

# Motion Bias and Structure Distortion induced by Calibration Errors

Marco Zucchelli<sup>†</sup>   Jana Košecká<sup>‡</sup>

<sup>†</sup> CVAP & CAS, Royal Institute of Technology, Stockholm, Sweden S100 44

<sup>‡</sup> Department of Computer Science, George Mason University, Fairfax, VA 22030

## Abstract

This article provides an account of sensitivity and robustness of structure and motion recovery with respect to the errors in intrinsic parameters of the camera. We demonstrate both analytically and in simulation, the interplay between measurement and calibration errors and their effect on motion and structure estimates. In particular we show that the calibration errors introduce an additional bias towards the optical axis, which has opposite sign to the bias typically observed by egomotion algorithms. The overall bias causes a distortion of the resulting 3D structure, which we express in a parametric form. The analysis and experiments are carried out in the differential setting for motion and structure estimation from image velocities. While the analytical explanations are derived in the context of linear techniques for motion estimation, we verify our observations experimentally on a variety of optimal and suboptimal motion and structure estimation algorithms. The obtained results illuminate and explain the performance and sensitivity of the differential structure and motion recovery techniques in the presence of calibration errors.

## 1 Introduction

While the basic geometric relationships governing the problem of structure and motion recovery from image sequences are well understood, the existence of robust automatic techniques for recovery of motion and structure is still elusive. Different aspects of the performance and sensitivity of the existing general techniques for motion and structure recovery have been addressed in the past. While the list of references is by no means exhaustive, we mention here few more recent representative works addressing the problem. The intrinsic sensitivity of the differential formulation of the problem has been studied thoroughly using analytical techniques as well as simulations [13, 11, 14]. These studies assumed calibrated cameras and focused on determining sensitive directions of motion, dependency on the depth variation and field of view [16]. The process of camera calibration introduces additional errors in the measurements, which affect the final estimates. This is the case both when the camera is calibrated off-line or when self-calibration techniques are used towards this end. With the exception of few, the study of these effects, has not received much attention. Various empirical observations regarding the stability of the estimation of intrinsic parameters and their effect on the structure estimation in the discrete setting have been made by Bougnoux [12]. He pursued the stratified approach to

Euclidean reconstruction and experimentally demonstrated that, in spite of depth distortions caused by calibration errors, the basic geometric relationships (orthogonality, parallelism) were preserved. The effect of calibration errors on motion estimates in the discrete setting have also been explored in [8], assuming noise free measurements of corresponding points. In [7] the authors derived the covariances of the parameters of an uncalibrated stereo system, both for fixed calibration parameters and for the hypothesis that an *a priori* Gaussian distribution for the calibration parameters is known. The effect of this prior knowledge on the quality of the final estimates was demonstrated in the context of non-linear optimization techniques. In the differential setting Cheong [11] characterized the depth distortions due to the free varying focal length using the analytical iso-distortion framework developed previously in the uncalibrated case. While the iso-distortion framework enables to study the intrinsic distortions as a family of transformation parametrized by errors in motion estimates, it does not assume any particular distribution of noise in the image measurements and camera parameters. Hence it is not suitable for quantifying the quality of the final estimates.

We present an analytical study of the sensitivity of egomotion and structure estimation assuming both noisy measurements and errors in intrinsic parameters of the camera. Resorting to some approximations, we both analytically and experimentally demonstrate that the errors in calibration introduce *an additional bias term* in motion estimation, which reduces the previously observed translation bias reported in [2]. The overall translation bias distorts the resulting structure proportionally to the distance of the measurements from the optical axis. These distortions are further accentuated by the errors in focal length and center of the projection. We offer an analytical explanation of the errors introduced by calibration in the context of linear techniques. The observations are verified by simulation of a variety of linear and nonlinear algorithms for structure and motion estimation and confirm the intrinsic nature of the errors, making them independent of the algorithm choice and objective function used.

