Reconstruction from two views using approximate calibration

Richard Hartley and Chanop Silpa-Anan,
Department of Systems Engineering,
RSISE,
Australian National University,
ACT, 0200, AUSTRALIA.

Email: {hartley,chanop}@syseng.anu.edu.au

Abstract

We consider the problem of Euclidean reconstruction from two perspective images. This problem is well studied for calibrated cameras, and good algorithms are known. On the other hand, if the cameras are known to have square pixels (no skew and unit aspect ratio) then the problem is also theoretically solvable, provided an estimate of the principal point is provided. The focal lengths of the cameras may be computed from the fundamental matrix, and then a calibrated reconstruction algorithm applied. In reality, however, it has been shown that this process is quite sensitive to the computed fundamental matrix and the assumed position of the principal point. In fact, sometimes the estimate of the focal length fails, and so Euclidean reconstruction is impossible using this method. In this paper, we investigate the cause of this problem, and suggest an algorithm that more reliably leads to a reconstruction. It is unnecessary to know the exact location of the principal point, provided weak bounds on the principal point locations, and the focal lengths of the cameras are provided. The condition that points must lie in front of the cameras gives a further constraint. It is shown that by suffering only a very small degradation in residual point-reprojection error, it is possible to compute a fundamental matrix that always leads to a plausible focal length estimate, and hence Euclidean reconstruction.

1 Introduction

Scene reconstruction from a number of perspective images of a scene is one of the fundamental problems of computer vision, and reconstruction from two views is the simplest case of this problem. Earliest work in this area concentrated on calibrated cameras, from which it is in principle possible to obtain a Euclidean (sometimes called metric) reconstruction of the scene. Such a reconstruction is unique apart from a choice of the Euclidean coordinate frame, and an overall scale, which it is impossible to determine. Notable algorithms for obtaining a reconstruction from two calibrated views include [LH81, Hor90, Hor91].

With a shift of interest towards uncalibrated cameras, it was natural to ask whether Euclidean reconstruction was possible from uncalibrated cameras, leading to a well-known negative result ([HGC92, Fau92]) showing that for arbitrary uncalibrated views, the best one can achieve is projective reconstruction. On the other hand, it was shown in [Har92] that given reasonable assumptions about the calibration, it is still possible to achieve Euclidean reconstruction. Specifically, if the pixels are assumed to be square, and the principal point known, then the focal lengths of the two cameras may be computed from the fundamental matrix. All camera parameters now being known, the problem is reduced to a calibrated reconstruction problem, which may be solved by the previously mentioned techniques. The method of reconstruction, therefore is as follows:

1. From point correspondences between the two images compute the fundamental matrix.

2. Assume square pixels, and guess the position of the principal point (usually the centre of the image).

3. Compute the focal lengths of the two cameras (the only remaining unknown internal camera parameters).

4. Now, knowing the calibration matrix for each camera, compute the essential matrix $\mathtt{E} = \mathtt{K}_2{}^{\top}\mathtt{F}\mathtt{K}_1$, and carry out reconstruction from calibrated cameras.

In [Har92] step 3 of this algorithm was done by a complicated technique, but a simple formula for the focal lengths was subsequently found by Bougnoux ([Bou98]).

This algorithm suffers from various deficiencies.

1. Computation of the focal lengths is not possible if the principal axes of the two cameras intersect ([NHBP96]).

2. In most cases, the computation of the focal lengths is sensitive to the computed fundamental matrix, and the assumed positions of the principal points ([nwxx]);

Indeed, the sensitivity is so severe that computation of the focal lengths can not be relied upon at all, and this algorithm is of doubtful practical value.

The need to provide an estimate of the principal points in the two images is a further problem with this algorithm, particularly since it strongly affects the computed focal lengths. Furthermore, it is not commonly appreciated that a fundamental matrix, (even those computed by the best known methods) may in fact be incompatible with *any* reasonable choice of the principal point. The most common problem is that the focal lengths found by the above technique will be imaginary (hence impossible) values for most positions of the principal point. Such a fundamental matrix must be considered as wrong. This point will be expanded in this paper.

The goal of this paper is to show that Euclidean reconstruction is possible from two views, even without exact knowledge of the principal point, provided some assumptions are admitted concerning the position of the principal points, and the focal lengths of the images. For instance, in the examples discussed below, an assumption that the principal point is weakly constrained to be near the image centre (let us say plus or minus half the image radius), and that the focal lengths of the two cameras are approximately equal (maybe within 5%), is sufficient for Euclidean reconstruction to succeed. Other assumptions are possible, as will be seen.

As part of this reconstruction process, a fundamental matrix is found that is compatible with reasonable estimates of the principal point. Thus, incorporation of this *a priori* knowledge of the probable principal points and focal lengths leads to an improved estimate of the fundamental matrix.

2 Impossible fundamental matrices

In this section, it will be shown that some fundamental matrices, even those computed from good quality data by the best algorithms (such as a Maximum Likelihood algorithm) are nevertheless incompatible with the data, and hence wrong. This is actually quite a common phenomenon, as the examples will show, particularly when the principal rays of the two cameras come close to intersecting. This situation occurs when the cameras are pointing approximately towards a common point in space, which must be so if they are to share matched points.

For a fundamental matrix to be accepted as being compatible with a set of matched points $\mathbf{x}_{2i} \leftrightarrow \mathbf{x}_{1i}$, three conditions are necessary.[1]

[1]Recall that we are making an assumption of square pixels (that is zero skew and unit aspect ratio) for both cameras.

1. The point correspondences must satisfy the coplanarity condition $\mathbf{x}_{2i}{}^{\top}\mathtt{F}\mathbf{x}_{1i}$, with a small residual error.

2. For at least some assumed locations of the principal points $\mathbf{p}_1$ and $\mathbf{p}_2$, the values of f_1^2 and f_2^2 computed from (1) are positive.

3. Given such assumed principal point positions, and corresponding focal length values, a calibrated reconstruction is possible, for which the reconstructed 3D points (apart from a small percentage of possible outliers) lie in front of the reconstructed cameras.

This last condition is related to the concept of "cheirality" discussed in [Har98], where it is shown that satisfying the coplanarity condition for *some* fundamental matrix is not sufficient for the set of matches to be realizable.

A set of matched points will often contain a small proportion of outliers, or false matches. Normally, these outliers must be eliminated from the matched-point set before the fundamental matrix is computed, otherwise the results will be bad. Nevertheless, we allow the possibility that some outliers may remain, which may cause 3D points to be displaced and end up behind the cameras.

It is important to understand the role of cheirality in calibrated reconstruction. It is well known (see for instance [May93, Har92, HZ00]) that based on the essential matrix alone there are four possible choices of camera pairs. However, a single matched point is sufficient to disambiguate the situation, since the reconstructed point will lie in front of both cameras for only one of the four pairs. If the estimated essential matrix is correct, then this will lead to a choice of the correct pair of camera matrices, and hence all reconstructed points (derived from correct point matches) will lie in front of the two cameras. However, if the essential matrix is estimated inaccurately, then it is possible for this condition to fail. Not all the reconstructed points will lie in front of both the reconstructed cameras. This may happen, for instance, if the assumed or computed calibration matrices for the two cameras are wrong.

This discussion is illustrated by the examples given in Fig 1. There a fundamental matrix is used, computed using a bundle-adjustment method (the Gold-standard method of [HZ00]), which is as good as any known method. The set of matched points used for this computation are of high quality, outliers having been previously removed, and the residual error from estimation of $\mathtt{F}$ being very small (see Fig 13). Nevertheless, it is seen that the possible positions for the principal point are very constrained.

