# Extracting Structure from an Affine View of a 3D Point Set with One or Two Bilateral Symmetries

Roger Fawcett, Andrew Zisserman and Michael Brady
Department of Engineering Science
University of Oxford, OX1 3PJ
email: { roger, az, jmb } @uk.ac.ox.robots

**Abstract**

We demonstrate that the structure of a 3D point set with a single bilateral symmetry can be reconstructed from an uncalibrated affine image, modulo a Euclidean transformation, up to a four parameter family of symmetric objects that could have given rise to the image. If the object has two orthogonal bilateral symmetries, the shape can be reconstructed modulo similarity. Both results are demonstrated using real images with uncalibrated cameras.

# 1 Introduction

Many man made objects and animals exhibit reflectional symmetry. We demonstrate that exploiting this symmetry facilitates 3D structure recovery modulo a linear transformation. Such structure has been shown to simplify a number of vision tasks such as model based recognition [9, 15, 19, 20], epipolar calibration [1, 5, 10] and motion transfer and point matching [4, 8, 13].

Many papers have dealt with the extraction of symmetries in images. These have generally assumed that the imaged object is within a similarity transformation of the image (viewed in a fronto-parallel plane) so that the imaging process does not destroy the mirror symmetry [2, 3, 16]. Some authors have included the effects of shear [7, 11, 19], but have generally limited consideration to 2D objects. Here we examine the images of 3D objects with one or more bilateral symmetries.

We assume the affine camera approximation [12]. This has been shown to be effective if the object dimensions are an order of magnitude less than its distance from the camera. A single view of a symmetric object is equivalent to two views, each of half of the object. Consequently, mathematical results established for affine stereo views [4, 8, 12, 17, 18] can be adapted to recover the 3D affine shape (i.e. 3D positions modulo an affinity). Building on this connection with stereo we use the term "epipolar lines" for the images of lines joining corresponding points on each side of the symmetry plane. We show that the extra constraints resulting from the object symmetry allow structure recovery to *better than* an affine transformation.

In the following sections we make precise the degree of ambiguity in the recovered structure, and demonstrate the method on real images. Two of the major features are that camera calibration is not required at any stage and that objects do not have to be of a restricted class, such as polyhedra.

**Notation** We adopt the notation that corresponding points in the world and image are distinguished by large and small letters (e.g. $x$ and $X$). Vectors are written in bold font (e.g $\mathbf{X}$).

BMVC 1993 doi:10.5244/C.7.35

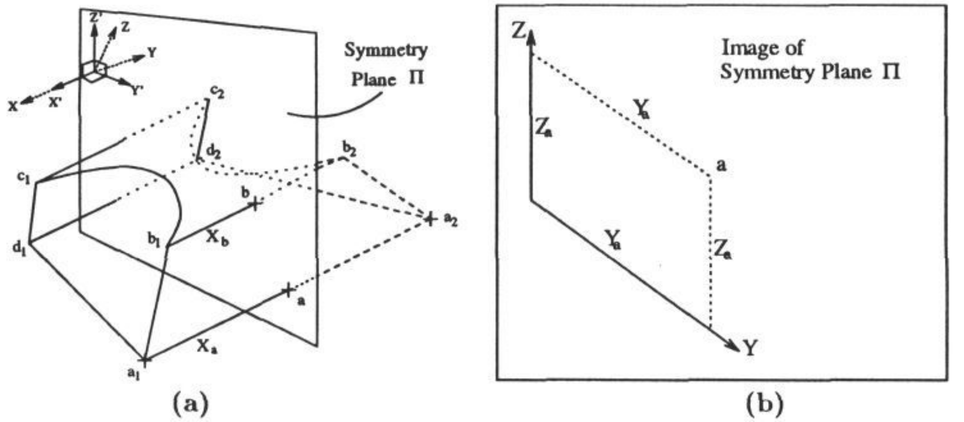



Figure 1: (a) Point $A_2$ is the reflection of point $A_1$ in $\Pi$. These points project to $a_2$ and $a_1$ respectively. Point $a$ is the mid point of the line $a_1a_2$. The $X'$ axis is normal to $\Pi$, and its projection is parallel to the line $a_1a_2$, which defines the $X$ axis. The $Y'$ and $Z'$ axes are in arbitrary orthogonal directions on $\Pi$. They form a 3D orthonormal coordinate frame with $X'$. $X_a$ is the length $a_1a$ measured in the image plane. (b) The affine coordinates $Y_a$ and $Z_a$ are obtained by parallel projection of $a$ onto the $Y$ and $Z$ axes.