## 2 Problem Formulation

In order to study the problem of motion and structure recovery for image sequences, when the individual frames are closely separated in time, we employ the differential formulation of the problem. The starting point of our analysis are the basic relationships between the motion of a 3D point  $\mathbf{X} = [X, Y, Z]^T$  induced by the camera motion  $(\omega, v)$  and the associated image plane motion field  $\mathbf{u}$ :

$$\dot{\mathbf{X}} = \hat{\omega}\mathbf{X} + v \quad \text{and} \quad \mathbf{u} = A(\mathbf{x})\omega + \frac{1}{\lambda}B(\mathbf{x})v \quad (1)$$

where  $\hat{\omega}$  denotes a skew symmetric matrix associated with the vector  $\omega$ , such that  $\hat{\omega} = \omega \times \mathbf{X}$ . The image point  $\mathbf{x} = [x, y, 1]^T$  and its 3D counterpart  $\mathbf{X}$  are related by an unknown scale  $\lambda$ , such that  $\lambda\mathbf{x} = \mathbf{X}$ , which in the case of perspective projection is  $\lambda = Z$ . The image plane motion field is  $\mathbf{u} = [\dot{x}, \dot{y}, 0]^T = [u_x, u_y, 0]^T$  and  $A(\mathbf{x})$ ,  $B(\mathbf{x})$  are functions of image coordinates defined as follows:

$$A = \begin{pmatrix} xy & -(1+x^2) & y \\ (1+y^2) & -xy & -x \\ 0 & 0 & 0 \end{pmatrix}; \quad B = \begin{pmatrix} -1 & 0 & x \\ 0 & -1 & y \\ 0 & 0 & 0 \end{pmatrix}.$$

The above relationship assumes so called normalized camera coordinate system  $\mathbf{x} = [x, y, 1]^T$  (with the retinal plane at  $Z = 1$ ). The retinal plane coordinates are related to the image (pixel) coordinates  $\mathbf{x}_p = [x_p, y_p, 1]^T$  via matrix  $K$  of camera intrinsic parameters. For the purpose of the analysis we consider simplified model of the calibration matrix assuming no skew  $s_\gamma = 0$  and the aspect ratio  $\alpha = \frac{k_x}{k_y} = 1$ , where  $1/k_x$  and  $1/k_y$  are horizontal and vertical pixel sizes respectively. Denoting  $o_x, o_y$  as the coordinates of the optical center projection and  $f$  as the focal length of the camera expressed in pixel units, the calibration matrix  $K$  is then defined as:

$$K = \begin{pmatrix} -f & 0 & o_x \\ 0 & -f & o_y \\ 0 & 0 & 1 \end{pmatrix}.$$

Camera's field of view (FOV)  $\theta$  is related to the dimension of the focal plane  $I$ , with  $\tan \theta = \frac{I}{f}$ .

The focus of this paper is to study the interplay between noisy image velocities and errors in calibration and their effects on the resulting motion and structure estimates. We initially assume that the calibration parameters are obtained by an off-line calibration procedure. For the purpose of sensitivity analysis assume that the entries of  $K$ :  $f, o_x, o_y$  are corrupted by zero-mean Gaussian noise  $n_f \sim N(0, \sigma_f^2)$ ,  $n_{o_x} \sim N(0, \sigma_{o_x}^2)$ ,  $n_{o_y} \sim N(0, \sigma_{o_y}^2)$  respectively. Since the focal length and optical center position can be measured independently and the choice of the axis in the retinal plane is arbitrary, we assume that the focal length  $f$  and  $o_x, o_y$  are uncorrelated random variables and the errors in  $o_x$  and  $o_y$  are identical with  $\sigma_{o_x} = \sigma_{o_y} = \sigma_o$ . The errors in calibration affect the coordinates of feature positions  $\mathbf{x}$  as well as image velocities  $\mathbf{u}$ . These are measured in the retinal plane and are expressed as nonlinear functions of the intrinsic parameters of the camera in the following way:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{f} \begin{pmatrix} o_x - x_p \\ o_y - y_p \end{pmatrix} \equiv \frac{1}{f} \begin{pmatrix} \Delta_x \\ \Delta_y \end{pmatrix} \quad \text{and} \quad \mathbf{u} = \frac{1}{f} \mathbf{u}_p = \frac{1}{f} (\mathbf{x}_p(t+1) - \mathbf{x}_p(t)) \quad (2)$$

where  $\Delta_x = o_x - x_p$  and  $\Delta_y = o_y - y_p$ . Note that image velocities  $\mathbf{u}$  depend only on the focal length  $f$  and are unaffected by the knowledge of the center of the projection. We further assume that the calibration errors are much smaller than optical flow errors. This approximation is valid for cameras calibrated with a calibration grid where the camera parameters can be estimated with errors up to few percent [15]. The optical flow errors are typically on the order of a few tens of percent [1], hence  $\frac{\sigma_u^2}{u^2} / \frac{\sigma_f^2}{f^2} \approx 100$ . We will refer to the above assumption later in the text and will use it only to simplify the analytical derivation. The experimental results will be provided for a broader set of conditions.

