Affine Reconstruction from Monocular Vision in the Presence of A Symmetry Plane

D. Q. Huynh*
School of Information Technology
Murdoch University, Perth WA 6150, Australia
Email: *d.huynh@murdoch.edu.au* or *du@it.murdoch.edu.au*

Abstract

This paper reports a closed-form solution for reconstructing a scene up to an affine transformation from a single image in the presence of a symmetry plane. Unlike scene reconstruction in stereo vision, the affine reconstruction process discussed in this paper does not require any knowledge about camera parameters or camera orientation relative to the scene, so camera self-calibration is totally eliminated. By setting in the scene a plane mirror which creates lateral symmetric world points for an uncalibrated, perspective camera to capture, the linear equations involved in the reconstruction process can be derived from two sets of similar triangles. The affine reconstruction is relative to an arbitrary affine coordinated frame implicitly defined on the mirror plane. Also involved in the process are the estimation of the epipole and recovery of the image-to-mirror plane homography. Implementation on estimating the epipole is detailed. A real experiment is presented to demonstrate the reconstruction.

1. Introduction

Scene reconstruction from images captured by uncalibrated cameras has caught the attention of many Vision and Photogrammetry researchers. Research on camera self-calibration originated with Faugeras et al's paper [3]. Since then, numerous research works on self-calibration and reconstruction have emerged. While Euclidean reconstruction contains all the information that we strive to achieve, affine structure has a fruitful amount of information that has been known to be useful in many Vision applications. In affine reconstruction, parallelism, ratio of regions on parallel planes, ratio of lengths of parallel lines are all preserved.

*The author thanks the anonymous reviewers for their suggested improvement to the paper and Prof. Kanatani for helpful comments. This research was in part supported by the Ministry of Education, Science, Sports and Culture, Japan under the Grant-in-Aid for JSPS Fellows (No. P98162).

Reconstructing a scene from monocular vision has not been substantially studied because the depth information that is lost from the 3D scene to the 2D image transformation is known to be unrecoverable from one image alone. However, when symmetric scene information is known to be present, both projective and affine reconstructions are retrievable. In Rothwell et al's work [13], projective reconstruction of a set of 3D points is investigated and the rank of the matrix involved is used to verify the class of polyhedral the object belongs (the object itself need not be a polyhedron). Following that, Fawcett et al [4] report an affine reconstruction algorithm for an affine image of objects that have bilateral symmetries. In their work, reconstruction modulo an Euclidean transformation to a three parameter family of symmetric shapes is also discussed. In this paper we describe a simple affine reconstruction method for objects of arbitrary shapes from a single perspective image. We create lateral symmetric world points by explicitly setting a plane mirror in the scene. With an arbitrary affine coordinate frame defined on the mirror, affine structure can be recovered. Our approach reported here is different from [4, 13] in the following respects: a perspective camera model is used here rather than an affine one as in [4]; the structure recovered in our system is affine rather than projective as in [13].

We organise the paper as follows. Section 2 describes the geometry and then affine reconstruction algorithm. Section 3, which is another major section of the paper, describes the estimation of the epipole. Due to the limitation of space, Section 4 presents one of the experiments conducted for this research. Section 5 discusses the issues on occlusion and applications of Vision systems that include mirrors. Lastly, Section 6 gives the conclusion.

2. Affine reconstruction

A brief description on notations: we use upper case bold letters for world points, lower case bold letters for image

points and lines. Entities with a $'$ are virtual world or image points. Matrices are in upper case; scalars are in lower case.

2.1. The geometry

Let a mirror plane Π_m be placed in the scene. If $\mathbf{A} \leftrightarrow \mathbf{A}'$ is a pair of corresponding real and virtual world points in the scene then the symmetry of points about the mirror plane ensures that $\mathbf{A}$ and $\mathbf{A}'$ are on either side of Π_m and are equidistant from it (Fig. 1), i.e. the midpoint $\bar{\mathbf{A}}$ of the segment is on the mirror plane. Furthermore, the segment $\mathbf{A}\mathbf{A}'$ is orthogonal to Π_m. The image points $\mathbf{a}$ and $\mathbf{a}'$ are the perspective projections of $\mathbf{A}$ and $\mathbf{A}'$ onto the image plane Π_i under the optical centre $\mathbf{O}$ of the camera. Let $\mathbf{A}_m$ and $\mathbf{A}'_m$ be, respectively, the intersection points of the rays $\mathbf{OA}$ and $\mathbf{OA}'$ with Π_m. A study of Fig. 1 reveals that $\mathbf{A}$, $\mathbf{A}'$, $\mathbf{A}_m$ $\mathbf{A}'_m$, $\bar{\mathbf{A}}$, $\mathbf{O}$, $\mathbf{a}$, and $\mathbf{a}'$ are coplanar in space.