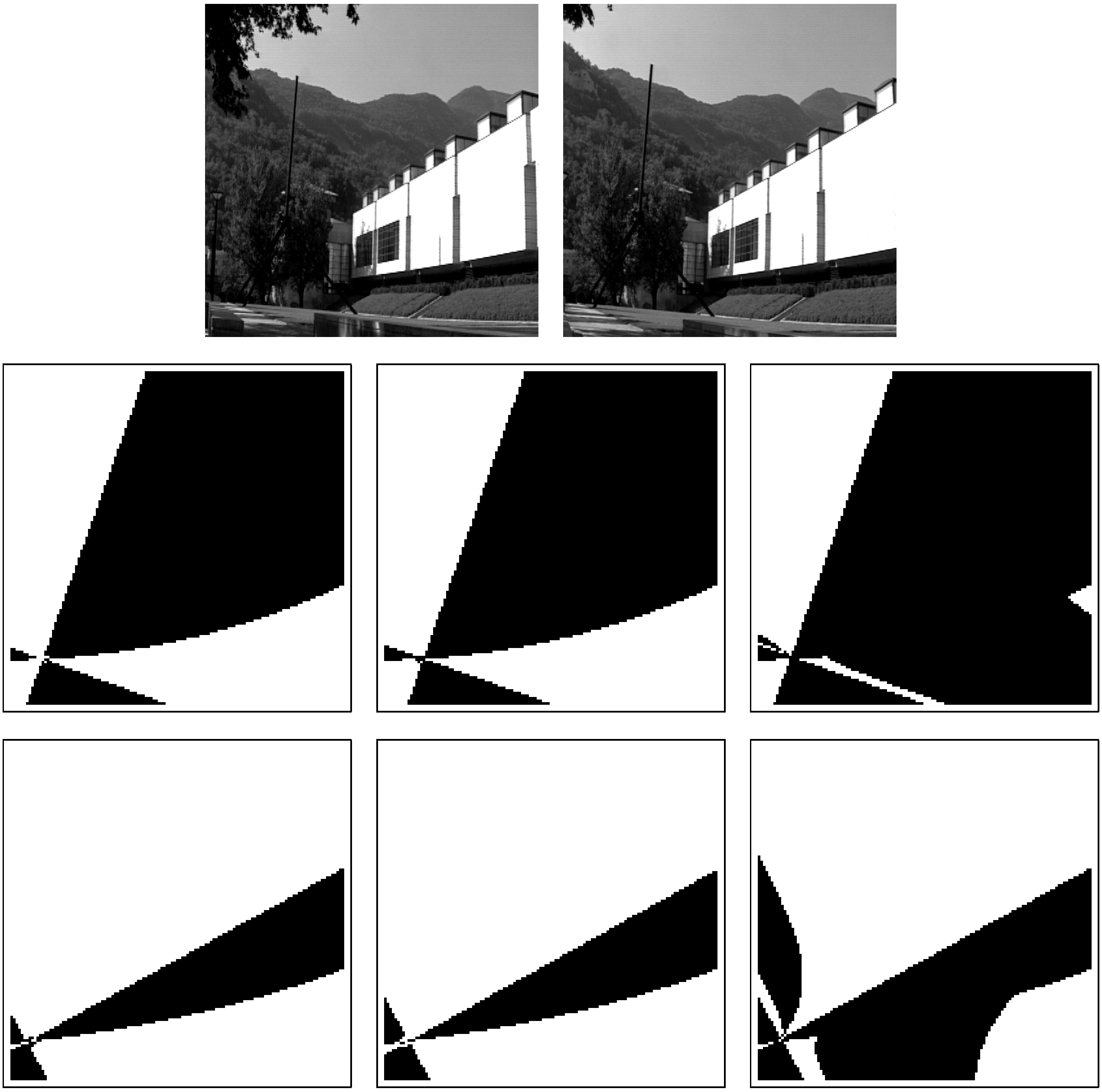

Figure 1: *At the top are a pair of images used for computing a fundamental matrix. The fundamental matrix was computed using the Gold-standard algorithm of [HZ00]. For this example it is assumed that the principal point is the same in both images. In the lower set of figures, the possible positions for the principal point are shown. White shows possible positions of the principal point, and black impossible. The criteria used for the three diagrams are:* f_1^2 *positive (left),* f_2^2 *positive (centre) and* $< 10\%$ *points behind cameras. The graph on the right shows those positions for the principal points where all three conditions are satisfied. As seen, there are very few possible positions for the principal points consistent with this fundamental matrix – certainly not the centre of the image.*

The third row of figures shows the same results for the fundamental matrix computed using the method described in this paper. The obtained fundamental matrix is consistent with the assumption of principal point near the centre of the image.

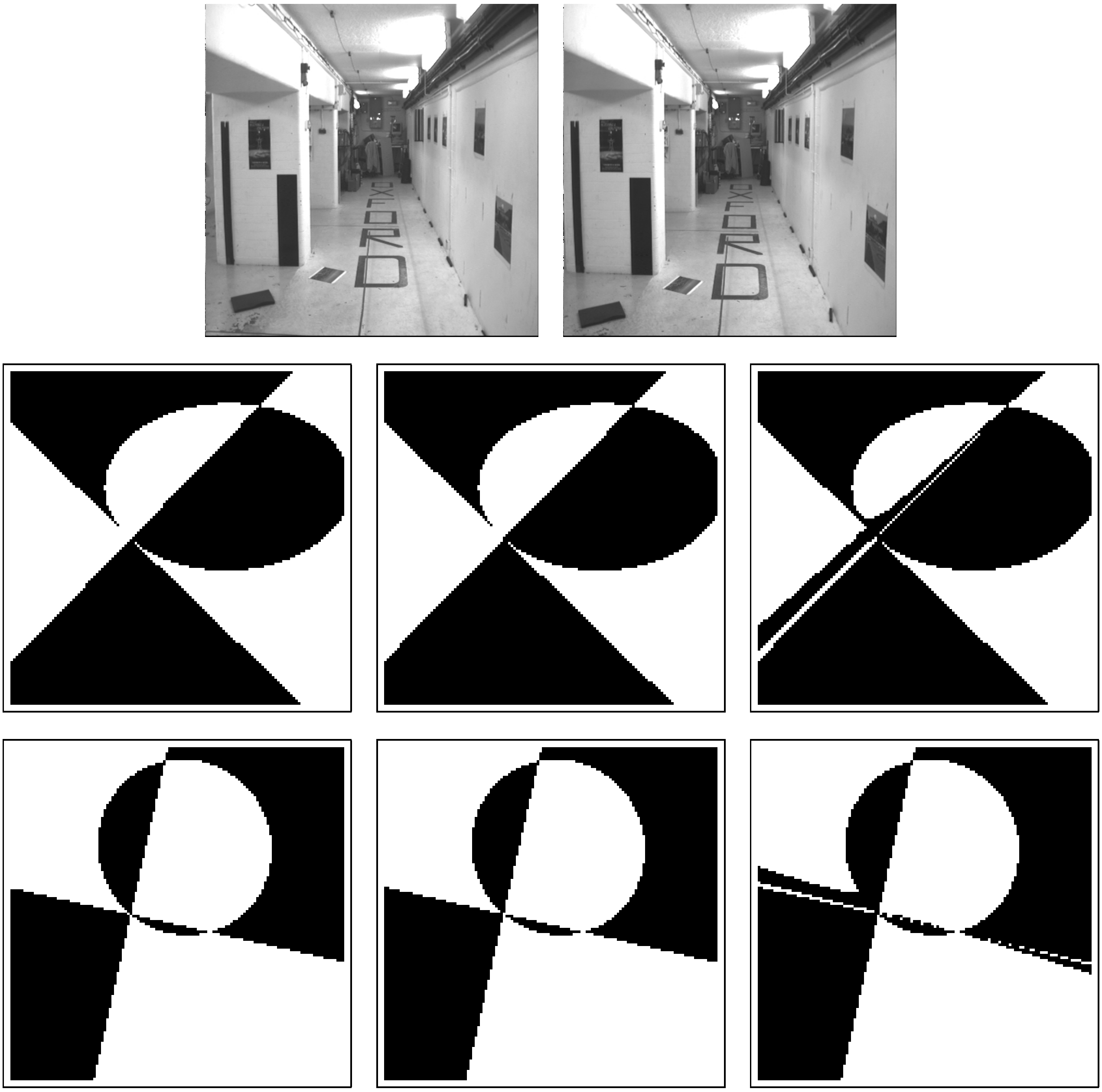

Figure 2: *This figure shows the same diagrams for a different set of images. At the top the images, in the middle possible principal points for the fundamental matrix found using the Gold-standard algorithm, and at the bottom the diagrams found using the algorithm of this paper.*