The study of the intrinsic sensitivity of the motion and structure estimation in the analytical setting is a difficult task due to the nonlinear nature of the problem. With the exception of few, the studies of sensitivity and robustness typically resort to simulations [18]. In the following paragraph we will demonstrate that by making some approximations we can draw general conclusions regarding the biasness and sensitivity analytically. In order to do so, we start with a linear formulation of the structure and motion estimation based on the bilinear epipolar constraint. In Section 5 we verify our observations by simulation, using a wide spectrum of different motion estimation algorithms.

## 2.1 Egomotion Estimation

In order to decouple the effect of measurement errors on the structure and motion estimates we employ the epipolar geometry formulation. The epipolar constraint is obtained by algebraic elimination of the unknown depth  $\lambda$  from the equation (1) and yields following relationships in calibrated and uncalibrated case:

$$\mathbf{u}^T \hat{v} \mathbf{x} + \mathbf{x}^T \hat{v} \hat{\omega} \mathbf{x} = 0 \quad \text{and} \quad \mathbf{u}^T K^{-T} \hat{v} K^{-1} \mathbf{x} + \mathbf{x}^T K^{-T} \hat{v} \hat{\omega} K^{-1} \mathbf{x} = 0 \quad (3)$$

Since the uncalibrated differential epipolar constraint does not provide enough constraints for the recovery of the unknown camera parameters from image measurements only, the complete camera self-calibration in the differential case is not possible. Equation (3) may be rewritten to highlight the dependence on the optical flow as:

$$[A_1(\mathbf{u}) \mid A_2] \mathbf{e} = A \mathbf{e} = 0 \quad (4)$$

where  $A_1(\mathbf{u}) \in \mathbb{R}^{m \times 3}$  is a linear function of the measured optical flow and  $A_2 \in \mathbb{R}^{m \times 6}$  is a function of the image points alone; vector  $\mathbf{e} \in \mathbb{R}^9$ ,  $\mathbf{e} = [v_x, v_y, v_z, s_1, s_2, s_3, s_4, s_5, s_6]^T$  is associated with the unknown parameters  $v$  and the symmetric part  $s = \frac{1}{2}(\hat{\omega} \hat{v} + \hat{v} \hat{\omega})$  of the matrix  $\hat{\omega} \hat{v}$ . Minimizing  $\|A \mathbf{e}\|^2$  leads to LLSE estimate of  $\mathbf{e}$ , which is obtained as the eigenvector of  $A^T A$  associated with the smallest eigenvalue. Translation estimate  $v$  is then directly available and angular velocity  $\omega$  can be obtained by decomposition of the special symmetric matrix  $s$ . This particular parametrization and the associated algorithm is described in greater detail in [9].

Assuming perfect calibration and measurement noise due to the temporal matching, only the columns of  $A$  related to matrix  $A_1$  are corrupted by noise. This causes the linear techniques to lead to biased estimates. Few techniques for removing the bias were suggested by [2, 3]. We will now demonstrate the source of this bias and its interplay with the bias induced by calibration errors. If we assume that calibration is perfect and each component of  $\mathbf{u}$  is corrupted by  $n \sim N(0, \sigma_u^2)$  due to the temporal matching, the noise in image velocities perturbs the matrix  $A$  by  $\delta A$ . This perturbation alters the eigenvectors of  $A^T A$ . It can be shown using perturbation theory for *Hermitian* matrices that if  $E[\delta(A^T A)] \neq 0$ , then the eigenvectors are biased (see [10] or [4]). First note that the rows of  $A_1$  have the form  $\tau = \hat{\mathbf{u}} \mathbf{x}$  while those of  $A_2$  are of the form  $p = (x^2, 2xy, 2xz, y^2, 2yz, z^2)$ . Lets write the noise free  $A^T A$  in terms of constraints  $(\tau_i, p_i)$  and denote its upper-left block as  $D$  :

$$A^T A = \sum_i \begin{pmatrix} \tau_i \tau_i^T & \tau_i p_i^T \\ p_i \tau_i^T & p_i p_i^T \end{pmatrix} \quad \text{and} \quad D = \sum_i \tau_i \tau_i^T$$