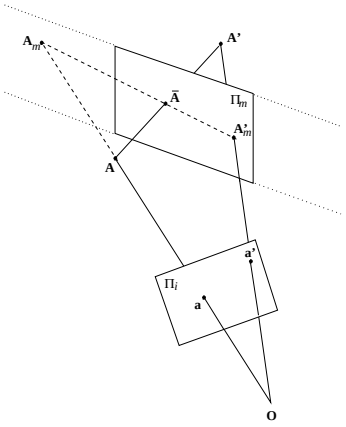


Figure 1. Geometry of a world point A with respect to the mirror plane Π_m and image plane Π_i.

Another set of coplanar world points are those on the mirror plane Π_m: $\mathbf{A}_m$, $\mathbf{A}'_m$, $\mathbf{O}_m$, $\bar{\mathbf{A}}$. Relative to an arbitrary 2D coordinate frame $\mathcal{F}_m$ defined on Π_m, they can be expressed as 2-vectors and be referred to as $\mathbf{a}_m$, $\mathbf{a}'_m$, $\mathbf{o}_m$, $\bar{\mathbf{a}}_m$ respectively.

2.2. An affine reconstruction algorithm

Construct the vector $\mathbf{OO}_m$ orthogonal to the mirror plane Π_m. Let d_m and d be, respectively, the distances of the perspective centre $\mathbf{O}$ and the world point $\mathbf{A}$ from Π_m (Fig. 2). Let $\mathbf{o}_m = (o_x, o_y)^\top$ be the coordinates of $\mathbf{O}_m$ relative to the arbitrary 2D coordinate frame $\mathcal{F}_m$. The world point $\mathbf{A} = (A_x, A_y, A_z)^\top$ has three unknown components: the first two components can be expressed relative to an *affine* coordinate frame $\mathcal{F}_m$ and the last component A_z can be expressed as the distance d from Π_m. From the two sets of similar right triangles, $\mathbf{OAQ}$ and $\mathbf{OA}_m\mathbf{O}_m$; $\mathbf{OA}'\mathbf{Q}'$ and $\mathbf{OA}'_m\mathbf{O}_m$, the following equations can be derived:

$$d_m\,(P\mathbf{A} - \mathbf{o}_m) \;=\; (\,d_m - d\,)\,(\mathbf{a}_m - \mathbf{o}_m) \qquad (1)$$

$$d_m\,(P\mathbf{A}' - \mathbf{o}_m) \;=\; (\,d_m + d\,)\,(\mathbf{a}'_m - \mathbf{o}_m) \qquad (2)$$

where $P = [\,I \,|\, \mathbf{0}\,]$ with I being the 2×2 identity matrix, so $P\mathbf{A} = (A_x, A_y)^\top$.

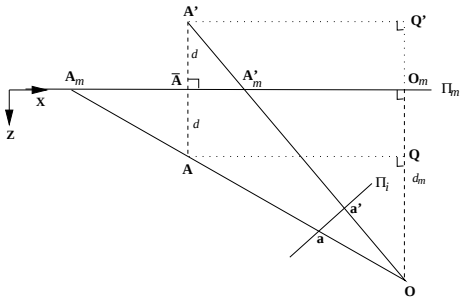


Figure 2. Top view of the geometry of a world point A and the mirror and image planes.

Since relative to the $\mathcal{F}_m$ coordinate frame $P\mathbf{A}$ and $P\mathbf{A}'$ have the same coordinates, (1) and (2) can be combined to eliminate $\mathbf{A}$ and $\mathbf{A}'$ and yield

$$d_m\,(\mathbf{a}_m - \mathbf{a}'_m) - d\,(\mathbf{a}_m + \mathbf{a}'_m) + 2\,d\,\mathbf{o}_m = 0. \qquad (3)$$

The unknown quantities in (3) are d, $\mathbf{o}_m$, and d_m provided that $\mathbf{a}_m$ and $\mathbf{a}'_m$ can be computed from other means. The total number of unknowns is therefore $n+3$, where n is the number of pairs of corresponding real and virtual world points, each of which has an unknown distance d from the mirror plane Π_m that needs to be solved; the remaining 3 unknowns are: the o_x and o_y components of $\mathbf{o}_m$, and the distance d_m of the perspective centre $\mathbf{O}$ to Π_m.

A close study of (3) reveals that multiplying d_m and d in the equation by an arbitrary scalar incurs no change to the equality. One can therefore conclude that d_m, and consequently the corresponding d value of each world point, can only be recovered up to scale. Dividing the equation through by d gives

$$\tilde{d}\,(\mathbf{a}_m - \mathbf{a}'_m) - (\mathbf{a}_m + \mathbf{a}'_m) + 2\,\mathbf{o}_m = 0 \qquad (4)$$

where $\tilde{d} = d_m/d$.

The total number of unknowns that needs to be solved in (4) for n corresponding pairs of image points is $n+2$. With two equations provided by each pair of corresponding real and virtual world points, the total number of equations available is $2n$.

Geometrically, the quantity $\hat{d} \equiv 1/\tilde{d} = d/d_m$ has a more meaningful interpretation: It is the distance d of each world point being scaled by the factor $1/d_m$. The estimated $\hat{d}$'s are now taken as the z-components of the reconstructed world points.