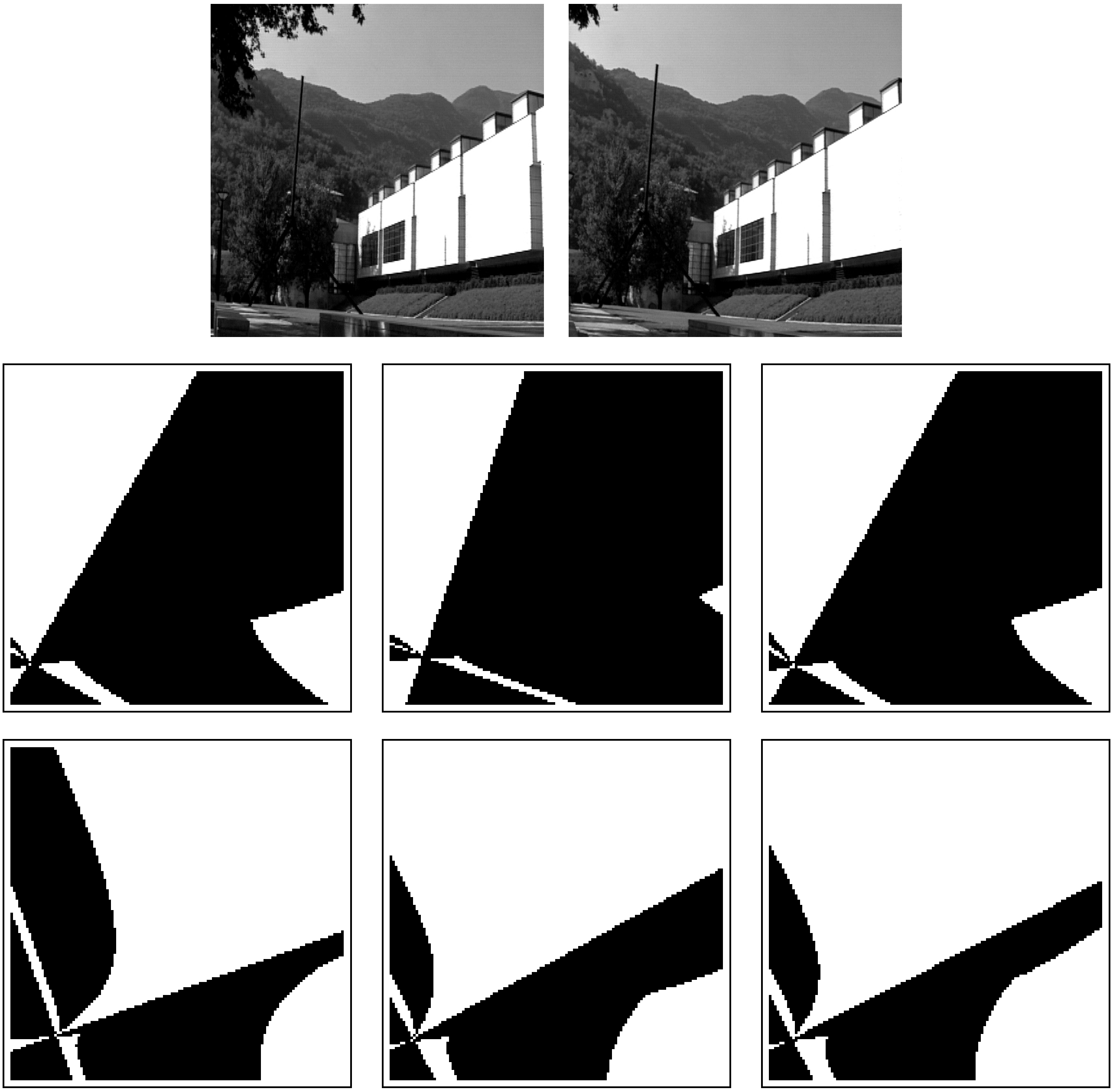

Figure 3: *This figure shows the possible principal point positions compatible with fundamental matrices computed using several different methods. The methods used are (from left to right, top to bottom) : normalized 8-point algorithm, gold-standard algorithm, algebraic distance algorithm, Sampson-error algorithm, algorithm of this paper, and calibrated reconstruction algorithm (see text for description).*

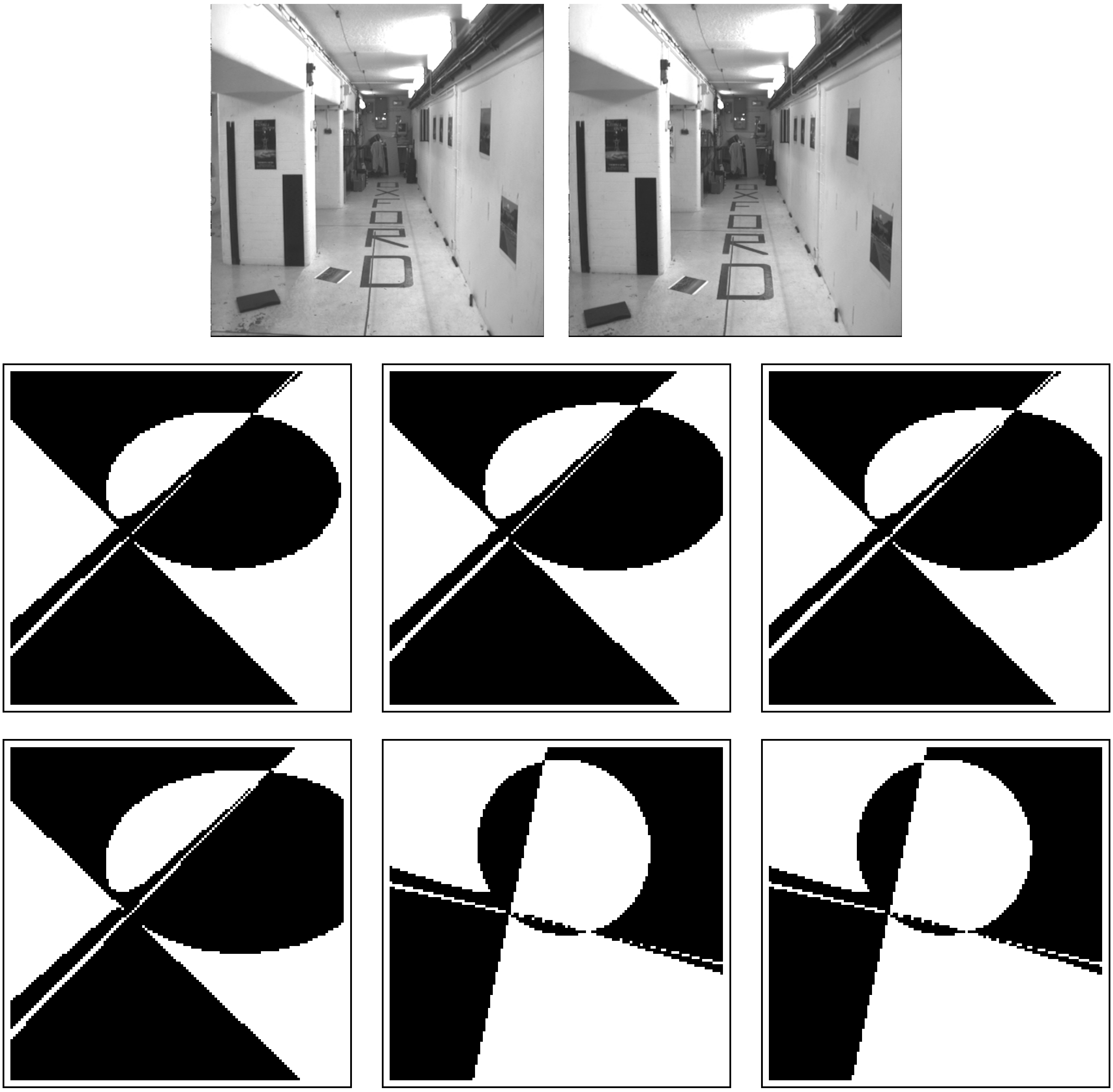

Figure 4: *This figure shows the possible principal point positions compatible with fundamental matrices computed using several different methods (same as Fig 3 with a different image. The methods used are (from left to right, top to bottom) : normalized 8-point algorithm, gold-standard algorithm, algebraic distance algorithm, Sampson-error algorithm, algorithm of this paper, and calibrated reconstruction algorithm (see text for description).*

3 Using a priori knowledge to compute the fundamental matrix

It is often said that the computation of the fundamental matrix is unstable, and to an extent this is true. This is a characteristic of the problem itself, and not the algorithms used, since many algorithms that obtain close to the Maximum Likelihood estimate are known.