# 2 Objects with Single Bilateral Symmetry

Lines joining corresponding object points (on either side of the symmetry plane) are parallel, and orthogonal to the plane of symmetry. Since affine projection preserves parallelism, the imaged lines are also parallel. The optimal extraction of their orientation is discussed in section 2.2. In the following we use a natural object coordinate system provided by the correspondence direction and symmetry plane. Using a construction based on affine projection properties we establish that:

**Theorem 1** *Given an uncalibrated affine image of an object with a single bilateral symmetry, its shape can be reconstructed, modulo a Euclidean transformation, to a four parameter family of symmetric shapes that could have given rise to the image.*

## 2.1 Proof By Construction

We define two 3D coordinate systems. The first is a possibly non-orthogonal affine system, $XYZ$, derived directly from image measurements. The $X$ axis is along the correspondence direction. The origin and the other two axes lie in the symmetry plane. The second system is an object centred orthonormal system, $X'Y'Z'$. Again, the $X'$ axis along the correspondence direction. The other two axes and the origin lying in the symmetry plane. These two coordinate systems are illustrated in figure 1a. They are related by a four parameter linear transformation. Structure recovery in the $XYZ$ system is in two stages, as illustrated in figure 1:

1. $X$ **coordinate** The mid-point, $a$, of two corresponding image points, $a_1a_2$ is constructed. Since mid-points are preserved by affine transformations, this is the projection of the actual mid-point, $A$, of the 3D points, $A_1$ and $A_2$. The point $A$ lies on the symmetry plane. The affine $X$ coordinate of $a_1$ is

the distance $aa_1$ in pixels (a pixel is the unit length). As ratios of lengths on parallel line segments are preserved under affine transformations, the ratio of lengths $X_a : X_b$ is equal to the ratio of world distances $X'_A : X'_B$, i.e. a common scale factor, $\lambda_1$, relates all the coordinates $X$, constructed in this way, to the Euclidean distances $X'$.

2. **$YZ$ coordinates** As mid-points of corresponding points all lie on the symmetry plane, there is a plane-to-plane affine transformation (symmetry plane to image plane) between world mid-points and their images. This transformation is represented by the matrix $M$ (see below). The mid-point construction projects points onto the symmetry plane. The affine $YZ$ coordinates of these points are determined as shown in figure 1b.

In general, a zero translation affine transformation has four degrees of freedom. However the rotation of $X'Y'$ about $Z'$ is not significant and can be disregarded. This is achieved by noting that any non-degenerate matrix can be decomposed into the product of a rotation and a symmetric matrix [6]:

$$M = R(\phi)\begin{bmatrix} e & f \\ f & g \end{bmatrix} \tag{1}$$

Hence affine transformation between the two systems can therefore be accomplished by a symmetric matrix. There only remains then a Euclidian transformation from this orthonormal system to a world Euclidian frame. To summarise:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = T_1 \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad where \ \ T_1 = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & e & f \\ 0 & f & g \end{bmatrix} \tag{2}$$

## 2.2 Epipolar Line Orientation

Since all the imaged correspondence directions are parallel, one point correspondence is sufficient to determine the epipolar structure. However, because of image localisation errors and limits of the affine approximation a more accurate estimate is obtained by aggregating many correspondences. It can be shown that the minimum variance solution for the epipolar line orientation is approximately given by:

$$\omega = \frac{\sum_{i=1}^{n} \omega_i x_i^2}{\sum_{i=1}^{n} x_i^2} \tag{3}$$

where $x_i$ is half the length of the line on the image plane joining the i'th pair of symmetrically related points, $\omega_i$ is its orientation, and $n$ the number of correspondences.