Having recovered the coordinates of $\mathbf{o}_m$ and all the world points' z-components up to scale, the next step is to compute their x- and y-components. Let $\mathbf{A} = (A_x, A_y, A_z)^\top$ be the world point to be reconstructed. Then dividing (1) and (2) by d_m and rearranging yields

$$\begin{bmatrix} A_x \\ A_y \end{bmatrix} = \left(1 - \hat{d}\right)(\mathbf{a}_m - \mathbf{o}_m) + \mathbf{o}_m = \left(1 + \hat{d}\right)(\mathbf{a}'_m - \mathbf{o}_m) + \mathbf{o}_m. \quad (5)$$

Note that unlike the z-component A_z of $\mathbf{A}$, the components A_x and A_y are recovered up to their exact scale with respect to the coordinate axes chosen for $\mathcal{F}_m$.

The reconstruction described above is affine only if the 2D coordinate frame $\mathcal{F}_m$ is affine. That is, if 4 arbitrary points on Π_m that form a rectangle or a parallelogram can be identified, then, by setting these 4 points to $(0,0,1)^\top$, $(1,0,1)^\top$, $(0,1,1)^\top$, $(1,1,1)^\top$, an affine frame $\mathcal{F}_m$ is defined. There are two approaches that one can adopt: *(1)* Put explicit landmarks on the mirror plane to define this coordinate frame; *(2)* Identify in the image 4 coplanar world points (if any) whose orthogonal projections onto the mirror plane form a rectangle or parallelogram. Although the second approach requires the existence of 4 coplanar world points on the object that form an affine coordinate frame, the object itself does not need to be a polyhedron. This approach is adopted in the paper and will be discussed in Section 3.2.

The reconstruction procedure assumes that the coordinates of $\mathbf{a}_m$ and $\mathbf{a}'_m$ on Π_m are available from other means. While these points are not known a priori, their image projections are identical to $\mathbf{a}$ and $\mathbf{a}'$ of the real and virtual world points $\mathbf{A}$ and $\mathbf{A}'$. If the homography [14] between the image plane Π_i and the mirror plane Π_m is known then $\mathbf{a}_m$ and $\mathbf{a}'_m$ can be synthesised. The determination of this image-to-mirror homography and the affine frame $\mathcal{F}_m$ will be outlined in Section 3.2.

3. The real/virtual correspondence problem

Since it is more convenient to express image points as homogeneous 3-vectors (by adding 1 to the third component), from here onward, image points and image lines are 3-vectors unless explicitly stated otherwise.

3.1. The epipole

For each pair of real and virtual image points, $\mathbf{a}_i \leftrightarrow \mathbf{a}'_i$, the epipolar line joining $\mathbf{a}_i$ and $\mathbf{a}'_i$ is simply $\mathbf{a}_i \times \mathbf{a}'_i$ i.e. the cross product of the homogeneous coordinates of the two image points. Ideally, the epipole $\mathbf{e}_\infty$ satisfies $(\mathbf{a}_i \times \mathbf{a}'_i)^\top \mathbf{e}_\infty = 0$, for all i. In the presence of noise, the dot product above may not identically vanish and one must tackle the problem of optimally determining $\mathbf{e}_\infty$. In this paper, we assume that the correspondences between the real and virtual image points have been identified, i.e. the *real/virtual correspondence problem* has been solved. We will focus on the recovery of the epipole, the determination of the image-to-mirror homography, and the final affine reconstruction.

3.2. Determining the image-to-mirror plane homography and $\mathcal{F}_m$

Given that the image points $\mathbf{a}_i \leftrightarrow \mathbf{a}'_i$ of 4 corresponding pairs of real and virtual world points $\mathbf{A}_i \leftrightarrow \mathbf{A}'_i$, for $1 \le i \le 4$, have been correctly identified and the epipole $\mathbf{e}_\infty$ has been computed, the image projection $\bar{\mathbf{a}}_i$ of the midpoint $\bar{\mathbf{A}}_i$ of the world line segment $\mathbf{A}_i\mathbf{A}'_i$ can be determined from the cross ratio. Since the cross ratio $\{\mathbf{a}_i, \mathbf{a}'_i; \bar{\mathbf{a}}_i, \mathbf{e}_\infty\}$ is known to have the value -1 (see [14]), one can deduce the image coordinates $\bar{\mathbf{a}}_i$ that corresponds to $\bar{\mathbf{A}}_i$ on the mirror plane. If the world points $\{\mathbf{A}_i \,|\, 1 \le i \le 4\}$ form a rectangle or parallelogram then the midpoints $\{\bar{\mathbf{A}}_i \,|\, 1 \le i \le 4\}$ on Π_m define an affine frame $\mathcal{F}_m$. The 3×3 non-singular matrix H that defines the image-to-mirror homography can then be determined from

$$H\bar{\mathbf{a}}_i \sim \bar{\mathbf{A}}_i, \qquad 1 \le i \le 4 \quad (6)$$

where $\sim$ denotes equality up to scale.