The fundamental matrix is however basically a projective object, suitable for applying in situations where the camera calibration is unknown. However, it has been used in many situations in which partial calibration of the camera is known, at least approximately. There are many examples of this.

1. In reconstruction from two views, the camera calibration is usually not completely unknown. For instance, the pixels are usually square, or at least rectangular with known aspect ratio. The principal point is usually near the centre of the image, and the focal length is at least known within some reasonable bounds. In addition the projective reconstruction algorithm based on the fundamental matrix has been used in cases where the calibration of the camera is known. Oliensis [Oli00] has undertaken a study of this, and argues that sometimes calibrated camera algorithms give better results.

2. The fundamental matrix has been used in self-calibration from two views ([Bou98, Har92]). To do this, it is necessary to assume square pixels, and to make an estimate of the principal point in the two images. Unfortunately, the estimate of the focal lengths is quite strongly dependent on the assumed locations of the principal points.

In this paper, we concentrate particularly on this second scenario, in which the focal lengths of the cameras may be computed from the fundamental matrix. However, we are interested in the inverse problem of using *a priori* estimates of the focal lengths of the cameras and also the principal points to improve the estimate of the fundamental matrix, and hence the 3D reconstruction. This method, therefore lies between two extremes – computation of the fundamental matrix for totally uncalibrated cameras, and for completely calibrated cameras (in which case it is the essential matrix that is computed). In this case, we compute the fundamental matrix for the case where the camera calibration is approximately know.

A simple formula for the focal lengths, given the fundamental matrix and principal point has been given by Bougnoux ([Bou98]):

$$f_2^2 = -\frac{(\mathbf{p}_1{}^\top [\mathbf{e}_1]_\times \mathtt{I}\mathtt{F}^\top \mathbf{p}_2)(\mathbf{p}_1{}^\top \mathtt{F}^\top \mathbf{p}_2)}{\mathbf{p}_1{}^\top ([\mathbf{e}_1]_\times \mathtt{I}\mathtt{F}^\top \mathtt{I}\mathtt{F})\mathbf{p}_1} \qquad (1)$$

where $\mathtt{I}$ is the diagonal matrix $\text{diag}(1,1,0)$. A similar formula for f_1^2 is obtained by interchanging the roles of the two images.

However, this method has an important failure configuration, when the principal rays of the two cameras meet in space ([NHBP96]). In this case, the principal points of the two cameras satisfy the coplanarity constraint $\mathbf{p}_2{}^\top \mathtt{F}\mathbf{p}_1 = 0$. For many image pairs this condition is nearly satisfied. Thus, being close to a critical configuration, one might expect that the results will be quite unstable. A stability analysis for estimation of f in this instance was undertaken in [nwxx], in which it was verified that focal length estimate is indeed often unreliable. It will be shown here that for many image pairs, the estimates of the focal length are effectively useless. Worse, the algorithm often fails entirely. This is because the formula of Bougnoux gives an expression for f_1^2 or f_2^2, the squares of the focal lengths of the two images. In the presence of noise, or an inaccurate estimate of the principal point, the value of f^2 so obtained is negative, and so f is an imaginary (hence impossible) value. Note that this is intrinsic to the problem as formulated, and is not just an artefact of Bougnoux's formula. In this failure case, the given fundamental matrix is just not compatible with the assumed values of the two principal points. Either the fundamental matrix or the principal point assumption is wrong.

The idea of this paper is to take advantage of this fact to obtain a better estimate of the fundamental matrix, and at the same time a better estimate of the focal lengths of the cameras. The way this is done is to add additional terms in a cost function used to estimate $\mathtt{F}$. These weight terms discriminate against improbable, or impossible values of the principal points of focal lengths of the cameras. Thus, a very small, or (worse) negative value of f^2 incurs a high cost, which makes it very unlikely to be accepted. If the weights are related to a probability distribution for principal point or focal length, then the new estimate of $\mathtt{F}$ may be thought of as a maximum a priori (MAP) estimate of the fundamental matrix.

4 A cost function

Given a set of point correspondences $\mathbf{x}_{2i} \leftrightarrow \mathbf{x}_{1i}$, our object is to estimate the fundamental matrix subject to prior assumptions about the distribution of the focal lengths and principal points of the two cameras. The method is applicable easily to any assumed distributions, but most commonly a normal (Gaussian) distribution will be assumed. An iterative (Levenberg-Marquardt) method is used to minimize a cost function of the following form:

$$\text{Cost}(\mathtt{F}, f_1^2, f_2^2, \mathbf{p}_1, \mathbf{p}_2) = C_{\mathtt{F}}(\mathtt{F}) + C_f(f_1^2, f_2^2) + C_p(\mathbf{p}_1, \mathbf{p}_2) \ . \qquad (2)$$

Thus the total cost is the sum of three cost functions, measuring the cost of estimates of the fundamental matrix $\mathtt{F}$, the squared focal lengths f_1^2 and f_2^2, and the principal points respectively. The reason for expressing the second cost function in terms of the *squared* focal lengths is because of the form of Bougnoux's formula (1), as will be seen later. We want to define a cost (very high) for a negative value of f^2.

The cost function $C_{\mathtt{F}}(\mathtt{F})$ is also a function of the point correspondences. Although various cost functions are possible, we prefer to use the Sampson cost function

$$\sum_i \frac{(\mathbf{x}_{2i}{}^\top \mathtt{F}\mathbf{x}_{1i})^2}{(\mathtt{F}\mathbf{x}_{1i})_1^2 + (\mathtt{F}\mathbf{x}_{1i})_2^2 + (\mathtt{F}^\top \mathbf{x}_{2i})_1^2 + (\mathtt{F}^\top \mathbf{x}_{2i})_2^2} . \qquad (3)$$

This cost function is a first-order approximation to the correct geometric (or Maximum Likelihood) cost function.

This cost function is minimized over a set of parameters by a parameter-minimization algorithm (LM). The set of parameters is divided into two parts.

1. A set of parameters parametrizing the fundamental matrix $\mathtt{F}$, and

2. A set of 4 (or 2) parameters defining the positions of the principal points. Two parameters may be used if it is assumed (and enforced) that the principal point is the same in both images.

The minimization algorithm is as follows.

1. Given $\mathtt{F}$ and $\mathbf{p}_i$, compute f_1^2 and f_2^2 using Bougnoux's formula (1).

2. Compute the cost function $\mathrm{Cost}(\mathtt{F}, f_i^2, \mathbf{p}_i^2)$.

3. Vary the parameters of $\mathtt{F}$ and the $\mathbf{p}_i$ to minimize the cost function.

Form of $C_\mathbf{p}$ and C_f. The specific form of the cost functions for $\mathbf{p}_i$ and f_i^2 may be chosen in various ways, and it is not the purpose of this paper to undertake a detailed investigation of all possible cost functions. However, a natural form for $C_\mathbf{p}$ is the squared Euclidean distance of the estimated principal point from the nominal value. Thus

$$C_\mathbf{p}(\mathbf{p}_i) = w_p^2 d(\mathbf{p}_i, \bar{\mathbf{p}}_i)^2 \qquad (4)$$

where $\bar{\mathbf{p}}_i$) represents the nominal position of the principal point, and w_p is a weight. There is a cost term of this form for each of the two principal points.