The entries of  $A^T A$  are nonlinear functions of  $\mathbf{x}$  and  $\mathbf{u}$ , which means that if these are corrupted by zero-mean Gaussian noise, the expectation value of the error  $\delta(A^T A)$  in  $A^T A$  is not zero. Considering the errors in temporal matching only, it is only the block of the error matrix  $\delta(A^T A)$  associated with  $D$  that has expectation value different from zero<sup>1</sup>. In fact the non-diagonal block elements are linear in  $\mathbf{u}$  and the  $p_i p_i^T$  block is independent of  $\mathbf{u}$ . Denote the noisy constraints  $\tilde{\tau}_i$ ,  $\tilde{\tau}_i = \tau_i + n$ . By propagating the errors

<sup>1</sup>The expected values for the remaining blocks are zero. In fact  $E[p_i p_i^T] = 0$  since  $x_i, y_i$  are noise free and  $E[p_i \tau_i^T] \propto E[u_x] = E[u_y] = 0$  with the assumption that components of  $\mathbf{u}$  are IID zero-mean Gaussian,  $n \sim N(0, \sigma_u^2)$ .

in image velocities  $\mathbf{u}$  due to temporal matching to the constraint coefficients  $\tau_i$ , one can easily compute the covariance matrix  $\Sigma_{\tau_i}$  of  $\tau_i$ . Let  $\tilde{D}$  be the noisy matrix  $D$  and  $E[\tilde{D}]$  its expectation value, such that  $E[\tilde{D}] = E[D] + \sigma_u^2 \Sigma$  where  $\Sigma$  is:

$$\Sigma = \sum_i \begin{pmatrix} 1 & 0 & -x_i \\ 0 & 1 & -y_i \\ -x_i & -y_i & x_i^2 + y_i^2 \end{pmatrix} = \sum_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & x_i^2 + y_i^2 - 1 \end{pmatrix} \quad (5)$$

Assuming uniformly distributed features,  $E[\sum_i x_i]$  and  $E[\sum_i y_i]$  are approximately zero and the part proportional to identity can be omitted since it does not change the eigenvectors of matrix  $D$ . We see that  $E[\Sigma] \neq 0$ , biasing the final estimate towards the optical axis. This has been previously observed both analytically and in simulations [5, 2].

### 3 Calibration and Egomotion

Suppose now that in addition to the errors due to flow computation, we want to understand how are the motion estimates affected by calibration errors. Assume that calibration parameters form a vector random variable  $\mathbf{k} = [o_x, o_y, f]^T$ , where the individual components are independent normally distributed random variables with  $N(0, \Sigma_c)$ , where  $\Sigma_c = \text{diag}([\sigma_o^2, \sigma_o^2, \sigma_f^2])$ . Each component of  $\mathbf{u}$  is also corrupted by errors  $n \sim N(0, \sigma_u^2)$  due to temporal matching. By propagating the errors in calibration to image positions and image velocities, we obtain characterization of the noisy feature positions  $\tilde{\mathbf{x}} \sim N(0, \Sigma_x)$  and noisy image velocities  $\tilde{\mathbf{u}} \sim N(0, \Sigma_u)$  as normal random variables. The form of these covariance matrices can be found in Appendix A. More detailed derivation can be found in [10]. Consequently the random variables  $\tilde{\mathbf{x}}$  and  $\tilde{\mathbf{u}}$  determine the structure of the error  $\delta(A^T A)$  via constraints  $(\tau_i, p_i)$ . Similarly as in the calibrated case we will justify, that the expected value of  $\delta(A^T A)$  is approximately block diagonal. This approximation is valid given the hypothesis that  $\mathbf{x}$  and  $\mathbf{u}$  are weakly correlated and  $E[u_i]$ , defined in Appendix A, is zero. In such case we obtain:

$$E[\tau_i p_i^T] \propto E[u_i] E[x_i^3] = 0 \Leftarrow E[u_i] = 0 \quad (6)$$

The block relative to  $p_i p_i^T$  contains coefficients of 4<sup>th</sup> powers in the image coordinates. Under the assumption that optical flow errors are bigger then calibration errors  $\frac{\sigma_u^2}{u^2} \gg \frac{\sigma_f^2}{f^2}$ , the relative errors of  $p_i p_i^T$  are small compared with the relative errors of  $\tau_i \tau_i^T$ . Then similarly as in the calibrated case we can approximate  $E[\delta(A^T A)]$  considering only the upper-left block  $D$ , with all the other entries being zero. In the presence of calibration errors the  $E[\tilde{D}]$  is augmented by two additional bias terms in the following way:

$$E[\tilde{D}] = E[D] + \sigma_1^2 \Sigma_1 + \sigma_2^2 \Sigma_2 \quad \text{where} \quad \Sigma_2 = \sum_i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & u_{x_i}^2 + u_{y_i}^2 \end{pmatrix} \quad (7)$$

where  $\Sigma_1 \equiv \Sigma$  as in the equation (5) and  $\sigma_1^2$  and  $\sigma_2^2$  have the following form:

$$\sigma_1^2 = \frac{1}{2} \frac{u^2}{f^2} \left( \frac{\sigma_u^2}{u^2} + \frac{\sigma_f^2}{f^2} \right) \quad \text{and} \quad \sigma_2^2 = \frac{1}{2} \frac{\Delta^2}{f^2} \left( \frac{\sigma_f^2}{f^2} + \frac{\sigma_o^2}{\Delta^2} \right) \quad (8)$$

where  $\Delta^2$  is the average distance of the features from the center of the projection, under the assumption that the features are uniformly distributed in the image plane. Its expected value can be related to the size of the image plane as  $E[\Delta^2] = I^2/12$ . Using the assumption that  $\frac{\sigma_u^2}{u^2} \gg \frac{\sigma_f^2}{f^2}$  and  $E[\Delta^2] = I^2/12$ , the expressions for  $\sigma_1^2$  and  $\sigma_2^2$  can be simplified to obtain:

$$\sigma_1^2 = \frac{1}{2} \frac{\sigma_u^2}{f^2} \quad \text{and} \quad \sigma_2^2 = \frac{1}{2} \frac{\sigma_f^2}{f^2} \left(1 + \frac{\tan^2 \theta}{12}\right) \quad (9)$$

where  $u^2$  and  $x^2$  are the average values of image velocities and feature positions. We can compare the two terms contributing to bias:  $\frac{1-x^2}{2} \frac{\sigma_u^2}{f^2}$  and  $\frac{u^2}{2} \frac{\sigma_f^2}{f^2} \left(1 + \frac{\tan^2 \theta}{12}\right)$ . Since  $1 - x^2$  is on the average negative, the two bias terms have opposite signs and dampen (or under favorable circumstances cancel) each other. This is demonstrated in Figure 1a where we computed the bias as a function of the calibration error assuming 70% error in the measurements of the optical flow. Similarly for 30% error in the optical flow, Figure 1b shows how the errors in calibration dominate the precision of the translation estimates. The details of the simulations are outlined in Section 5.

**FOV dependency.** Equation (9) also reveals the dependence of the motion estimates on the field of view  $\theta$  of the camera. As  $FOV \rightarrow 0$  the term  $1 - x^2$  becomes 'more' negative generating a stronger bias as previously shown [18]. The term  $\sigma_2^2$  arising from the noisy calibration also increases for bigger  $FOV$ . This effect is demonstrated in Figure 2 where we calculated the bias in  $v_z$  ( $z$ -component of the linear velocity) as a function of  $FOV$  for noiseless optical flow and 30 % error on the calibration parameters.

## 4 3D Structure Distortion

In the following section we study the effect of the calibration errors on the resulting 3D structure. In order to separate the sensitivity issues pertaining to the translation bias from those related to the structure of the scene we assume that the motion is purely translational in the direction of the optical axis. Given the translation estimate  $v_z$  the unknown structure is given by the equation (10). Assuming perfect calibration the estimation of the depth  $Z$  is affected both by errors on  $(v_z, x, \dot{x}, f)$  and the translation bias  $\delta v_z$  in the estimation of  $v_z$  itself. This bias will induce an error in the scene depth in the following way:

$$Z = \frac{v_z x}{\dot{x} f} \quad \text{and} \quad \delta Z = \frac{v_z \|\mathbf{x}\|}{\|\mathbf{u}\| f} \delta v_z \quad (10)$$