If explicit landmarks on the mirror plane are used and their image projections are identified then the homography H can be directly computed without knowing the epipole.

3.3. The epipolar constraint

Once the epipole has been estimated, given a real (or virtual) image point $\mathbf{p}$, the corresponding virtual (or real) image point $\mathbf{p}'$ must be on the line defined by $\mathbf{p} \times \mathbf{e}_\infty$. This is the epipolar line that constrains the search space for the corresponding virtual (or real) image point $\mathbf{p}'$.

Proposition 1. *Let the pencil of planes with axis at $\mathbf{OO}_m$ be denoted by the symbol pencil($\mathbf{OO}_m$). Then for any corresponding pair of real and virtual world points $\mathbf{A} \leftrightarrow \mathbf{A}'$ there exists a plane Π in pencil($\mathbf{OO}_m$) such that $\mathbf{A}$, $\mathbf{A}' \in \Pi$.*

Proof:
The vector $\mathbf{OO}_m$ is constructed to be orthogonal to Π_m. As the line segment $\mathbf{AA}'$ is also orthogonal to Π_m, it follows immediately that the points $\mathbf{A}$, $\mathbf{A}'$, $\mathbf{O}$, $\mathbf{O}_m$ are coplanar in space. Since $\mathbf{A}$ and $\mathbf{A}'$ are arbitrarily chosen, every plane determined by $\mathbf{O}$ and any corresponding pair of real and virtual world points must contain the point $\mathbf{O}_m$. All these planes thus form a pencil with axis at $\mathbf{OO}_m$. The plane in the pencil that contains $\mathbf{A}$ and $\mathbf{A}'$ is known as the *epipolar plane* of $\mathbf{A}$ (or $\mathbf{A}'$). Fig. 3 shows the geometry involved. □

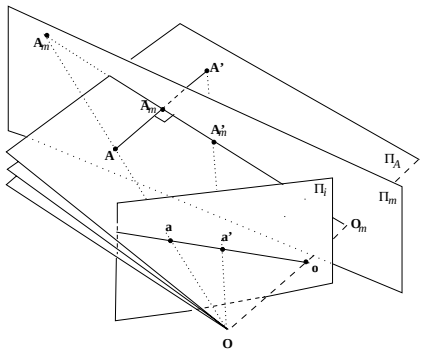


Figure 3. A pencil of planes with axis $\mathbf{OO}_m$ intersecting the image plane Π_i and mirror plane Π_m. Π_A is the epipolar plane of A (or A′).

Proposition 2. *The axis $\mathbf{OO}_m$ meets Π_i in the epipole $\mathbf{e}_\infty$.*

Proof:
Let $\mathbf{o}$ be the intersection point of line segment $\mathbf{OO}_m$ with the image plane Π_i (Fig. 3). From Proposition 1, all the epipolar planes intersect at the axis $\mathbf{OO}_m$. These epipolar planes intersect Π_i at the epipolar lines, which are concurrent at the epipole $\mathbf{e}_\infty$. It follows that $\mathbf{o}$ and $\mathbf{e}_\infty$ must coincide. Thus, the axis $\mathbf{OO}_m$ meets Π_i at the epipole $\mathbf{e}_\infty$. □

3.4. Computing the epipole

Since getting a good estimate of the epipole is essentially for our affine reconstruction, we will explain at length in this subsection the various approaches that we have implemented.

Normalisation

In computing the epipole line $\mathbf{l}_i \equiv \mathbf{a}_i \times \mathbf{a}'_i$, the large magnitude (in pixel unit) of the first two components of $\mathbf{a}_i$ and $\mathbf{a}'_i$ relative to their third component (a constant 1) often causes undesirable numerical instability. A sensible choice is to normalise the first two components of each $\mathbf{a}_i$ and $\mathbf{a}'_i$ before computing $\mathbf{l}_i$. Such normalisation has been employed by Hartley [6] in estimating the fundamental matrix. Here we define a *column scaling matrix* [5] $D \equiv \text{diag}(1/m_x, 1/m_y, 1)$, where $(m_x, m_y, 1)^\top = \frac{1}{2n}\sum_{i=1}^{n}(\mathbf{a}_i + \mathbf{a}'_i)$, to normalise $\mathbf{a}_i$ and $\mathbf{a}'_i$.

Three methods for estimating the epipole are presented below. All these methods are preceded by the above normalisation process (i.e. multiplied by D) so the computation is performed in *scaled* pixel coordinates. When the epipole has been computed, the unnormalisation process (multiplied by D^{-1}) is then applied. Our simulations show that this normalisation process gave the methods significant improvement in epipole estimation. Golub and Van Loan [5] also suggest *row-column equilibration* (page 125) by combining a row scaling matrix. In this paper, we focus only on column scaling.

Method 1: SVD

1. Construct
$$L = \begin{bmatrix} \mathbf{l}_1{}^\top \\ \vdots \\ \mathbf{l}_n{}^\top \end{bmatrix}$$
where $\mathbf{l}_i \equiv D\mathbf{a}_i \times D\mathbf{a}'_i$.
2. Apply the singular value decomposition (SVD) to L. Define $\mathbf{e}_\infty$ to be the singular vector that corresponds to the smallest singular value of L.

