Characterizing Depth Distortion Due to Calibration Uncertainty

Loong-Fah Cheong and Chin-Hwee Peh

Department of Electrical Engineering
National University of Singapore
10 Kent Ridge Cresent, Singapore 119260
Tel: 65-874-2290 Fax: 65-779-1103
{eleclf, elepehch}@nus.edu.sg

Abstract. There have been relatively little works to shed light on the effects of errors in the intrinsic parameters on motion estimation and scene reconstruction. Given that the estimation of the extrinsic and intrinsic parameters from uncalibrated motion apts to be imprecise, it is important to study the resulting distortion on the recovered structure. By making use of the iso-distortion framework, we explicitly characterize the geometry of the distorted space recovered from 3-D motion with freely varying focal length. This characterization allows us: 1) to investigate the effectiveness of the visibility constraint in disambiguating uncalibrated motion by studying the negative distortion regions, and 2) to make explicit those ambiguous error situations under which the visibility constraint is not effective. An important finding is that under these ambiguous situations, the direction of heading can nevertheless be accurately recovered and the structure recovered experienced a well-behaved distortion. The distortion is given by a relief transformation which preserves ordinal depth relations. Thus in the case where the only unknown intrinsic parameter is the focal length, structure information in the form of depth relief can be obtained. Experiments were presented to support the use of the visibility constraint in obtaining such partial motion and structure solutions.

Keywords: Structure from motion, Depth distortion, Space perception, Uncalibrated motion analysis.

1 Introduction

While there have been various works on the self-calibration problem, most face difficulties in estimating the intrinsic parameters accurately. Two courses are open to researchers. One approach is to enforce special camera displacements to obtain better estimates of the intrinsic parameters [1,5,11,16,20]. Another approach argues that as far as scene reconstruction is concerned, several weaker structures (Projective, Affine) can be obtained without complete recovery of the intrinsic parameters. While the mainstay of the research efforts adopts the discrete approach, [3,20] have recently formulated the problem in the continuous

D. Vernon (Ed.): ECCV 2000, LNCS 1842, pp. 664–677, 2000.

domain. Most of the schemes presented assume that the intrinsic parameters across the frames are constant. A more general treatment of the problem, allowing for varying intrinsic parameters, is given in [2,3,10,12,15,17,20].

There have been relatively little works to shed light on the effects of errors in the intrinsic parameters on motion estimation and scene reconstruction. Florou and Mohr [8] used the statistic approach to study reconstruction errors with respect to calibration parameters. Svoboda and Sturm [18] studied how uncertainty in the calibration parameters gets propagated to the motion parameters. Viéville and Faugeras studied the partial observability of rotational motion, calibration, and depth map in [20]. Bougnoux [2] offered a critique of the self-calibration problem, finding that the estimation of various intrinsic parameters are unstable. However, it was observed, partly empirically, that despite uncertainty in the focal length estimation, the quality of the reconstruction does not seem to be affected. Certain geometrical properties such as parallelism seemed to be preserved. Aside from this observation, there has not really been an in-depth geometrical characterization of the errors in the reconstructed depth given some errors in both the intrinsic and the extrinsic parameters.

In this paper, we consider the common situation where all of the intrinsic parameters are fixed except the focal length. The focal length can be freely varying across frames, resulting in a zoom field (considering infinitesimal motion) which is difficult to separate from that of a translation along the optical axis. This, together with the perennial problem of the coupling between translation and rotation, means that distortion in the recovered structure is likely to be present.

This paper attempts to make the geometry of this distortion explicit by using the iso-distortion framework introduced in [4]. The motivation for performing this analysis is twofold: first, to seek to characterize the distortion in the perceived depth; second, to extend our understanding on how depth distortion in turn interacts with motion (including zoom) estimation. It is an alternative look at the problem of depth representation from the usual stratified viewpoint [6], but one that will inform one another.

This paper is structured along the following lines. First comes some preliminaries regarding the iso-distortion framework in Section 2, followed by an extension of this framework to the self-calibration problem. Several major features of the resulting distortion are then made explicit. The main goals of Section 3 are (1) to elaborate the relations between the depth distortion and the estimation of both the intrinsic and extrinsic parameters; and (2) to study certain well-behaved depth distortion resulting from ambiguous solutions. Section 4 presents experiments to support the use of the visibility constraint in obtaining partial solutions to the estimation of both motion and structure. The paper ends with a summary of the work.