# 3 Objects with Two Orthogonal Bilateral Symmetries

In this case there is a natural *orthogonal* coordinate system consisting of the two orthogonal correspondence directions, and the intersection line of the symmetry planes. This structure can be determined from the image by affine constructions. We prove the following:

**Theorem 2** *Given an uncalibrated affine image of a 3D object with two orthogonal bilateral symmetries, its shape can be reconstructed, modulo a Euclidean transformation, to a three parameter family of symmetric shapes that could have given rise*

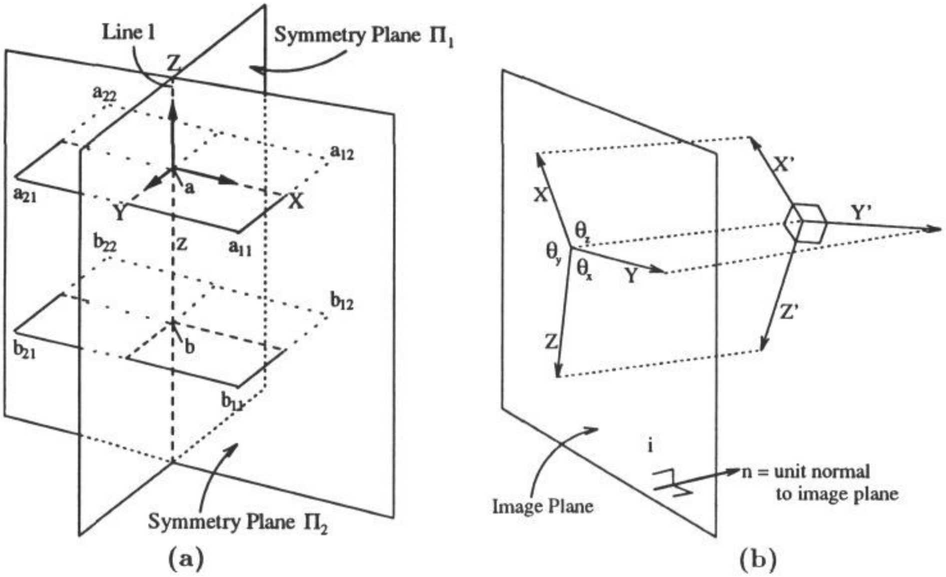



Figure 2: (a) The line $l$ is the image of the intersection of the two symmetry planes $\Pi_1$ and $\Pi_2$. $a_{ij}$ and $b_{ij}$ are the images of reflections of points in the symmetry planes. Their mid-points, $a$ and $b$, lie on $l$. (b) $\mathbf{X}'$, $\mathbf{Y}'$ and $\mathbf{Z}'$ are three orthogonal vectors. $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are the scaled orthographic projections of the orthogonal vectors onto the image plane. The angles on the image plane between $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are $\theta_x$, $\theta_y$ and $\theta_z$, as shown. A unit vector perpendicular to the image plane is denoted by $\mathbf{n}$.

*to the image. If the aspect ratio of the camera is known, the three parameter family reduces to a single scale parameter and the orientation of the object can also be determined.*

## 3.1 Proof By Construction

We first describe recovery up to the three parameter family. The reduction to a single scale factor is described in section 3.2. Here the $XYZ$ and $X'Y'Z'$ coordinate systems are both orthogonal with $X$ and $Y$ axes parallel to the correspondence directions, and $Z$ defined by the symmetry plane intersection. As in section 2, the $X'Y'Z'$ is orthonormal, and the $XYZ$ affine coordinates are determined directly from image measurements.