Both methods 2 and 3 below take into account the noise involved in the corresponding image points. Before we go on to describe these two methods, we first introduce notations and general ideas.

Let $\mathbf{l}_i = \bar{\mathbf{l}}_i + \Delta\mathbf{l}_i$, for $1 \le i \le n$, be the observed homogeneous coordinates of the epipolar lines, with $\bar{\mathbf{l}}_i$ being the true, unknown line coordinates and $\Delta\mathbf{l}_i$ the noise component involved. If $\Lambda_{\mathbf{a}_i}$ and $\Lambda_{\mathbf{a}'_i}$ are the covariance matrices of the corresponding image points $\mathbf{a}_i$ and $\mathbf{a}'_i$, then the first order approximation of the covariance matrix (see [1] for a general discussion) of $\Delta\mathbf{l}_i$ is simply

$$\begin{aligned}\Lambda_{\mathbf{l}_i} \;=\; & P_i \left([\,D\mathbf{a}_i\,]_\times D\Lambda_{\mathbf{a}'_i} D^\top [\,D\mathbf{a}_i\,]_\times^\top + \right. \\ & \left. [\,D\mathbf{a}'_i\,]_\times D\Lambda_{\mathbf{a}_i} D^\top [\,D\mathbf{a}'_i\,]_\times^\top \right) P_i^\top \end{aligned} \tag{7}$$

where $[\,.\,]_\times$ denotes the skew symmetric matrix of the vector concerned and $P_i = I - \mathbf{l}_i\mathbf{l}_i^\top$, with I being the 3×3 identity matrix. Assuming that the noise $\Delta\mathbf{l}_i$ is Gaussian of 0 mean, then maximising the likelihood function of the $\Delta\mathbf{l}_i$'s is equivalent to minimising the quantity $\sum_{i=1}^{n} \Delta\mathbf{l}_i^\top \Lambda_{\mathbf{l}_i}^- \Delta\mathbf{l}_i$ subject to the linearised constraints $\mathbf{l}_i^\top \mathbf{e}_\infty - \Delta\mathbf{l}_i^\top \mathbf{e}_\infty = 0$, for $1 \le i \le n$, where $\Lambda_{\mathbf{l}_i}^-$ denotes the rank-2 generalised inverse of $\Lambda_{\mathbf{l}_i}$. The noise terms $\Delta\mathbf{l}_i$'s can be eliminated from the system to yield a nonlinear unconstrained minimisation system given as

$$\min_{\mathbf{e}_\infty} J = \sum_{i=1}^{n} \frac{(\mathbf{l}_i^\top \mathbf{e}_\infty)^2}{\mathbf{e}_\infty^\top \Lambda_{\mathbf{l}_i} \mathbf{e}_\infty}. \tag{8}$$

The common approach to take from here is to replace the denominator term with a constant term w_i often referred to as the *weighting factor* so that the least-squares approach can be adopted. That is,

$$\min_{\mathbf{e}_\infty} J = \sum_{i=1}^{n} \frac{(\mathbf{l}_i^\top \mathbf{e}_\infty)^2}{w_i}. \tag{9}$$

Unfortunately, the solution that minimises J in (9) is known to be not optimal to the original nonlinear problem given in (8), as is now widely known in the literature. Due to the space limit, we cannot cite all the known references. Interested reader is referred to [16, 10] for a similar problem in estimating the fundamental matrix.

Method 2: Generalised eigenvector

The generalised eigenvector method described in [8] is investigated and incorporated into our method is some modification that has shown to improve the estimate of the epipole in hundreds of conducted simulations (not reported here). The main idea is outlined below.

The (weighted) moment matrix M involved in our least-squares problem in (9) is defined as $\frac{1}{n}\sum_{i=1}^{n}\mathbf{l}_i\mathbf{l}_i^\top/w_i$. Since each $\mathbf{l}_i$ that constitutes the moment matrix M can be decomposed into $\bar{\mathbf{l}}_i + \Delta\mathbf{l}_i$, matrix M can be written as the sum of two matrices: $M = \bar{M} + \epsilon^2\Delta M$, where $\bar{M} = \frac{1}{n}\sum_{i=1}^{n}\bar{\mathbf{l}}_i\bar{\mathbf{l}}_i^\top/w_i$ and $\Delta M = \frac{1}{n}\sum_{i=1}^{n}\Lambda_{\mathbf{l}_i}/w_i$. Here, we assume that the covariance matrix $\Lambda_{\mathbf{l}_i}$ estimated in (7) is equal to the true covariance matrix of $\Delta\mathbf{l}_i$ up to an unknown scale factor ϵ^2 (the noise level). Defining

$$\hat{M} \equiv M - \epsilon^2\Delta M$$

gives $\mathtt{E}(\hat{M}) = \bar{M}$, where $\mathtt{E}(.)$ denotes the expected value. That is, $\hat{M}$ is an unbiased estimator of $\bar{M}$. This is a generalised eigenvector problem, with the unknown noise level ϵ^2, as discussed in [8], being chosen as the smallest generalised eigenvalue. The epipole is then estimated as the corresponding generalised eigenvector.