2 Iso-Distortion Framework

2.1 Pre-requisites

If a camera is moving rigidly with respect to its 3D world with a translation (U, V, W) and a rotation (α, β, γ), together with a zooming operation, the resulting optical flow (u, v) at an image location (x, y) can be extended from its basic form to include the following terms:

$$u = \frac{W}{Z}(x - x_0) + \frac{xy}{f}\alpha - f(1 + \frac{x^2}{f^2})\beta + \gamma y + \frac{\dot{f}}{f}x \tag{1}$$

$$v = \frac{W}{Z}(y - y_0) - \frac{xy}{f}\beta + f(1 + \frac{y^2}{f^2})\alpha - \gamma x + \frac{\dot{f}}{f}y \tag{2}$$

where f is the focal length of the camera;
$\dot{f}$ is the rate of change of the focal length;
$(x_0, y_0) = (f\frac{U}{W}, f\frac{V}{W})$ is the Focus of Expansion (FOE) of the flow field;
and Z is the depth of scene point.

2.2 Space Distortion Arising from Uncalibrated Motion

In a recent work [4], the geometric laws under which the recovered scene is distorted due to some errors in the viewing geometry is represented by a distortion transformation. It was called the iso-distortion framework whereby distortion in the perceived space can be visualized by families of iso-distortion lines. In the present study, this framework has been extended to characterize the types of distortion experienced by a visual system where a change in the focal length may result in further difficulties and errors in the estimation of its calibration.

From the well-known motion equation, the relative depth of a scene point recovered using normal flow with direction (n_x, n_y) may be represented by

$$Z = \frac{(x - x_0, y - y_0).(n_x, n_y)}{(u_n - (u_r + u_f).(n_x, n_y))} \tag{3}$$

where u_n is the normal flow magnitude;
u_r is the rotational flow; and
u_f is the zoom flow caused by a change in the focal length.

If there are some errors in the estimation of the intrinsic or the extrinsic parameters, this will in turn cause errors in the estimation of the scaled depth, and thus a distorted version of space will be computed. By representing the estimated parameters with the hat symbol ($\hat{}$) and errors in the estimated parameters with the subscript e (where error of any estimate p is defined as $p_e = p - \hat{p}$), the estimated relative depth $\hat{Z}$ may be expressed in terms of the actual depth Z as follows:

$$\hat{Z} = Z\left(\frac{(x-\hat{x}_0)n_x + (y-\hat{y}_0)n_y}{(x-x_0, y-y_0).(n_x,n_y) + u_{re}.(n_x,n_y)Z + u_{fe}.(n_x,n_y)Z + NZ}\right) \quad (4)$$

where $u_{re} = u_r - \hat{u}_r$
$= (\frac{xy}{f}(\alpha-\hat{\alpha}) - f(1+\frac{x^2}{f^2})(\beta-\hat{\beta}) + y(\gamma-\hat{\gamma}),$
$\frac{-xy}{f}(\beta-\hat{\beta}) + f(1+\frac{y^2}{f^2})(\alpha-\hat{\alpha}) - x(\gamma-\hat{\gamma}));$
$u_{fe} = u_f - \hat{u}_f = (\frac{\hat{f}-\dot{f}}{f}x, \frac{\hat{f}-\dot{f}}{f}y)$; and
N is a noise term representing error in the estimate for normal flow value.

Equation (4) shows that errors in the motion estimates distort the recovered relative depth by a factor D, given by the terms in the bracket:

$$D = \frac{(x-\hat{x_0})n_x + (y-\hat{y_0})n_y}{(x-x_0, y-y_0).(n_x,n_y) + u_{re}.(n_x,n_y)Z + u_{fe}.(n_x,n_y)Z + NZ} \quad (5)$$