The form of the cost function C_f may be a little more complicated, since we want to ensure that the value of f^2 does not end up negative. In addition, it may be appropriate to enforce a condition that the two focal lengths are the same (or approximately so). Accordingly, the cost function $C_f(f_1^2, f_2^2)$ used in our implementation of the algorithm has several components:

$$\begin{aligned} C_f(f_1^2 \quad , \quad f_2^2) &= w_1^2(f_1^2 - \bar{f}_1^2)^2 + w_2^2(f_2^2 - \bar{f}_2^2)^2 \\ &+ \ w_d^2(f_1^2 - f_2^2) \\ &+ \ w_{z1}^2(f_{\mathrm{min}}^2 - f_1^2)^2 + w_{z2}^2(f_{\mathrm{min}}^2 - f_2^2)^2 \end{aligned} \qquad (5)$$

Recall here that f_i^2 is the value returned by Bougnoux's formula (1), and may be negative. The final two terms of this equation involve a "minimum" value f_{min} for the focal length, and are only included if $f_i^2 < f_{\mathrm{min}}^2$. (That is w_{z1} and w_{z2} are zero unless f_i^2 is small, or negative.). This term grows rapidly for small or negative values of f_i^2, and effectively prevent f_i^2 taking on negative values. A reasonable minimum value of f_{min} can be deduced from the size of the image. The field of view of the camera is equal to $2\arctan(\mathrm{dim}/f)$, where dim is the radius of the image. For small values of f, this becomes unrealistically large. Most images encountered (except for extreme wide angled views) do not have field of view exceeding 75°.

The choice of the weight values may be chosen according to taste. The values of w_{zi} are not critical, and normally it is sufficient to apply quite weak weights for the other values.

5 Initialization

The input data for the reconstruction problem includes an estimate of the focal length and principal point of the cameras. Therefore, it makes sense to use these estimates to carry out a calibrated reconstruction to obtain an initial estimate of the fundamental matrix.

Thus, given a set of point correspondences, $\mathbf{x}_{2i} \leftrightarrow \mathbf{x}_{1i}$ and initial estimates $\bar{\mathbf{p}}_i$ and $\bar{f}_i$ of the principal points and focal lengths, an initial value for the fundamental matrix is found as follows. Let $\mathtt{K}_1$ and $\mathtt{K}_2$ be the initial calibration matrices for the two images. Thus,

$$\mathtt{K}_1 = \begin{bmatrix} \bar{f}_1 & 0 & \bar{x}_1 \\ & \bar{f}_1 & \bar{y}_1 \\ & & 1 \end{bmatrix} \qquad \mathtt{K}_2 = \begin{bmatrix} \bar{f}_2 & 0 & \bar{x}_2 \\ & \bar{f}_2 & \bar{y}_2 \\ & & 1 \end{bmatrix} \qquad (6)$$

where $\bar{\mathbf{p}}_i = (\bar{x}_i, \bar{y}_i)$ are the principal points. Let $\mathtt{F}$ be an estimate of the fundamental matrix computed from the point correspondences using any desired method. We used a method ([HZ00]) that minimizes algebraic error. The essential matrix $\mathtt{E}$ may then be computed as

$$\mathtt{E} = \mathtt{K}_2{}^\top \mathtt{F} \mathtt{K}_1 \ .$$

This value of the essential matrix will not generally be quite correct for the assumed calibration matrices, since it will not satisfy the necessary condition for a essential matrix, namely that it have two equal singular values. To correct this, the singular value decomposition of $\mathtt{E}$ is computed as $\mathtt{E} = \mathtt{UDV}^\top$, and a corrected essential matrix is computed, by setting $\widehat{\mathtt{E}} = \mathtt{UIV}^\top$, where $\mathtt{I} = \text{diag}(1, 1, 0)$. Finally a corrected value of the fundamental matrix is computed by setting

$$\widehat{\mathtt{F}} = \mathtt{K}_2{}^{-\top}\widehat{\mathtt{E}}\mathtt{K}_1^{-1} \ .$$

The resulting fundamental matrix $\widehat{\mathtt{F}}$ is the best approximation to $\mathtt{F}$, compatible with the assumed calibration matrices.

Iteration will start with this estimate $\widehat{\mathtt{F}}$ for the fundamental matrix, and the assumed values $\bar{\mathbf{p}}_1$ and $\bar{\mathbf{p}}_2$ for the two principal points. The cost function $C_{\mathtt{F}}(\widehat{\mathtt{F}})$ will be slightly greater than $C_{\mathtt{F}}(\mathtt{F})$, but usually the difference is not too great. However, because of the way $\widehat{\mathtt{F}}$ is defined, the cost functions $C_{\mathbf{p}}(\bar{\mathbf{p}}_1, \bar{\mathbf{p}}_2)$ and $C_f(f_1^2, f_2^2)$ will both have initial values zero.

The alternative is to start iteration with fundamental matrix $\mathtt{F}$, and the two principal point estimates. In this case, $C_{\mathtt{F}}(\mathtt{F})$ will have a smaller value than $C_{\mathtt{F}}(\widehat{\mathtt{F}})$, but the value of $C_f(f_1^2, f_2^2)$ will often be quite high, particularly if the values of f_1^2 or f_2^2 are negative. Thus, the initial cost can be very high, and we are starting far from the minimum. This can lead to problems with correct convergence.

Notice the difference between the two initial methods. Recall that the values of f_1^2 and f_2^2 are derived from the formula (1) from the current (in this case the initial) estimates of $\mathtt{F}$ and $\mathbf{p}_i$. Since $\widehat{\mathtt{F}}$ is a correct fundamental matrix for the calibration matrices given in (6), it follows that (1) applied to $\widehat{\mathtt{F}}$ and $\bar{\mathbf{p}}_i$ will give values of $f_i^2 = \bar{f}_i^2$, and so the cost will be zero. However, (1) applied to $\mathtt{F}$ and the initial values of $\bar{\mathbf{p}}_i$ can give values of f_i^2 that differ very greatly from the nominal values $\bar{f}_i^2$.

6 Experiments

The algorithm was carried out on both synthetic and real data. The algorithms used for computing the fundamental matrix were as follows. For details of these algorithms, the reader is referred to [HZ00].

1. A normalized 8-point algorithm.
2. The gold-standard algorithm (bundle adjustment).
3. Iterative minimization of algebraic error (algebraic minimization algorithm).
4. Iterative minimization of Sampson error, (3).
5. The algorithm of this paper (denoted by **a-priori** in the graphs and tables).
6. Calibrated reconstruction, given fixed assumed values of the principal points and focal length.

The final algorithm is the one described in section 5, which is used as an initial point for the **a-priori** algorithm.

6.1 Synthetic image experiments

Evaluation of the method was carried out on two sets of data. In the first set, a set of scattered points on the surface of a cube were used, and data were generated by projecting the points into a pair of nominal cameras. In the second set, a pair of images were used to build a realistic-looking model of a house (using the images in Fig 8). The model was then projected back into two images, approximating the original images. This data was the synthetic data, corresponding to a realistic imaging setup, that was used for experiments.

For each synthetic data set, noise was added in varying degrees, and Euclidean reconstruction was carried out. The results are shown in figures 5 – 7, which give the reconstruction errors in each case.

The experiments were carried out with two different a-priori estimates for the principal point and focal length. In the first set, the exact values were given. This of course gives a significant advantage to the calibrated reconstruction algorithm, since it is provided with exact values for these parameters. In the second set of experiments, the value of the principal point is shifted by 30 pixels in each dimension (in a 512×512 image), and a slightly modified value for the focal length is given. This is perhaps more realistic, since these parameter values may not be known exactly in advance.

Choice of weights. The **a-priori** algorithm was run with very weak weights, in order to place minimal constraints on principal point and focal length. The value of w_p was equal to 0.01, which means that a variation of 100 pixels (in a 512×512 image) was given as much weight as one pixel reprojection error for a single matched point. The principal point was, however, assumed to be the same in both images.

As for the focal length, the weights w_1 and w_2 were set to zero, which means that no constraint was placed on these parameters individually. However, a value of 0.001 was chosen for w_d. This means that a value of 1000 for $f_1^2 - f_2^2$ was equivalent to one pixel reprojection error. Since f_i was around 500, this means approximately 2 pixels difference $f_1 - f_2$.

As stated, the values of w_{zi} are not very critical. A value of $f_{\text{min}} = 100$ was taken, and a weight $w_{zi} = 0.01$ was chosen.

Findings. Perhaps because of the controlled nature of the synthetic algorithms, in no case did the values of f_i^2 turn out to be negative. As a result, the reconstruction results were good for all the algorithms used. The **a-priori** algorithm performed significantly better than the calibrated reconstruction algorithm, except in the case where exact parameters were given, and for high noise values.