The modification that we apply to the above method is based on the following observation: the definition of M is determined by the weighting factors w_i's, which may not be optimal to start with. Since the weighting factors w_i's can be computed once a reasonably good estimate of the epipole is known, a new copy of matrices M and ΔM, based on the new weighting factors w_i's, should be used. Our simulations show that this modification improves the epipole estimation and, on average, this modified method is comparable to the nonlinear minimisation that will be described later. The procedure for Method 2 is given below:

1. Define each weighting factor w_i in (9) as trace$(\Lambda_{\mathbf{l}_i})$. Compute M and ΔM.
2. Compute the smallest generalised eigenvalue λ such that $M - \lambda\Delta M$ is singular. Let $\mathbf{e}_o$ be the corresponding generalised eigenvector.
3. Recompute each weighting factor w_i as $\mathbf{e}_o^\top\Lambda_{\mathbf{l}_i}\mathbf{e}_o$. Reconstruct M and ΔM.
4. Redo step 2 above. Let ϵ^2 be the smallest generalised eigenvalue. The epipole $\hat{\mathbf{e}}_\infty$ is estimated as the corresponding eigenvector.

Routines for generalised eigendecomposition can be found in most Mathematics software libraries. In Matlab, the `qz` function does the job. We note that in some odd cases both ϵ^2 and $\hat{\mathbf{e}}_\infty$ can be complex; however, our simulations show that the imaginary components of these two entities are far from significance and can be ignored.

Finally, the covariance matrix of the estimated $\hat{\mathbf{e}}_\infty$ is defined as (see [8] for derivation)

$$\Lambda_{\hat{\mathbf{e}}_\infty} = \frac{\epsilon^2}{n}(P\hat{M}P^\top)^- \tag{10}$$

where P=diag$(1, 1, 0)$. We call this covariance matrix $\Lambda^{(2)}_{\hat{\mathbf{e}}_\infty}$.

Method 3: Nonlinear minimisation

This method targets directly at the nonlinear problem given in (8). Given that an initial estimate of the epipole can be obtained from Method 1, our objective is to refine this estimate by minimising the nonlinear system of equations. Again, most standard Mathematics software libraries provide routines to do the job. In this paper, we use the Matlab `fminunc` function whose minimisation algorithm is based on the interior-reflective Newton method. We set all the appropriate options so that the user defined gradient and Hessian matrix G of the residual J are used in the minimisation procedure. The inverse of the Hessian matrix G at the estimated $\hat{\mathbf{e}}_\infty$ is taken to be the covariance matrix of $\hat{\mathbf{e}}_\infty$, i.e. $\Lambda_{\hat{\mathbf{e}}_\infty} = (PGP^\top)^-$, where P is given in (10). We call this covariance matrix $\Lambda^{(3)}_{\hat{\mathbf{e}}_\infty}$.

We note that Propositions 1 and 2 imply that $\mathbf{e}_\infty \sim H^{-1}[o_x\ o_y\ 1]^\top$. Thus, if the image-to-mirror homography is known because explicit landmarks are used, then one can compare the epipoles obtained from these three methods with that produced by the inverse homography.

4. Experiment

The affine reconstruction algorithm and the computation for the epipole described have been implemented in Matlab. One of the conducted experiments is reported here.

Fig. 4 shows the image of a toy. The image has 480×640 pixels and was taken with a Casio QV-100 hand-held digital camera. Superimposed onto the image are the 48 pairs of corners manually detected and, to avoid crowdedness, the epipolar line for every 7th corresponding pair of real and virtual image points. The landmarks on the mirror are for an earlier version of our reconstruction algorithm and are not used in this paper.

The affine coordinate frame $\mathcal{F}_m$ was defined by manually selecting 4 pairs of image points whose world points are known to form a rectangle. The coordinates of the epipole computed by the 3 described methods are shown graphically

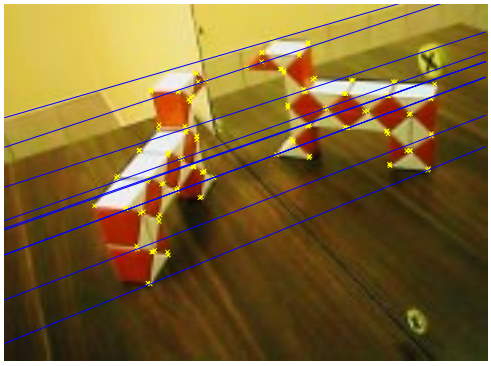
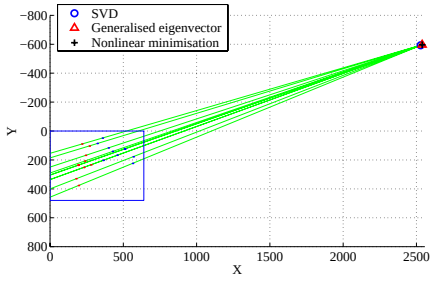