Equation (5) describes, for any fixed direction (n_x, n_y) and any fixed distortion factor D, a surface $f(x, y, Z) = 0$ in xyZ-space, which has been called the iso-distortion surface. For specific values of the parameters x_0, y_0, $\hat{x}_0$, $\hat{y}_0$, α_e, β_e, γ_e, f, $\dot{f}$, $\hat{\dot{f}}$ and (n_x, n_y), this iso-distortion surface has the obvious property that points lying on it are distorted in depth by the same multiplicative factor D. The distortion of the estimated space can be studied by looking at these iso-distortion surfaces. In order to present the these analyses visually, most of the investigation will be conducted by initially considering (n_x, n_y) to be in the horizontal direction. Ignoring the noise term N, we get the following set of equations for different values of D:

$$\begin{cases} x = \hat{x_0} & \text{if } D = 0 \\ Z = \frac{x_0 - x}{u^x_{re} + u^x_{fe}} & \text{if } D \to \pm\infty \\ Z = \frac{1-D}{D}(\frac{x-x_0}{u^x_{re}+u^x_{fe}}) + \frac{1}{D}(\frac{x_{0e}}{u^x_{re}+u^x_{fe}}) & \text{otherwise} \end{cases} \quad (6)$$

where the superscript x indicates the projection of vector onto the horizontal direction.

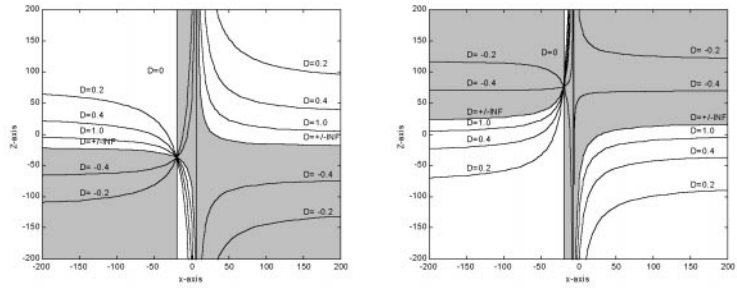
Now, if we were to consider the field of view of the camera to be small and ignore the effect of γ_e so that u_{re} becomes $(-\beta_e f, \alpha_e f)$, we have

$$\begin{cases} x = \hat{x_0} & \text{if } D = 0 \\ Z = \frac{x_0 - x}{-f\beta_e + \delta_e x} & \text{if } D \to \pm\infty \\ Z = \frac{1-D}{D}(\frac{x-x_0}{-f\beta_e+\delta_e x}) + \frac{1}{D}(\frac{x_{0e}}{-f\beta_e+\delta_e x}) & \text{otherwise} \end{cases} \quad (7)$$

where $\delta_e = \frac{\hat{f}-\dot{f}}{f}$.

These equations describe the iso-distortion surfaces as a set of surfaces perpendicular to the x-Z plane. Much of the information that equations (7) contain

can thus be visualized by considering a family of iso-distortion contours on a two-dimensional x-Z plane. Each family is defined by four parameters: $x0$ and the three error terms $x0_e$, β_e and δ_e. Within each family, a particular D defines an iso-distortion contour. Figure 1 corresponds to two particular cases of the iso-distortion contours. In the next subsection, we shall determine the salient geometrical properties of the iso-distortion contours.



(a) (b)

Fig. 1. Iso distortion contours for (a) $x_0 = 30$, $x_{0e} = 50$, $\beta_e = 0.001$, $\delta_e = 0.05$ and $f = 350$. (b) When $\delta_e = -0.05$. (Shaded region corresponds to the negative depth region.)

2.3 Salient Properties

Several salient features can be identified from the plot.

1) The $D = 0$ curve is a vertical line that intersects the x-axis at $\hat{x_0}$. Any change in the estimated FOE will slide this line along the x-axis.
2) The $D = \pm\infty$ curve intersects the x-axis at x_0 and approaches to $Z = \frac{1}{\delta_e}$ as x tends to infinity. The structure of the $D = \pm\infty$ contour is independent of the position of the estimated FOE but on the true FOE location.
3) The contours intersect at a singular point where $x = \hat{x_0}$ and $Z = \frac{x_{0e}}{\delta_e \hat{x_0} - f\beta_e}$. At this point, the depth of the scene is undefined.
4) The vertical asymptote for all contours where $D \neq 0$ is $x = \frac{f\beta_e}{\delta_e}$.
5) The horizontal asymptote for each contours where $D \neq 0$ is the line $Z = \frac{1-D}{D\delta_e}$. Hence, each contour has a different horizontal asymptote depending on the values of D. However, the horizontal asymptote for the contour $D = 1$ is always the x-axis, independent of other parameters. A diminishing value in δ_e simultaneously moves the horizontal and vertical asymptotes away from the image center. It approaches the iso-distortion configuration under the case of calibrated motion [4].

