. -Y. Shum, and G. Xu, "Error Analysis of Pure Rotation-Based Self-Calibration," MSR technical report, Microsoft Corp., 2002.
- [18] L. Wang, S. B. Kang, H.Y. Shum, and G.Y. Xu, "Error Analysis of Pure Rotation-Based Self-Calibration," Proc. Int'l Conf. Computer Vision, vol. 1, pp. 464-471, July 2001. Z.Y. Zhang, "A Flexible New Technique for Camera Calibration," *Proc. Int'l*
- [19] Conf. Computer Vision, vol. 1, pp. 666-673, Sept. 1999.

> For more information on this or any other computing topic, please visit our Digital Library at http://computer.org/publications/dlib.