1. **$X$ and $Y$ coordinates** Refer to figure 2a. As in section 2, $X_a$ is obtained from the mid-point of $a_{11}a_{21}$ (since $A_{21}$ is the reflection in $\Pi_1$ of $A_{11}$). Similarly, $Y_a$ is obtained from the mid-point of $a_{11}$ and $a_{12}$ (since $A_{12}$ is a reflection in $\Pi_2$ of $A_{11}$).
2. **$Z$ coordinates** We first construct the $Z$ axis, which lies on the line $l$, the projection of the intersection of the $\Pi_1$ and $\Pi_2$ planes. The four symmetry related points form a parallelogram in the image. Consider two such parallelograms (these are marked, $a_{ij}$ and $b_{ij}$ in figure 2a. The centre of any such parallelogram will lie on $l$ (since the centre is preserved by affine transformations). Two centres (e.g. $a$ and $b$) define $l$. When there are more than

two sets of symmetry points, a least squares fit is used to obtain line $l$. The parallelogram centre is the projection of the intersection of the plane containing the four symmetry related points with the $Z$ axis. This determines the $Z$ coordinate. Again, image measured $Z$ coordinates are within a constant factor of $Z'$, since there is a linear affine transformation between world and image.

It should be noted that if only three of the four symmetry related points can be seen, the position of the fourth can be constructed. The transformation between the affine and object orthonormal frame is now a scaling along each axis:

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = T_2 \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad where \ T_2 = \begin{bmatrix} \lambda_x & 0 & 0 \\ 0 & \lambda_y & 0 \\ 0 & 0 & \lambda_z \end{bmatrix} \tag{4}$$

## 3.2 Reducing the scaling parameter Ambiguity

The $XYZ$ coordinate system is a projected *orthogonal* coordinate system. In essence, this constraint and the angles between the projected axes are sufficient to determine the 3D orientation of the orthogonal system (up to Necker ambiguity), and reduce the three scaling parameters in equation (4) to a single overall scaling. Measuring angles requires image aspect ratio to be known.

Suppose there are three orthonormal vectors $\mathbf{X}', \mathbf{Y}'$ and $\mathbf{Z}'$ in 3D which project to three image vectors $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ respectively, as shown in figure 2b. From equation (4) the norms of these vectors are related by:

$$X' = \lambda_x X, \qquad Y' = \lambda_y Y, \qquad Z' = \lambda_z Z \tag{5}$$

Under scaled orthography the *vectors* are related as:

$$\mathbf{X}' = \lambda(\mathbf{X} + \alpha\mathbf{n}) \qquad \mathbf{Y}' = \lambda(\mathbf{Y} + \beta\mathbf{n}) \qquad \mathbf{Z}' = \lambda(\mathbf{Z} + \gamma\mathbf{n}) \tag{6}$$

where $\lambda$ is the scaling and $\alpha, \beta, \gamma$ are unknowns. Using the orthogonality, i.e. $\mathbf{X}'.\mathbf{Y}' = \mathbf{X}'.\mathbf{Z}' = \mathbf{Y}'.\mathbf{Z}' = 0$, we obtain the following three equations:

$$\mathbf{X}.\mathbf{Y} = -\alpha\beta \qquad \mathbf{Y}.\mathbf{Z} = -\beta\gamma \qquad \mathbf{Z}.\mathbf{X} = -\alpha\gamma \tag{7}$$

which can be solved for $\alpha$, $\beta$ and $\gamma$:

$$\alpha = -\xi X \stackrel{+}{\sqrt{\frac{\cos\theta_y cos\theta_z}{\cos\theta_x}}} \quad \beta = -\xi Y \stackrel{+}{\sqrt{\frac{\cos\theta_z cos\theta_x}{\cos\theta_y}}} \quad \gamma = -\xi Z \stackrel{+}{\sqrt{\frac{\cos\theta_x cos\theta_y}{\cos\theta_z}}} \tag{8}$$

where $\xi = \pm 1$ and corresponds to the Necker reversal ambiguity. Substituting this into (5) and using (7) we obtain[1]:

$$\begin{bmatrix} \lambda_x \\ \lambda_y \\ \lambda_z \end{bmatrix} = \lambda \begin{bmatrix} \psi_x \\ \psi_y \\ \psi_z \end{bmatrix} \qquad where \tag{9}$$

$$\psi_x = \stackrel{+}{\sqrt{1 - \frac{\cos\theta_y \cos\theta_z}{\cos\theta_x}}} \quad \psi_y = \stackrel{+}{\sqrt{1 - \frac{\cos\theta_z \cos\theta_x}{\cos\theta_y}}} \quad \psi_z = \stackrel{+}{\sqrt{1 - \frac{\cos\theta_x \cos\theta_y}{\cos\theta_z}}}$$

Substituting (9) into (4), it can be seen that the three scale factors in (5) reduce to the single scale, $\lambda$. Thus there is a similarity transformation between recovered structure and the $X'Y'Z'$ frame. As $\alpha$,$\beta$ and $\gamma$ are known, equation (6) can be used to determine the orientation of the object relative to the image plane.