3 What Can the Distortion Contours Tell Us?

3.1 Confusion between Translation, Rotation, and Zoom

It is well-known that we cannot numerically raise the ambiguity between several intrinsic and extrinsic parameters. For instance, there is a strong ambiguity between a translation along the optical axis and a zoom. In this subsection, we briefly look at the numerical aspect of this problem before studying how the iso-distortion framework can be utilized to yield useful information.

One straightforward way of estimating the motion parameters would be to hypothesize a FOE. Flow vectors are then projected to the gradient direction $\boldsymbol{n}^i$ perpendicular to the emanated lines from the hypothesized FOE so that they contain only the rotational and the zoom field (if the hypothesized FOE were correct). The least square formulation is as follows:

$$\begin{bmatrix} (\frac{x^i y^i}{f}, f(1+\frac{y^{i2}}{f^2})).\boldsymbol{n}^i & (-f(1+\frac{x^{i2}}{f^2}), -\frac{x^i y^i}{f}).\boldsymbol{n}^i & (y^i, -x^i).\boldsymbol{n}^i & (x^i, y^i).\boldsymbol{n}^i \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \frac{\dot{f}}{f} \end{bmatrix} = \begin{bmatrix} (u^i, v^i).\boldsymbol{n}^i \\ \vdots \end{bmatrix} \tag{8}$$

The conjecture is that, for the correct FOE candidate, the residual should be the smallest among all candidates, since the least square equations correctly model the situation. Thus, the least square residual furnishes a feasible measure upon which to base the FOE candidate selection. However, this formulation has several problems. Consider the true FOE candidate: as far as this candidate is concerned, the last column of the the matrix in Equation (8) can be rewritten as $(x_0, y_0).\boldsymbol{n}^i$ since in this case $(x^i, y^i).\boldsymbol{n}^i = (x_0, y_0).\boldsymbol{n}^i, \forall i$. Thus, when the FOE is at the image center (i.e. (x_0, y_0)=(0,0)), the least square estimation in Equation (8) becomes rank deficient and this gives an infinite set of solutions. Hence, from the least-square fitting formulation, we are not able to separate a pure forward translation from a zoom. Now, if the FOE lies near to the image center, the least square estimation becomes unstable, which can be analyzed by using the concept of condition number. Similar conclusions can be derived for various other algorithms, for instance that of Brooks [3].

Thus, the least-square fitting residual cannot be the final criteria for choosing the correct FOE. Rather, we have to look for additional constraint to prune the set of possible solution candidates.

3.2 The Visibility Constraint

Direct motion algorithms [13,7] often attempt to find the solution by minimizing the number of negative depth found. This is known as the visibility constraint.

However its usage in the estimation of uncalibrated motion is relatively unexplored. We would now like to examine this constraint in the light of the negative distortion region. The geometry of the negative distortion region allows us to examine these questions: does the veridical solution have the minimum number of negative depth? Are there combination of estimation errors such that the visibility constraint is not sufficient to discriminate between alternative solutions? Do these ambiguous solutions exhibit any peculiar properties in terms of their recovered structure or their motion estimates? To study these questions, we consider the distortion plot for the horizontal flow first, before considering those in other directions.

3.3 Constraints on Motion Errors

Comparing Figures 1 (a) and (b), the first observation is that if a particular solution has a negative δ_e, the distortion plot will be such that the majority of the negative distortion regions lie in front of the image plane. Irrespective of the actual scene structure, there will be a large number of negative depth estimates obtained, thereby ruling out that particular solution. Thus the first condition on the motion errors for ambiguity to arise is that the zoom error flow must be such that:

$$\delta_e > 0 \tag{9}$$

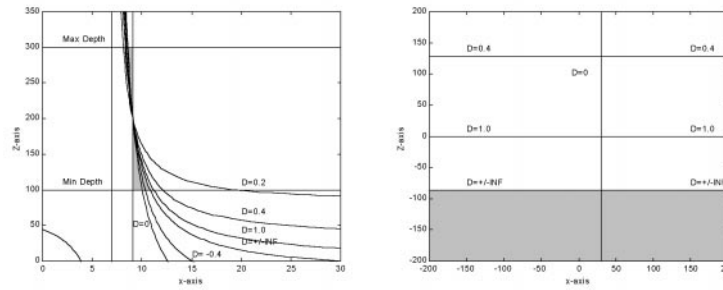
in which case most of the negative distortion region lie behind the image plane. What remains in front of the image plane is a band of negative distortion region, bounded by two contours, the $D = 0$ and the $D = -\infty$ contours, whose equations are respectively $x = \hat{x}_0$ and $Z = \frac{x_0 - x}{-f\beta_e + \delta_e x}$. The latter cuts the horizontal axis at $x = x_0$ and its vertical asymptote is given by $x = \frac{f\beta_e}{\delta_e}$. We now derive the combination of errors such that this negative band will be minimized (i.e. ambiguity is maximized).

To derive these combinations, we first arbitrarily fix the error δ_e and suppose it satisfies the constraint given in (9). The constraints on the other parameters $\hat{\beta}$ and $\hat{x}_0$ that will yield minimum negative distortion region depend on whether an algorithm solves for these parameters separately or simultaneously:

1) If $\hat{\beta}$ is solved first and the estimate contains an error β_e, then the $\hat{x}_0$ that minimizes the negative depth region, given these fixed δ_e and β_e, is:

$$\hat{x}_0 = \frac{2x_0 + (Z_{max} + Z_{min})f\beta_e}{2 + (Z_{max} + Z_{min})\delta_e}$$

where we have assumed that depths in the scene are uniformly distributed between Z_{min} and Z_{max}. See Figure 2(a). This $\hat{x}_0$ always lies between x_0 and the vertical asymptote. Similar condition on β_e can be derived if we solve for $\hat{x}_0$ first.



(a) (b)

Fig. 2. Configuration for Minimum Negative Depth. (a)fixed δ_e and β_e. (b) fixed δ_e.

2) If both $\hat{\beta}$ and $\hat{x}_0$ are solved together, then the solution that minimizes the negative depth region is given by

$$x_0 = \hat{x}_0 = \frac{f\beta_e}{\delta_e}$$

that is, the x-component of the direction of heading is recovered veridically, and the vertical asymptote $x = \frac{f\beta_e}{\delta_e}$ coincides with the line $x = x_0 = \hat{x}_0$ so that the $D = 0$ and the $D = -\infty$ contours coincide. In this case, the negative band in front of the image plane vanishes. See Figure 2(b).

3) Furthermore, if the lines $x = x_0, x = \hat{x}_0$ and $x = \frac{f\beta_e}{\delta_e}$ are out of the image (and on the same side), then even if they do not meet, the negative distortion band would be outside the field of view. Thus, this solution will not yield any negative depth estimates and would be totally ambiguous too.

To both summarize and to complete the analysis, we consider flows in any other gradient direction. The preceding conditions can be generalized as:

$$(x_0, y_0) = (\hat{x}_0, \hat{y}_0) = \frac{(\beta_e, -\alpha_e)f}{\delta_e}; \qquad \delta_e > 0 \tag{10}$$

a solution which correctly estimates the direction of heading. $(\beta_e, -\alpha_e)$ must be in the same direction as (x_0, y_0) so that for any gradient direction (n_x, n_y), the condition $\frac{f(\beta_e, -\alpha_e).(n_x, n_y)}{\delta_e} = (x_0, y_0)$ can be satisfied for the same δ_e. In this case, the negative distortion region vanishes in front of image plane for all gradient directions.

3.4 Distortion of Recovered Structure

The preceding analysis shows that the use of the visibility constraint does not lift the ambiguities that exist among various kinds of motions. However it does restrict the solution set so that those yielding the minimum negative depth estimates possess certain nice properties, such as the direction of heading is correctly

estimated. Furthermore, as can be seen from Figure 2, the iso-distortion contours become horizontal, resulting in well-behaved distortion. Indeed, the distortion factor D in this case has the form of a relief transformation $\frac{1}{a+bZ}$, where $a = 1$ and $b = \delta_e$. This relief transformation preserves the ordering of points; its general properties were recently discussed and analyzed by [9,14]. As a result of this distortion, the reconstructed scene may appear visually perfect even though the depths have been squashed to various degrees. It is of interest to compare this result with that demonstrated by Bougnoux [2]: that the uncertainty on the focal length estimation leads to a Euclidean calibration up to a quasi anisotropic homothety, which in turn yields visually good-looking reconstruction.