We see that the bias in depth  $\delta Z$  is proportional to the distance of the feature from the optical center with the same sign as  $\delta v_z$ . In the case  $v_z$  is overestimated, the further points are from the optical axis, the further they are projected. The effect of calibration errors on the resulting structure can be demonstrated by means of error propagation:

$$\sigma_Z^2 = \frac{\|\mathbf{x}\|^2 v_z^2}{\|\mathbf{u}\|^2 f^2} \left( \frac{\sigma_{v_z}^2}{v_z^2} + \frac{\sigma_f^2}{f^2} + \frac{\sigma_o^2}{\|\mathbf{x}\|^2} + \frac{\sigma_u^2}{\|\mathbf{u}\|^2} \right) \quad (11)$$

We see that the error is again proportional to  $\|\mathbf{x}\|$ , so the features far from the optical center are affected more. In [10] we also showed that  $\delta v_z \gg \sigma_{v_z}$  which implies that the

errors due to both calibration and optical flow mostly tend skew the structure than dither it. This conclusion is independent of the algorithm used for the reconstruction and holds for general camera displacement and rotations. Figures 3 and 4 show the skew for a general motion on synthetic and real sequences.

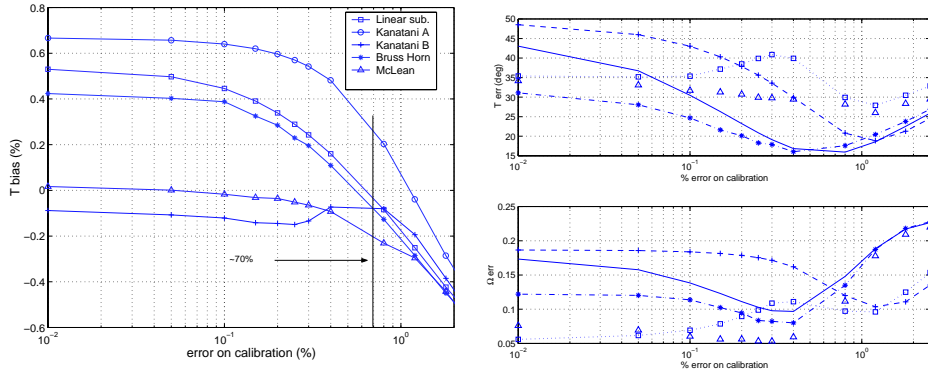


Figure 1: **(a)** Relationship between translation bias as a function of calibration error, assuming that the measurement errors were fixed at 70%. Note that the bias is approximately 0 for an error of calibration parameters of about 70%. This is the situation when the two bias terms cancel each other. *FOV* for these experiments was 90°. **(b)** The bias and sensitivity of translation estimates expressed in angular units as a function of the calibration error for a optical flow noise level of 30%. Note that the bias tends to decrease for amounts of noise smaller compared with the 30% noise on optical flow and increase again when the calibration error becomes dominating.

## 5 Experiments

The experiments verifying the observations derived in the previous sections were performed both on synthetic and real sequences. For synthetic data we adopted approach pursued in [18]. The simulations were performed for random clouds of 100 points. The focal length was 1 and the focal plane dimension of 512×512 pixels. The *FOV*, unless otherwise stated, was chosen to be 90°. Motion was realized for various combinations of translation and rotation. Zero-mean Gaussian noise was added to the components of image velocities and the standard deviation of such noise was chosen to be constant. The rotational speed was constant chosen to be 0.23° per frame. The magnitude of the linear velocity was chosen in order to fixate the point at the center of the random cloud of points. The tested algorithms were those of Horn [17], Heeger and Jepson [5], Kanatani [3] and MacLean [2], with the implementations made available by [18]. The tested algorithms differ in the choice of objective function, leading to linear or nonlinear optimization problems. The linear algorithms of [2] and [3] provide a solution for correcting the bias due to errors in temporal matching. Bias and sensitivity were measured over one thousand trials. The camera was first calibrated off-line several times using the Matlab calibration toolbox [15] and planar calibration grid, to justify the assumptions about the average precision that can be commonly obtained. The skew results were approximately zero in all the tests done. The reconstruction of unknown depths was obtained as LLSE problem,