Noise Level	Cube gold-standard	Cube a-priori	Shed gold-standard	Shed a-priori
0.0	43.82	4.40	20.01	42.77
0.1	43.99	5.00	25.52	45.22
0.2	44.18	6.31	51.76	43.82
0.3	44.38	7.99	71.68	47.66
0.5	44.81	15.27	102.9	82.06
0.7	45.33	35.79	127.8	74.83
1.0	46.46	47.56	148.6	90.86
1.5	51.42	55.35	171.2	117.5
2.0	60.73	66.00	177.9	125.4
3.0	85.50	97.36	191.6	161.0
4.0	117.3	123.0	203.1	186.8

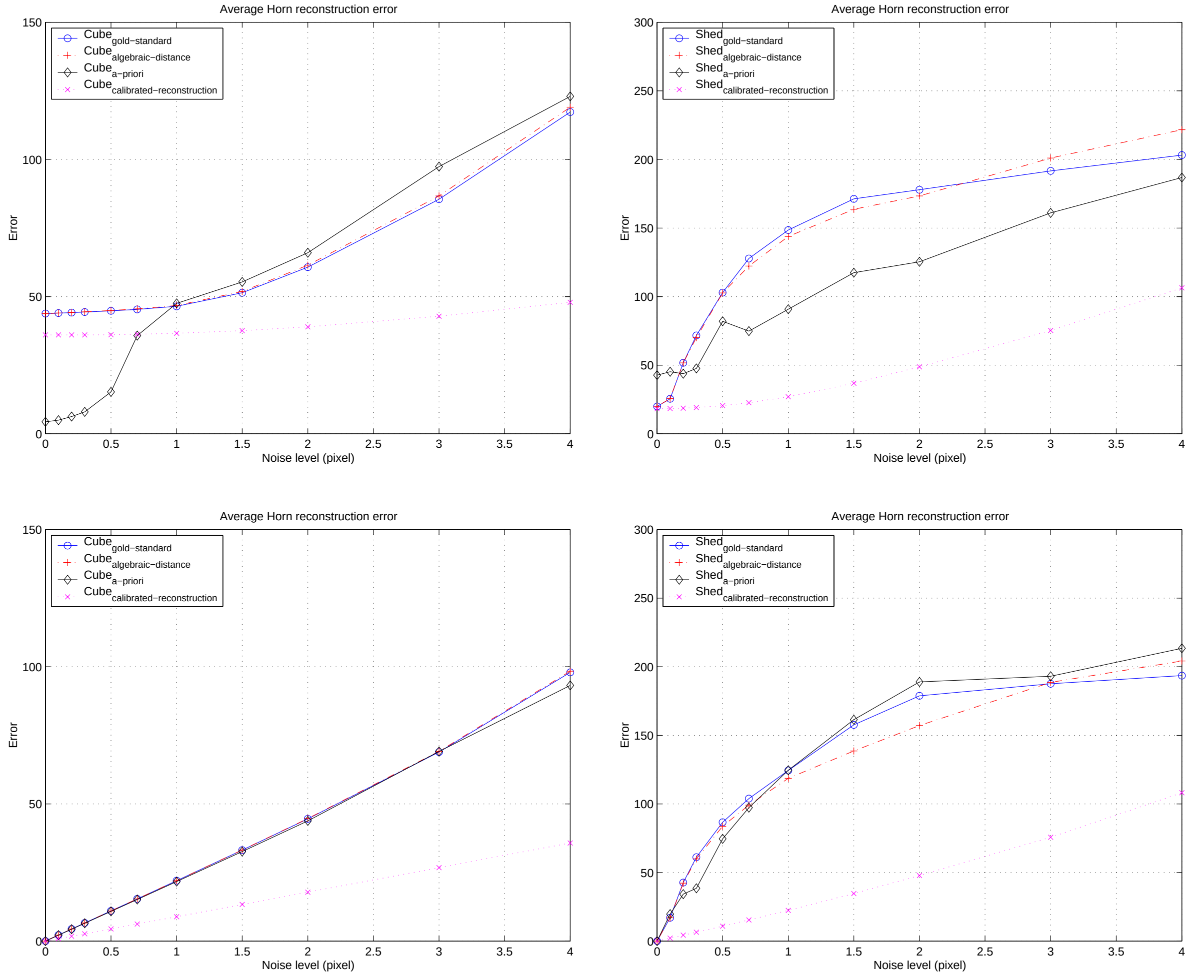

Figure 5: **Mean errors for synthetic data.** *This graph shows the average reconstruction error for the two synthetic data set, cube and shed images. For this example, four algorithms are used to estimated the fundamental matrix, gold-standard algorithm, algebraic distance algorithm, algorithm in this paper, and calibrated For the middle two graphs, an incorrect initial principle point (225.5,225.5) and focal length (590) are given to the algorithm. For the lower pair of graphs, the true initial principle point (255.5,255.5) and focal length (500) are given.*

Noise Level	Cube		Shed	
	gold-standard	a-priori	gold-standard	a-priori
0.0	43.82	4.40	20.01	42.77
0.1	43.12	4.47	22.38	42.46
0.2	42.40	5.42	37.28	43.82
0.3	41.70	6.40	57.84	43.01
0.5	40.27	10.19	93.17	39.74
0.7	38.88	14.98	166.8	37.07
1.0	37.00	23.44	168.2	43.78
1.5	34.49	33.98	170.234	59.21
2.0	35.88	43.62	169.6	60.8
3.0	68.46	76.68	170.7	131.2
4.0	111.0	86.18	172.7	187.8

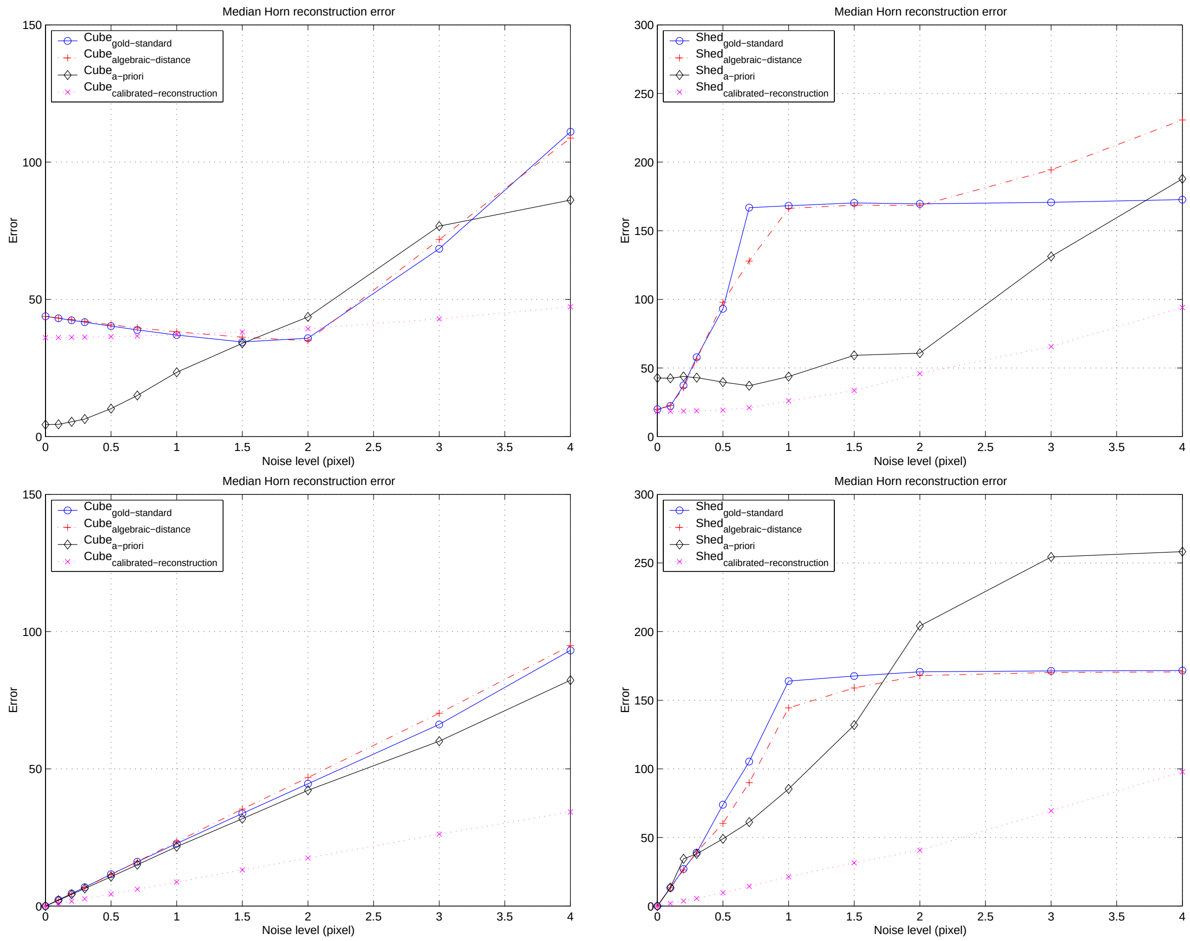

Figure 6: **Median errors for synthetic data.** *This gives the same results as in Fig 5 except that median error is given, instead of average.*