---

[1] only +ve square root is used in (9) because the vector moduli in (5) are positive.

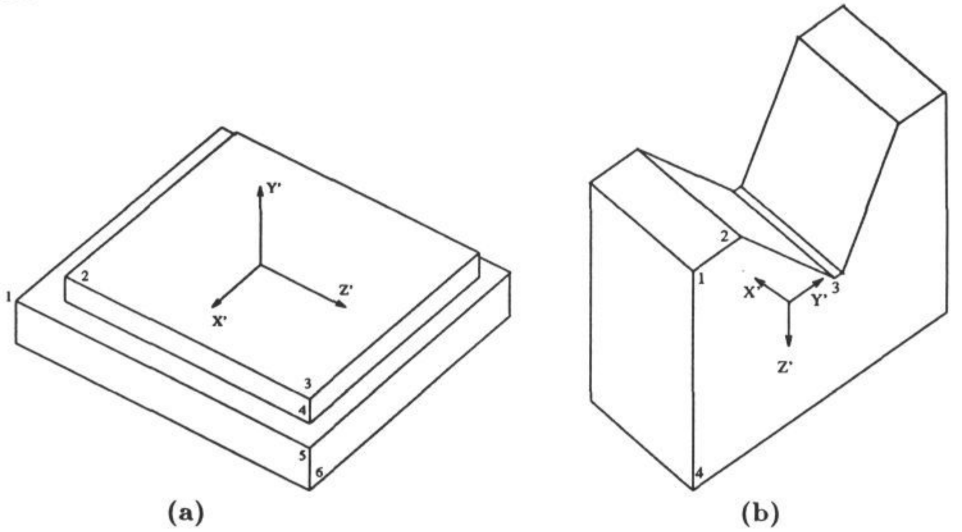



Figure 3: The coordinate system used to measure the $X'$, $Y'$ and $Z'$ values for the table (a) and the v-block (b). The numerical values are shown in tables 1 and 2. In both cases the axes origins were placed at the centre on mass of the corner points, but are shown displaced for clarity.

# 4 Results

The methods described in sections 2 and 3 have been implemented for a number of objects. Symmetrically opposite points are currently chosen by hand and the structure in the $XYZ$ frame calculated. The recovered 3D points are rotated and reprojected at various orientations. Figure 4 shows the single symmetry reconstructions and figure 5 the bisymmetric.

To evaluate the accuracy of the reconstructions, we computed the optimal transformation matrix, T (equation (2) for the single symmetry and the scale factor $\lambda$ for the bisymmetry) between the recovered $XYZ$ coordinates and their actual corresponding $X'Y'Z'$ values *measured* on the object. The T matrix, is then used to transform the calculated coordinates into the $X'Y'Z'$ frame so they can be compared with measured values. The $T$ matrix is evaluated by a least squares minimisation, of equations (2) and (4), yielding: :

**Single Symmetry Solution**

$$M^T = (A^T A)^{-1} B \qquad \lambda_1 = \frac{\sum_{i=1}^n X_i X_i'}{\sum_{i=1}^n X_i^2} \tag{10}$$

$$where \quad A = \begin{bmatrix} Y_1 & Z_1 \\ \vdots & \vdots \\ Y_n & Z_n \end{bmatrix} \qquad B = \begin{bmatrix} Y_1' & Z_1' \\ \vdots & \vdots \\ Y_n' & Z_n' \end{bmatrix}$$

Note that we are also calculating the $R(\phi)$ in equation 1.