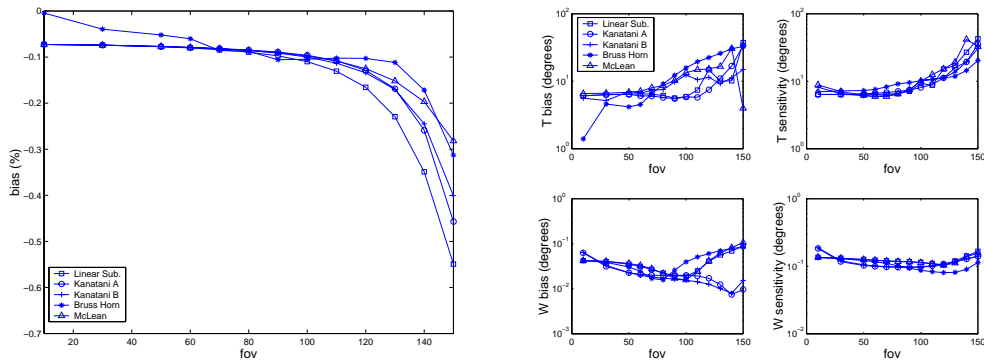


Figure 2: (a) Dependence of the bias due to noisy calibration of the  $FOV$  of the camera. 30% noise on camera parameters is generated while optical flow is noiseless. The magnitude of the bias increases with increasing  $FOV$ . (b) Dependence of velocity estimation on  $FOV$  for 30% error on calibration parameters and noise-free tracking.

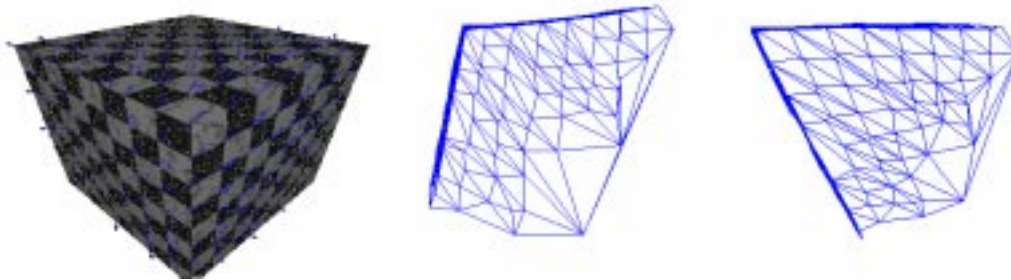


Figure 3: Synthetic sequence : Original model (top figures), model distorted by an underestimated focal length (bottom-left), model distorted by an overestimated focal length (bottom-right).

with automated rescaling and is described in [6]. The experiments on both real and simulated sequences are in Figures 3 and 4. The figures demonstrate the effect of over and under estimation of the focal length on the computed 3D structure. The flattening and elongation are clearly noticeable.

## 6 Conclusions

The analysis carried out in this paper demonstrates the sensitivity of structure and motion estimation with respect to the errors in camera calibration. As the main contribution we demonstrated that the calibration errors introduce an additional bias in the direction of the optical axis with the opposite sign to the one produced by errors due to temporal matching. Under favorable circumstances the two bias terms cancel each other leading to unbiased estimates. Moreover the bias produced by erroneous calibration increases in magnitude for increasing  $FOV$ . This is in contrary to what happens to the bias produced





Figure 4: Original model (left) with estimated  $v_z = 0.5530$ ; model distorted by an underestimated focal length by 50% (center) with  $v_z = 0.7597$ ; model distorted by an overestimated focal length by 50% (right) and  $v_z = 0.3975$ .

by noisy image velocities, which in addition to that, depends on the direction of translation. The relationship between  $FOV$  and the errors caused by noisy calibration is mostly independent of the choice of translation direction. The linear velocity bias propagates to the structure estimation and distorts the resulting 3D structure. The observed distortions are proportional to the distance of the feature from the optical axis. This is shown analytically for a simple motion configuration and tested on real sequences on variety of motions in order to justify the generality of the previous assessment.

The above observations were derived analytically, resorting to some approximations, in the context of linear techniques. More extensive simulations confirmed the reasonability of the approximations and justified the intrinsic nature of the calibration errors independent of the algorithm choice.