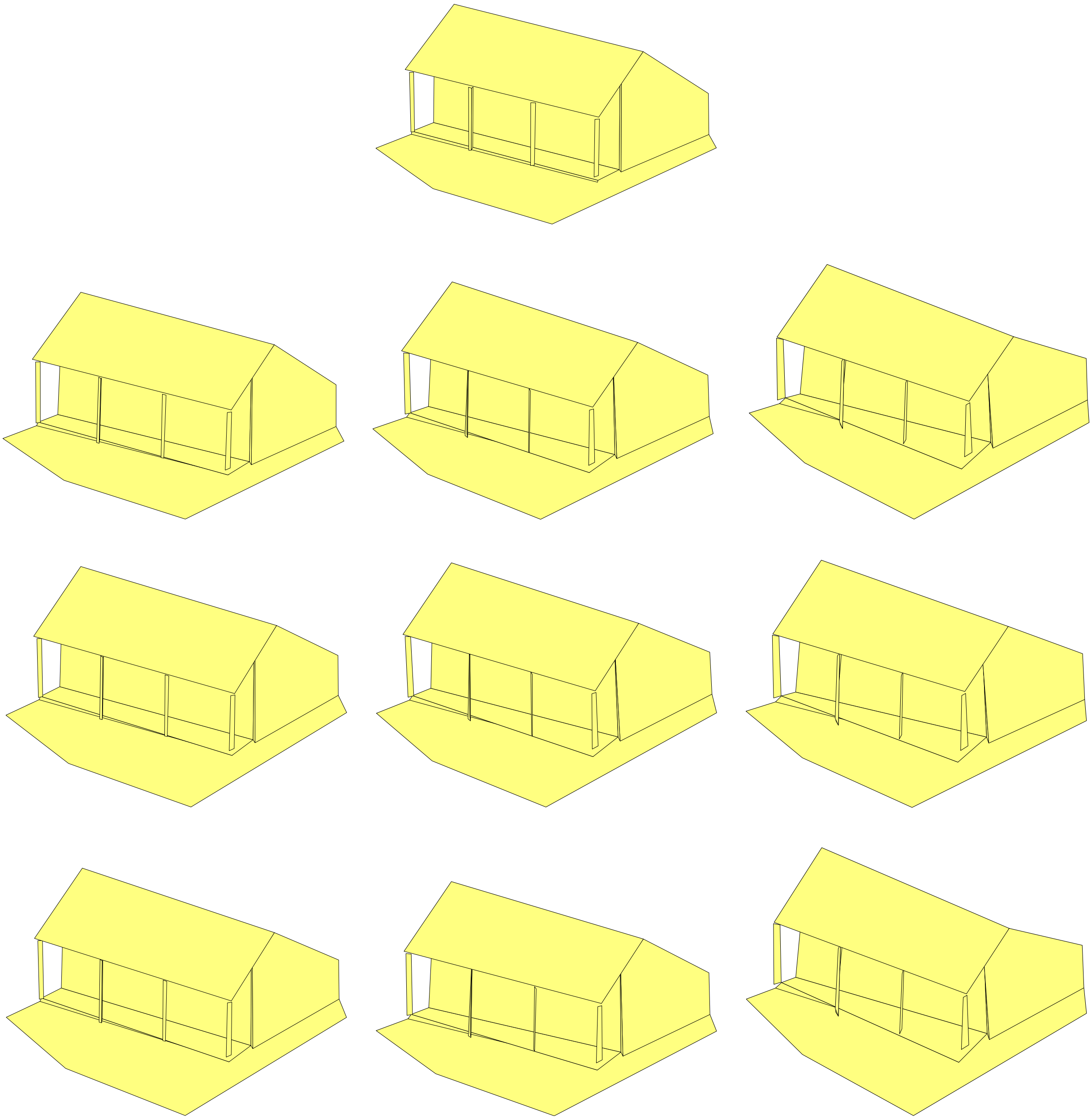

Figure 7: **Reconstructed models.** *This figure shows the reconstruction result of the synthetic shed images derived from the images in Fig 8. On the top row, a reconstruction from noise-free data is shown. On the second row a reconstruction using the gold-standard algorithm for noise levels of 0.4, 1.0 and 2.0 pixels (left to right). The third row shows the recontruction with algebraic distance algorithm. The last row shows the reconstruction with the algorithm in this paper.*

Figure 8: *Pair of images used to create the synthetic shed data.*

6.2 Real image experiments

Experiments were also carried out on real images, for which, however, no ground truth structure was known. The image pairs used were the ones used in [HZ00] for evaluation of fundamental matrix computation, and kindly supplied to us (see Fig 9). They may be taken as reasonably representative of real image sets. The experiments were carried out with the same weight values and a-priori parameter estimates as the synthetic images. The results are given in the following figures 10 – 13.

Findings. Most of the algorithms used gave values for the fundamental matrices for which the focal length values f_i^2 were negative, under the reasonable assumption that the principal points were in the centre of the images. Thus, in this case it was not possible to proceed with Euclidean reconstruction. For each of the images, calibrated reconstruction algorithm was also used to compute the fundamental matrix. Since the calibration was only guessed at, understandably the residual reprojection error in this case was much worse than for the present **a-priori** method. The consequence of this will be a significantly degraded reconstruction, since the 3D points will not be estimated accurately. In addition, in one case (the statue example, see Fig 12), the calibrated reconstruction resulted in points being found behind the camera. This also occured for some other algorithms and image sets, but not for the **a-priori** algorithm. The consequence of points ending up behind the camera will be a severely distorted Euclidean reconstruction.

It is important to note that the residual projection error is only very slightly greater for the **a-priori** algorithm than it is for the gold-standard algorithm (or the other algorithms). Despite this, the **a-priori** algorithm gives fundamental matrices that are much more realistic, in fact usable for subsequent 3D reconstruction.

7 Conclusion

This paper points out some of the difficulties involved in Euclidean reconstruction from two views. It is shown that most of the standard algorithms for computing the fundamental matrix will result in matrices that are wrong, and in fact unusable for 3D reconstruction. However, adding some weak constraint terms to the cost function used to compute the fundamental matrix can lead to vastly improved results, particularly as far as estimation of the focal lengths, and subsequent 3D reconstruction is concerned. The effect is far more noticeable for real images than synthetic ones. The cost of adding these constraint terms is very small in terms of increased reprojection residual.

In our experiments, we did not attempt to add any extra cost terms to ensure that the reprojected points lie in front of the cameras. In all the cases shown here, the new algorithm resulted in points being reconstructed in front of the cameras. This is not guaranteed, however, and we occasionally found cases in which this condition was violated.

Further work in this area can include determination of the best weight functions for principal point and focal-length errors, and an investigation of adding terms to the cost function to ensure points lie in front of the cameras.

References

[Bou98] S. Bougnoux. From Projective to Euclidean space under any practical situation, a criticism of self-calibration. In *Proc. 6th International Conference on Computer Vision, Bombay, India*, pages 790–796, January 1998.