4 Experiments

This section presents the experiments carried out to support the theoretical findings established in the preceding section. Specifically, we demonstrate the ability to correctly estimate the heading direction of the camera based on minimizing the number of negative depth estimates. The distortion effects due to erroneous motion estimates on simple surfaces were also tested. In our experiments, both synthetic and real images were used.

4.1 Synthetic Images

A set of noise-free synthetic images with dimension 240 pixels by 320 pixels were generated. The focal length of the projection was fixed at 600 pixels. This gave a viewing angle of near 30^o. Three simple planes with different orientation at different 3-D depth were constructed in the image. The true FOE was located at $(65, 0)$ of the image plane and the rotational parameters (α, β, γ) have the values $(0, -0.00025, 0)$. There was no change in the focal length (i.e. $\frac{\dot{f}}{f} = 0$).

We first arbitrarily fixed the error δ_e to be some positive number. We then solved for the rotational parameter $(\hat{\beta})$ and the FOE $(\hat{x_0})$ in the following manner: For each hypothesized $\hat{\beta}$, we selected the best $\hat{x_0}$ candidate such that the minimum number of negative depth estimates was obtained. The search range for $\hat{\beta}$ lies between -0.005 to 0.005. The top ten candidates with the least amount of negative depth estimates are tabulated in Table 1.

In this experiment, all ten candidates gave no negative depth estimates. One possible explanation is that there is a lack of depth variations in the synthetic image (the range of the motion flow lies between 0.000167 to 1.170667). Another observation is that the selected $\hat{x_0}$ for any $\hat{\beta}$ always resides between the vertical asymptote $\frac{f\beta_e}{\delta_e}$ and the true FOE x_0.

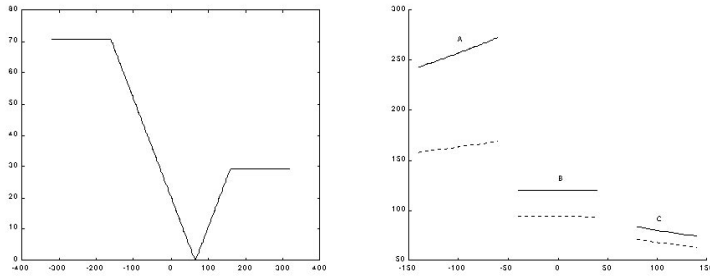
Figure 3(a) depicts the variations of the percentage of negative depth estimates as a function of the estimated $\hat{x_0}$ when the vertical asymptote is at the veridical FOE position. The figure corroborates our theoretical predictions that the least amount of negative depth estimates (in this case 0) is obtained when $x_0 = \hat{x_0}$. Similar curves are obtained when the vertical asymptote is not at

the veridical FOE position. In theory, if there were sufficient feature points, the minimum points of these curves should be higher than those of the former.

Using the erroneous motion estimates $\hat{x_0}$, $f\hat{\beta}$ and $\hat{\delta}_e$ that resulted in the least amount of negative depth estimates, we attempted to reconstruct the synthetic planes. Figure 3(b) shows the plan view of the three synthetic planes, together with their reconstructed versions. It can be seen that the relief of the plane remained unchanged after the transformation, i.e., the ordinal depth were preserved. Note that the metric aspect of the plane orientations (their slants) was altered. This change can be related to the calibration uncertainties via the complex rational function given in Equation (4).

Table 1. Top ten candidates in the synthetic image experiment. $x_0 = 65$.

$\frac{f\beta_e}{\delta_e}$	58	59	60	61	62	63	64	65	66	67
$\hat{x_0}$	63	64	64	64	64	65	65	65	65	66



(a) (b)

Fig. 3. (a) Variations of amount of negative depth estimates as a function of $\hat{x_0}$ for synthetic images. (b) Synthetic planes (A,B and C) with their reconstructed versions (dotted lines)

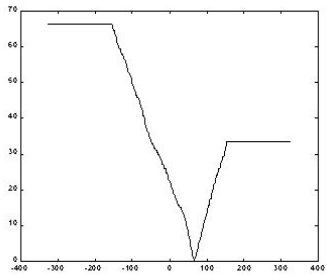
4.2 Real Images

For real images, we used a sequence whose dimension has been scaled down to 287 pixels by 326 pixels. The focal length and the field of view of the camera were respectively 620 pixels and approximately 30^o. The rotational parameter (α, β, γ) was (-0.00013,-0.00025,0) and $\dot{f} = 0$. The true FOE was at image location (65,73). A similar experimental procedure to the synthetic experiment was applied to the real images. The top 10 candidates obtained are shown in Table 2.