## Appendix A

By propagating the errors in calibration to image positions and image velocities we present here covariance matrices for noisy feature positions and noisy velocities. The noisy feature position  $\tilde{\mathbf{x}} \sim N(0, \Sigma_x)$  and noisy image velocity  $\tilde{\mathbf{u}} \sim N(0, \Sigma_u)$  are normal random variables, with the following covariance matrices:

$$\Sigma_x \approx \frac{1}{2} \frac{\Delta^2}{f^2} \begin{pmatrix} \frac{\sigma_f^2}{f^2} + \frac{\sigma_v^2}{\Delta^2} & \frac{\sigma_f^2}{f^2} \\ \frac{\sigma_f^2}{f^2} & \frac{\sigma_f^2}{f^2} + \frac{\sigma_v^2}{\Delta^2} \end{pmatrix} \quad \Sigma_u \approx \frac{1}{2} \frac{u^2}{f^2} \begin{pmatrix} \frac{\sigma_f^2}{u^2} + \frac{\sigma_f^2}{f^2} & \frac{\sigma_f^2}{f^2} \\ \frac{\sigma_f^2}{f^2} & \frac{\sigma_f^2}{v^2} + \frac{\sigma_f^2}{f^2} \end{pmatrix} \quad (12)$$

where  $\Delta^2$  is the average distance of the features from the center of the projection under the assumption that the features are uniformly distributed in the image plane ( *i.e.* on the average  $\Delta_x^2 \approx \Delta_y^2$ ) and  $u^2 = u_x^2 + u_y^2$  is the average image velocity.

## References

- [1] Barron J. L. , Fleet D. J. , Beuchemlein, S. S. Performance of Optical Flow Techniques, *IJCV*, Vol 12:1, pp. 43-77, 1994.
- [2] MacLean W. J. Removal of Translational Bias when using Subspace Methods, *Proceedings from ICCV*, pp. 753-758, 1999.

- [3] Kanatani K. Renormalization for Unbiased Estimation, *Proceedings from ICCV*, pp. 599-606, 1993.
- [4] Kanatani K. Geometric Computation for Machine Vision. *Oxford University Press*, 1993.
- [5] Heeger D. J. , Jepson A. D. Subspace Methods for Recovering Rigid Motion, *IJCV*, Vol 7:2, pp. 95-117, 1992.
- [6] Zucchelli M. , Christensen H. I. Flow Based Structure from Parallax with Automatic Rescaling, *in submission to BMVC*.
- [7] Grossman E. Santos-Victor J. Uncertainty Analysis of 3D Reconstruction from Uncalibrated Views, *Image and Vision Computing* Vol 18, pp. 685-696, 2000.
- [8] Svodoba T. Sturm P. What can be done with a badly calibrated Camera in Ego-motion estimation?, *Research Report CTU-CMP-1996-01* Czech Technical University.
- [9] Ma Y., Kosecka J., Sastry S. Linear Differential Algorithm for Motion Recovery:A Geometric Approach. *IJCV* Vol. 36 pp. 71-89, January 2000.
- [10] Kosecka J., Zucchelli M. Motion Bias and Structure Distortion induced by Calibration Errors, *George Mason University TR-04-01*.
- [11] Cheong L.-F.and Peh Ch.-H. Characterizing Depth Distortion due to Calibration Uncertainty. *Proceedings of ECCV, 2000*, Vol. 2 pp. 665-667.
- [12] Bognoux S. From Projective to Euclidean Space Under any Practical Situation, a Criticism of Self-Calibration. *Proceedings of ICCV, 1998*, Vol. 2 pp. 790-795.
- [13] K. Danilidis and M. Spetsakis. Understanding noise sensitivity in structure from motion. *Visual Navigation*, Eds: Y. Aloimonos, pp. 61-88, 1996.
- [14] Oliensis J. A Structure from Motion Ambiguity. *PAMI* Vol. 22, No. 7, pp. 685-700, 2000.
- [15] [http : //www.vision.caltech.edu/bouguetj/calib\\_doc/index.html](http://www.vision.caltech.edu/bouguetj/calib_doc/index.html)
- [16] Zhang T., Tomasi C. Fast, Robust, and Consistent Camera Motion Estimation. *Proceedings of CVPR*, pp. 164-170, 1999.
- [17] Bruss A.R., Horn B.K.P. Passive Navigation. *CVGIP* 21:3-20, 1983.
- [18] Tian T.Y., Tomasi C., Heeger D. J. Comparison of Approaches to Egomotion Computation. *CVPR* , pp. 315-320, 1994.