Figure 4. Image of a toy.

in Fig. 5 and numerically in Table 1. If the epipolar lines are parallel then the estimated epipole from any of these methods will have its third component insignificant, signifying that the epipole is in the ideal plane. For Methods 2 and 3, the covariance matrices $\Lambda_{\mathbf{a}_i}$ and $\Lambda_{\mathbf{a}'_i}$ of all corresponding points were set to $\mathrm{diag}(1, 1, 0)$, based on the assumption that the image noise is independent, isotropic, and Gaussian. The covariance matrix $\Lambda_{\mathbf{l}_i}$ was then estimated using formula (7). The residuals J for both methods were estimated to 45.84, which is well below the 99% χ^2 value with $(n{-}2)$ degrees of freedom. The noise level ϵ^2 was estimated to 0.69 from Method 2. The confidence ellipses of the estimated $\hat{\mathbf{e}}_\infty$ from these methods are shown in Fig. 5, using the estimated covariance matrices $\Lambda^{(2)}_{\hat{\mathbf{e}}_\infty}$ and $\Lambda^{(3)}_{\hat{\mathbf{e}}_\infty}$.

(a)



(b)

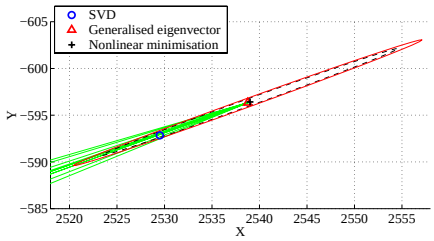


Figure 5. The epipole estimated from the three methods and the corrected epipolar lines. (a) The rectangular box is the boundary of the image. (b) A magnified portion of (a) showing the two confidence ellipses centred at the epipoles computed by Methods 2 and 3.

The two confidence ellipses are almost concentric in this experiment because the epipoles from Methods 2 and 3 are almost identical. However, the confidence ellipse from Method 3 is smaller (dashed black line) than that from Method 2 (solid red line). The reason is that, when differentiating J of (8) with respect to $\mathbf{e}_\infty$ for Method 3, the squared term in the numerator becomes a scale factor of 2 for the gradient and the Hessian matrix. Since this scale factor has no effect on finding the minimum in the iteration process, the common strategy is to save computation time by discarding it. In our implementation for Method 3, however, this scale factor was retained. So our Hessian matrix is twice as large and $\Lambda^{(3)}_{\hat{\mathbf{e}}_\infty}$, being the inverse Hessian, is twice as small. Taking into account the estimated noise level ϵ^2 that Method 2 uses in (10) for rescaling each $\Lambda_{\mathbf{l}_i}$, we found the following relationship between $\Lambda^{(2)}_{\hat{\mathbf{e}}_\infty}$ and $\Lambda^{(3)}_{\hat{\mathbf{e}}_\infty}$ that explains the size difference of the two confidence ellipses: $\Lambda^{(2)}_{\hat{\mathbf{e}}_\infty} \approx 2\,\epsilon^2 \Lambda^{(3)}_{\hat{\mathbf{e}}_\infty}$ (agreement to the first 3 to 4 decimal places). A conclusion that one can draw from here is that it is not meaningful to talk about the size of a confidence ellipse unless all the factors that determine this number are well justified. We also note that while both $\hat{\mathbf{e}}_\infty$ and $\alpha\hat{\mathbf{e}}_\infty$, for any non-zero scalar α, are solutions to (8) and give the same residual, α plays the role of a scale factor to the Hessian matrix. To be consistent with Methods 1 and 2, we scaled the epipole in Method 3 to unit length.

Method	Estimated epipole (pixels)
1	$(2529.51, -592.84)^\top$
2	$(2538.75, -596.33)^\top$
3	$(2539.00, -596.42)^\top$

Table 1. The coordinates of $\mathbf{e}_\infty$ computed from the three methods.

(a) Top view (b) A perspective view

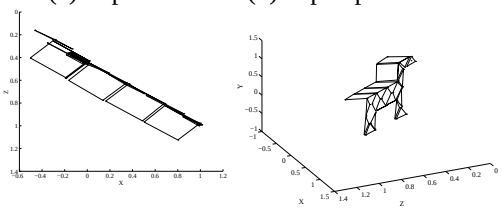


Figure 6. Affine reconstruction of the toy.

In our simulations, the estimates of the epipoles from Methods 2 and 3 were closer to the true epipole than Method 1. This finding is consistent with Luong and Faugeras' [10] report that the nonlinear method is more robust than the linear method in computing the fundamental matrix. Although the difference between the estimates of the epipole from Methods 2 and 3 is insignificant, we have chosen the result from Method 3 for subsequent computation and affine reconstruction. The image coordinates of the midpoints of

world line segments (see Section 3.2) that are required for determining the H matrix in (6) were estimated. The affine reconstruction of the toy is displayed in Fig. 6. Superimposed onto the reconstructed 3D points are 3D line segments for visualisation. It is evident that parallelism was well retained.