[Fau92] O. D. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In

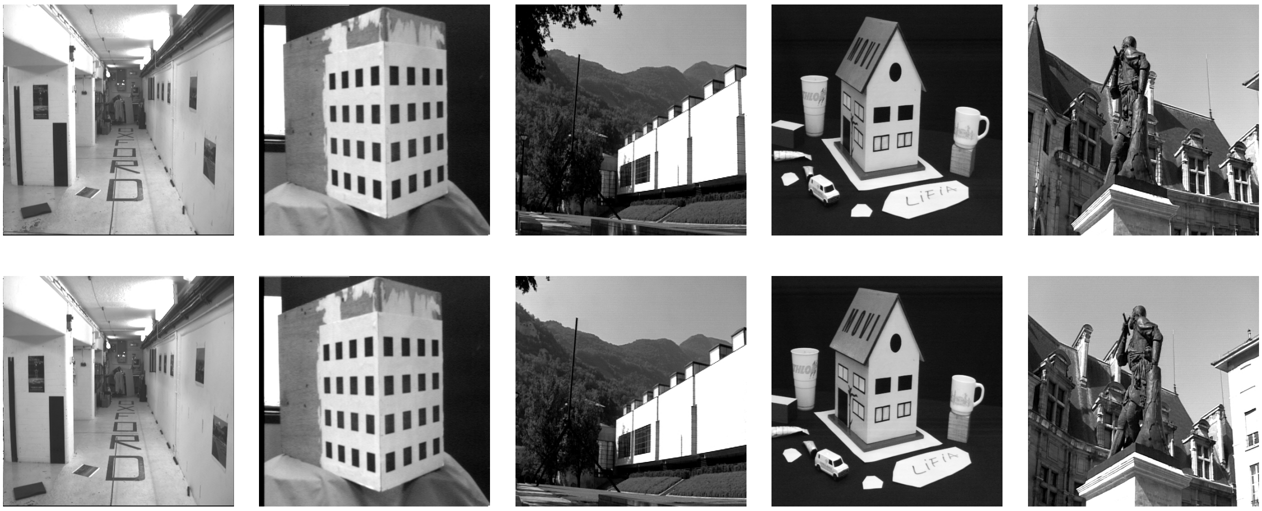

Figure 9: *Images used for the real image experiments.*

Method / Image	normalized 8-points, gold-standard, algebraic distance, Samson error, calibrated	a-priori
Basement	(255.5,255.5)/(255.5,255.5)	(255.5,255.5)/(255.5,255.5)
Calib	(255.5,255.5)/(255.5,255.5)	(255.9,254.6)/(255.9,254.6)
House	(255.5,255.5)/(255.5,255.5)	(254.2,234.9)/(255.4,234.9)
Museum	(255.5,255.5)/(255.5,255.5)	(232.0,229.8)/(232.0,229.8)
Statue	(255.5,255.5)/(255.5,255.5)	(248.5,206.9)/(248.5,206.9)

Figure 10: **Estimated principal points.** *In the experiment, the principle points are assumed to be (255.5,255.5) for both images for normalized 8-points algorithm, gold-standard algorithm, algebraic distance algorithm, and Sampson error algorithm. For algorithm in this paper an initial estimated for principle points are the same (255.5,255.5). The algorithm in this paper has moved the principle points away from the initial estimates in some image pairs.*

Method / Image	normalized 8-points	gold standard	algebraic distance	Sampson error	a-priori	calibrated
Basement	X/X	X/X	X/X	X/X	503.3/503.4	707.1/707.1
Calibration	872.3/864.3	510.7/509.2	X/X	X/X	279.9/279.7	707.1/707.1
House	X/X	X/X	X/X	X/X	451.6/451.5	707.1/707.1
Museum	X/X	X/X	X/X	2964/3001	654.8/654.8	707.1/707.1
Statue	X/X	X/X	X/X	X/X	1167/1167	707.1/707.1

Figure 11: **Estimated focal lengths.** *This table shows the estimated focal length for each image pair calculated from 5 algorithms. Note that some well known algorithms such as normalized 8-points algorithm and gold standard algorithm may give an impossible answer (imaginary focal length, denoted by X).*

Method Image	normalized 8-points	gold standard	algebraic distance	Sampson error	a-priori	calibrated
Basement	0/0	0/0	0/100	0/0	100/100	100/100
Calibration	100/100	100/100	100/0	100/100	100/100	100/100
House	91.7/91.7	100/0	100/0	100/0	100/100	100/100
Museum	0/0	0/100	0/100	94.9/94.9	100/100	100/100
Statue	100/100	100/0	100/0	100/100	100/100	52.4/52.4

Figure 12: **Percentage points in front of cameras.** *This table shows the percentage of points that lie in front of each camera in each image pair. It is clear that the algorithm in this paper gives a better result in that the estimated fundamental matrix yields estimated points in front of the camera. When wrong principle points and focal length are used in the calibrated reconstruction algorithm some points may end up behind the cameras, as seen for the statue pair.*

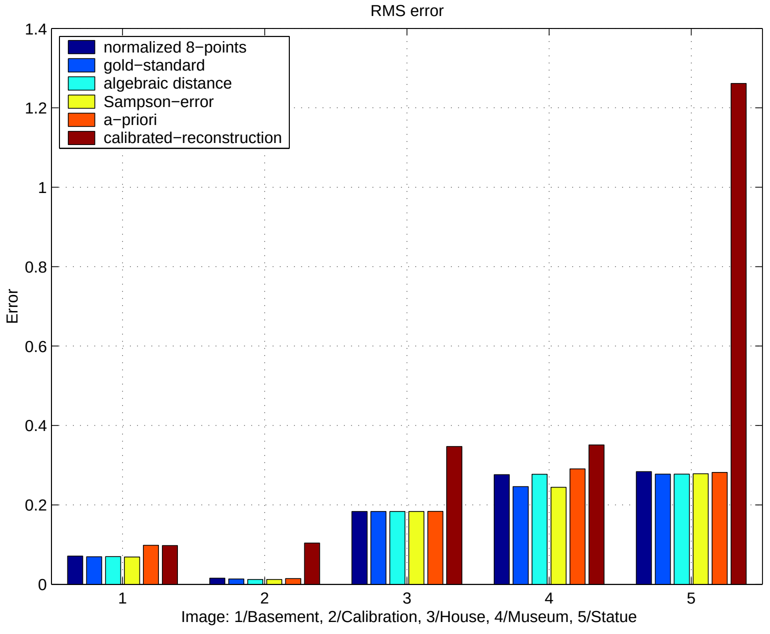

Figure 13: **Residuals.** *This figure summarizes the residual rms error of the estimated fundamental matrix obtained from 5 algorithms: normalized 8-points algorithm, gold-standard algorithm, algebraic distance algorithm, Sampson error algorithm, algorithm in this paper, and algorithm in this paper with constraints on the focal length and principle point. It is clear that the algorithm in this paper gives a similar level of residual RMS error when compared with the other uncalibrated algorithms. However, using a calibrated reconstruction algorithm gives significantly higher residual, because the internal parameters may not have been guessed correctly.*

Proc. European Conference on Computer Vision, LNCS 588, pages 563–578. Springer-Verlag, 1992.

[Har92] R. I. Hartley. Estimation of relative camera positions for uncalibrated cameras. In *Proc. European Conference on Computer Vision*, LNCS 588, pages 579–587. Springer-Verlag, 1992.

[Har98] R. I. Hartley. Chirality. *International Journal of Computer Vision*, 26(1):41–61, 1998.

[HGC92] R. I. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 1992.

[Hor90] B. K. P. Horn. Relative orientation. *International Journal of Computer Vision*, 4:59–78, 1990.

[Hor91] B. K. P. Horn. Relative orientation revisited. *Journal of the Optical Society of America*, 8(10):1630–1638, 1991.

[HZ00] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[LH81] H. C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, September 1981.

[May93] S. J. Maybank. *Theory of reconstruction from image motion*. Springer-Verlag, Berlin, 1993.

[NHBP96] G. Newsam, D. Q. Huynh, M. Brooks, and H. P. Pan. Recovering unknown focal lengths in self-calibration: An essentially linear algorithm and degenerate configurations. In *Int. Arch. Photogrammetry & Remote Sensing*, volume XXXI-B3, pages 575–80, Vienna, 1996.

[nwxx] Author name withheld. Paper title withheld. In *xx*, xx.

[Oli00] J. Oliensis. A critique of structure from motion algorithms. *Computer Vision and Image Understanding*, 80:172–214, 2000.