When the vertical asymptote was at 64, the variations of the percentage of negative depths obtained as a function of the estimated $\hat{x_0}$ is plotted in Figure 4(a). We obtained a similar curve to the synthetic case. However, due to the presence of noise, the minimum amount of negative depth obtained was not zero and the best $\hat{x_0}$ was located at 64 (about 1 pixel from the true x_0). The procedure was repeated to locate the y-component of the FOE using the vertical component of the normal flow vectors. The positions of the true and estimated FOE are plotted in Figure 4(b).

Table 2. Top ten candidates for the real image experiment. $x_0 = 65$.

$\frac{f\beta_e}{\delta_e}$	62	61	63	59	60	64	45	75	67	76
$\hat{x_0}$	64	63	65	61	62	66	46	78	69	79
$NegDepth(\%)$	0.097	0.110	0.117	0.130	0.145	0.147	0.152	0.154	0.157	0.159



(a)

(b)

Fig. 4. (a) Variations of negative depth as a function of $\hat{x_0}$ for real images. (b) Positions of true and estimated FOE in the real image. (+ : true FOE at (65,73); × : Estimated FOE at (64,74))

4.3 Discussions

The results obtained seem to corroborate the various predictions made in this paper. In particular, while the use of visibility constraint cannot be used to effect a full recovery of all the parameters, minimizing the number of negative depth estimates do result in certain nice properties of the solutions. It seems that at least in the case where the only unknown intrinsic parameter is the focal length, structure information in the form of depth relief can be obtained from the motion cue. The reconstructed depths did look visually alright due to the presentation

of the depth relief. The results also validate the assumptions made in this paper (that quadratic terms in the flow can be discarded), at least for field of view up to 30^o.

There are many problems that plague real image. Foremost among these is the presence of noise. The effect of any noise N at a particular image pixel is to replace the term $f\beta_e$ in the numerator of the vertical asymptote $x = \frac{f\beta_e}{\delta_e}$ by $f\beta_e - N$. Thus, to this particular image point, its effective vertical asymptote has shifted and part of the problem lies in that this shift has different effects on the various solution candidates. For the case of those solutions where the negative distortion region in front of the image plane would have vanished under noiseless conditions, this noise-induced shift away from $x = x_0$ may result in that particular depth estimate becoming negative again (depending on the sign and magnitude of that N). For the case of other solutions, this shift may have the contrary effect of moving that point out of the negative distortion region. It becomes plausible that the "desired" solutions (i.e. those satisfying (10)) may not have the minimum number of negative depth estimates. Thus, the overall effect of noise is to reduce the effectiveness of the visibility constraint in getting the "desired" solutions. In our experiment, we found that as long as $f\beta_e$, the error in the rotational flow, is large enough so that $f\beta_e - N \approx f\beta_e$, the location of the FOE (x_0, y_0) could be determined quite accurately.

Other confounding factors for real images include the sparse distribution of scene features. It holds that while the underlying negative distortion region may have increased in size, there may not be any increase in the number of negative depth estimates, due to a lack of scene point residing in the negative distortion regions. Evidently, under such circumstances, the number of negative depth estimates may not exhibit a monotonic increase as the error in the FOE increases. Inspecting the top candidates selected in Table 2, we observed that this is indeed the case as $\hat{x}_0$ moves away from x_0.

5 Conclusions and Future Directions

This paper represents a first look at the distortion in the perceived space resulting from errors in the estimates for uncalibrated motion. The geometry of the negative distortion region allows us to answer questions such as whether the visibility constraint is adequate for resolving ambiguity. It is also found that while Euclidean reconstruction is difficult, the resulting distortion in the structure satisfies the relief transformation, which means that ordinal depth is preserved.

A concluding caveat is in order concerning real zoom lens operation. The principal point often changes when the focal length varies. Hence, an analysis based on a more detailed distortion model should be carried out. Our iso-distortion model can be readily extended to take into account these changes and this would be our future work.