5. Discussions

Affine coordinate frame

The affine reconstruction reported in this paper is relative to an arbitrary affine coordinate frame defined on the mirror plane Π_m via the image-to-mirror homography. This is analogous to Irani et al [7] and Weinshall et al's [15] scene reconstruction relative to a floor plane. In particular, their identification of parallel lines on the floor plane for height computation (see also [2]) is also similar, except that in here these parallel lines are determined by four coplanar points on the object and then transferred to the mirror plane Π_m using the cross ratios of the corresponding points with the estimated epipole.

Occlusion and applications

In stereo vision, the occlusion problem refers to the situation where some prominent features that need to be reconstructed are not visible to both cameras; in a camera/mirror system, this refers to the situation where some features are not visible to both the camera and the mirror. Our conducted experiments show that occlusion is minimised when the optical axis of the camera is orthogonal or near orthogonal to the mirror plane normal. While the occlusion problem may be viewed as a drawback of a camera/mirror system, many industrial applications have in fact utilised multiple mirrors to uncover occluded regions of the imaged objects [11]. All these early camera/mirror(s) systems (e.g. [9, 11]) that we found in the literature had the camera explicitly calibrated for Euclidean reconstruction. More recent camera/mirror systems see the inclusion of ellipsoidal, hyperboloidal, and paraboloidal mirrors, e.g. [12]. Interestingly, Nene and Nayar's [12] camera/mirror systems capture two mirror reflections of the scene for analysis while the actual 3D scene is outside the field of view of the camera.

6. Conclusion

We have presented an affine reconstruction method from a single image using the real and virtual corresponding image points created by a plane mirror. We have also given a thorough treatment to the estimation of the epipole that is essential to the reconstruction method. We have demonstrated the reconstruction result on a real image.

References

[1] J. C. Clarke. Modelling Uncertainty: A Primer. Technical Report 2161/98, Univ. of Oxford, Dept. Engineering Science, 1998.

[2] A. Criminisi, I. Reid, and A. Zisserman. Duality, Rigidity and Planar Parallax. In *Proc. European Conf. on Computer Vision*, pages 846–861, Jun 1998.

[3] O. D. Faugeras, Q. T. Luong, and S. J. Maybank. Camera Self-Calibration: Theory and Experiments. In G. Sandini, editor, *Proc. European Conf. on Computer Vision*, pages 321–334, Santa Margherita Ligure, Italy, May 1992.

[4] R. Fawcett, A. Zisserman, and J. M. Brady. Extracting Structure from an Affine View of a 3D Point Set with One or Two Bilateral Symmetries. *Image and Vision Computing*, 12(9):615–622, Nov 1994.

[5] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 3rd edition, 1996.

[6] R. I. Hartley. In Defence of the 8-point Algorithm. In *Proc. Int. Conf. on Computer Vision*, pages 1064–1070, Massachusetts, USA, Jun 1995.

[7] M. Irani, P. Anadan, and D. Weinshall. From Reference Frames to Reference Planes: Multi-View Parallax Geometry and Applications. In *Proc. European Conf. on Computer Vision*, Freiberg, Germany, Jun 1998.

[8] K. Kanatani. *Statistical Optimization for Geometric Computation: Theory and Practice*. Elsevier Science, Amsterdam, 1996.

[9] S. Kaneko and T. Honda. Calculation of Polyhedral Objects Using Direct and Mirror Images. *J. Japan Soc. Precision Eng.*, 52(1):149–155, 1986.

[10] Q. T. Luong and O. D. Faugeras. The Fundamental matrix: Theory, Algorithms, and Stability Analysis. *Int. J. of Computer Vision*, 17:43–75, 1996.

[11] H. Mitsumoto, S. Tamura, K. Okazaki, N. Kajimi, and Y. Fukui. 3-D Reconstruction Using Mirror Images Based on a Plane Symmetry Recovering Method. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(9):941–946, Sep 1992.

[12] S. A. Nene and S. K. Nayar. Stereo with Mirrors. In *Proc. Int. Conf. on Computer Vision*, pages 1087–1094, Bombay, India, Jan 1998.

[13] C. A. Rothwell, D. A. Forsyth, A. Zisserman, and J. L. Mundy. Extracting Projective Structure from Single Perspective Views of 3D Point Sets. In *Proc. Int. Conf. on Computer Vision*, pages 573–582, Berlin, Germany, May 1993.

[14] J. G. Semple and G. T. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1952.

[15] D. Weinshall, P. Anadan, and M. Irani. From Ordinal to Euclidean Reconstruction with Partial Scene Calibration. In *Springer LNCS: 3D Structure from Multiple Images of Large Scale Environments*, Freiberg, Germany, Jun 1998.

[16] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A Robust Technique for Matching Two Uncalibrated Images through the Recovery of the Unknown Epipolar Geometry. *Artificial Intelligence*, 75(1-2):87–120, 1995.