To close this paper, the remark should be added that there are many potential applications of the results of our research to areas like multimedia video indexing, searching and browsing, where it is common practice to use zoom lenses. It is

desirable to incorporate partial scene understanding capabilities under freely varying focal length, yet without having to go through elaborate egomotion estimation to obtain the scene information. The conclusion of this paper is that while it is very difficult to extract metric scene descriptions from video input, qualitative representations based on ordinal representation constitute a viable avenue.

References

1. M. Armstrong, A. Zissermann, and P. Beardsley, "Euclidean structure from uncalibrated images," In Edwin Hancock, pp. 508–518, BMVA PressHancock, York, UK, September 1994.
2. S. Bougnoux, "From Projective to Euclidean Space under any practical situation, a criticism of self-calibration," *ICCV '98*, pp. 790-796, 1998.
3. M.J. Brooks and W. Chojnacki, "Determining the egomotion of an uncalibrated camera from instantaneous optical flow," *J. Opt. Soc. Am. A* **14(10)**, pp. 2670-2677, 1997.
4. L.F. Cheong, C. Fermüller and Y. Aloimonos, "Effects of errors in the viewing geometry on shape estimation," *Computer Vision and Image Understanding* **71(3)**, pp. 356-372, Sep 1998.
5. L. Dron, "Dynamic camera self-calibration from controlled motion sequences", In *Proc. IEEE Conf. on CVPR*, pp. 501-506, New York,NY, 1993.
6. O. Faugeras, "Stratification of three-dimensional vision: projective, affine, and metric representations," *J. Opt. Soc. of Am. A* **12(3)**, pp. 465-484, Mar 1995.
7. C. Fermüller, "Passive Navigation as a Pattern Recognition Problem," *Int'l Journal of Computer Vision* **14**, pp. 147-158, 1995.
8. G. Florou and R. Mohr, "What accuracy for 3D measurement with cameras?", In *ICPR,* pp. 354-358, Los Alamitos, CA, 1996.
9. J. Gårding, J. Porrill, J.E.W. Mayhew and J.P. Frisby, "Stereopsis, vertical disparity and relief transformations," *Vision Research* **35(5)**, pp. 703-722, 1995.
10. R.I. Hartley, "Calibration of cameras using the essential matrix", In *Proc. ARPA Image Understanding Workshop,* pp. 911-915, Morgan Kaufmann Publishers, Inc., 1992.
11. R.I. Hartley, "Self-calibration from multiple views with a rotating camera", In *Proc. European Conf. on Computer Vision* **1**, pp. 471-479, Stockholm, Sweden, 1994.
12. A. Heyden and K. Åström, "Euclidean reconstruction from image sequences with varying and unknown focal length and principal points", In *Proc. IEEE Conf. on CVPR*, pp. 438-443, 1997.
13. B.K.P. Horn and E.J. Weldon Jr., "Direct method for recovering motion," *Int'l Journal of Computer Vision* **2**, pp. 51-76, 1988.
14. J.J. Koenderink and A.J. van Doorn, "Relief: Pictorial and Otherwise," *Image and Vision Computing* **13(5)**, pp. 321-334, 1995.
15. M. Pollefeys, L. Van Gool, and A. Oosterlinck, "Euclidean reconstruction from image sequences with variable focal length", In B. Buxton and R. Cipolla, editors, ECCV'96, volume 1064 of *Lecture notes in Computer Science,* pp. 31-44, Springer-Verlag, 1996.
16. G.P. Stein, "Accurate internal camera calibration using rotation with analysis of sources of error", In *Proc. ICCV'95*, pp. 230-236, Boston, MA, 1995.

17. P. Sturm, "Self-calibration of a moving zoom lens camera by pre-calibration", In *BMVC*, pp. 675-684, Edinburgh, 1996.
18. T. Svoboda and P. Sturm "A badly calibrated camera in ego-motion estimation—propagation of uncertainty", In *Proc. Int'l Conf. on Computer Analysis of Images and Patterns,* pp. 183-190, 1997.
19. B. Triggs, "Autocalibration and the absolute quadric", In *Proc. IEEE Conf. on CVPR*, pp. 609-614, 1997.
20. T. Viéville and O. Faugera, "The first order expansion of motion equations in the uncalibrated case," *CVGIP: Image Understanding* **64(1)**, pp. 128-146, July 1996.