**Double Symmetry Solution**

$$\lambda = \frac{\psi_x \sum_{i=1}^n X_i X_i' + \psi_y \sum_{i=1}^n Y_i Y_i' + \psi_z \sum_{i=1}^n Z_i Z_i'}{\psi_x^2 \sum_{i=1}^n X_i^2 + \psi_y^2 \sum_{i=1}^n Y_i^2 + \psi_z^2 \sum_{i=1}^n Z_i^2} \tag{11}$$

| Point No. | Measured Position/mm | | | Reconstructed Position/mm | | | Error /mm | | |
|---|---|---|---|---|---|---|---|---|---|
| | X | Y | Z | X | Y | Z | X | Y | Z |
| 1 | 59.0 | -90.0 | -0.1 | 56.9 | -96.0 | -5.3 | 2.1 | 6.0 | 5.2 |
| 2 | 45.0 | -84.2 | 12.9 | 43.0 | -81.8 | 16.2 | 2.0 | 3.6 | 3.3 |
| 3 | 45.0 | 37.0 | 12.9 | 43.0 | 40.0 | 9.8 | 2.0 | 3.0 | 3.1 |
| 4 | 45.0 | 37.0 | -0.1 | 45.5 | 40.0 | 0.9 | 0.5 | 3.0 | 1.0 |
| 5 | 59.0 | 50.0 | -0.1 | 61.5 | 46.7 | 4.7 | 2.5 | 3.3 | 4.8 |
| 6 | 59.0 | 50.0 | -25.0 | 60.7 | 46.9 | -25.5 | 1.7 | 3.1 | 0.5 |

Table 1: Comparison of recovered and measured corner positions for object an object with a single symmetry. The point numbers correspond to the corners marked in Figure 3(a).

| Point No. | Measured Position/mm | | | Reconstructed Position/mm | | | Reconstructed /mm | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | left image | | | right image | | |
| | X | Y | Z | X | Y | Z | X | Y | Z |
| 1 | 40.0 | 100.0 | -52.5 | 39.6 | 101.1 | -53.4 | 41.2 | 98.7 | -53.7 |
| 2 | 40.0 | 65.0 | -52.5 | 39.1 | 66.4 | -53.1 | 40.0 | 65.6 | -53.4 |
| 3 | 40.0 | 3.0 | 7.5 | 39.4 | 3.6 | 9.7 | 39.1 | 3.0 | 9.3 |
| 4 | 40.0 | 100.0 | 97.5 | 37.3 | 99.0 | 97.2 | 40.3 | 98.9 | 97.9 |
| Angle | | | | Calculated Axis Angles | | | | | |
| | | | | X | Y | Z | X | Y | Z |
| $\sigma$ | | | | 45.6 | 80.2 | 46.1 | 64.6 | 55.0 | 45.7 |
| $\tau$ | | | | 109.4 | 9.6 | 90.0 | 152.4 | 142.1 | 90.0 |

Table 2: Comparison of recovered structure and corner positions measured on the v-block which has two orthogonal symmetries. The point numbers correspond to the corners marked in Figure 3(b). The slant angle, $\sigma$, is the angle between the axes and the image plane normal. The tilt angle, $\tau$, is the orientation of the axes when projected onto the image plane.

The results for the calibration table, shown in figure 4, are given in table 1. The $X$ values are more accurate than $Y$ and $Z$. The block shown in figure 5a and b is viewed from a fixed vantage point at two orientations, differing by a 45 degree rotation about the vertical Z axis. The results in table 2 show both object dimensions (which were calculated to within a scale) and the object orientation. It can be seen by inspection that the Z axis orientation is consistent to within one degree between the two images, and the rotation about the Z axis is calculated to be 45.6 degrees (cf. an actual rotation of 45.0 degrees).

There are two sources of error in the reconstruction. The first is due to the affine approximation to perspective, the second due to point localisation error.

# 5 Discussion

We have demonstrated that structure, modulo a linear transformation, can be successfully recovered from single images of 3D objects with bilateral symmetries. Invariants to this linear transformation (e.g. affine invariants) can be measured and used as indexes in an object recognition system (e.g. [15]).

We have only illustrated structure recovery for point sets. However, since the

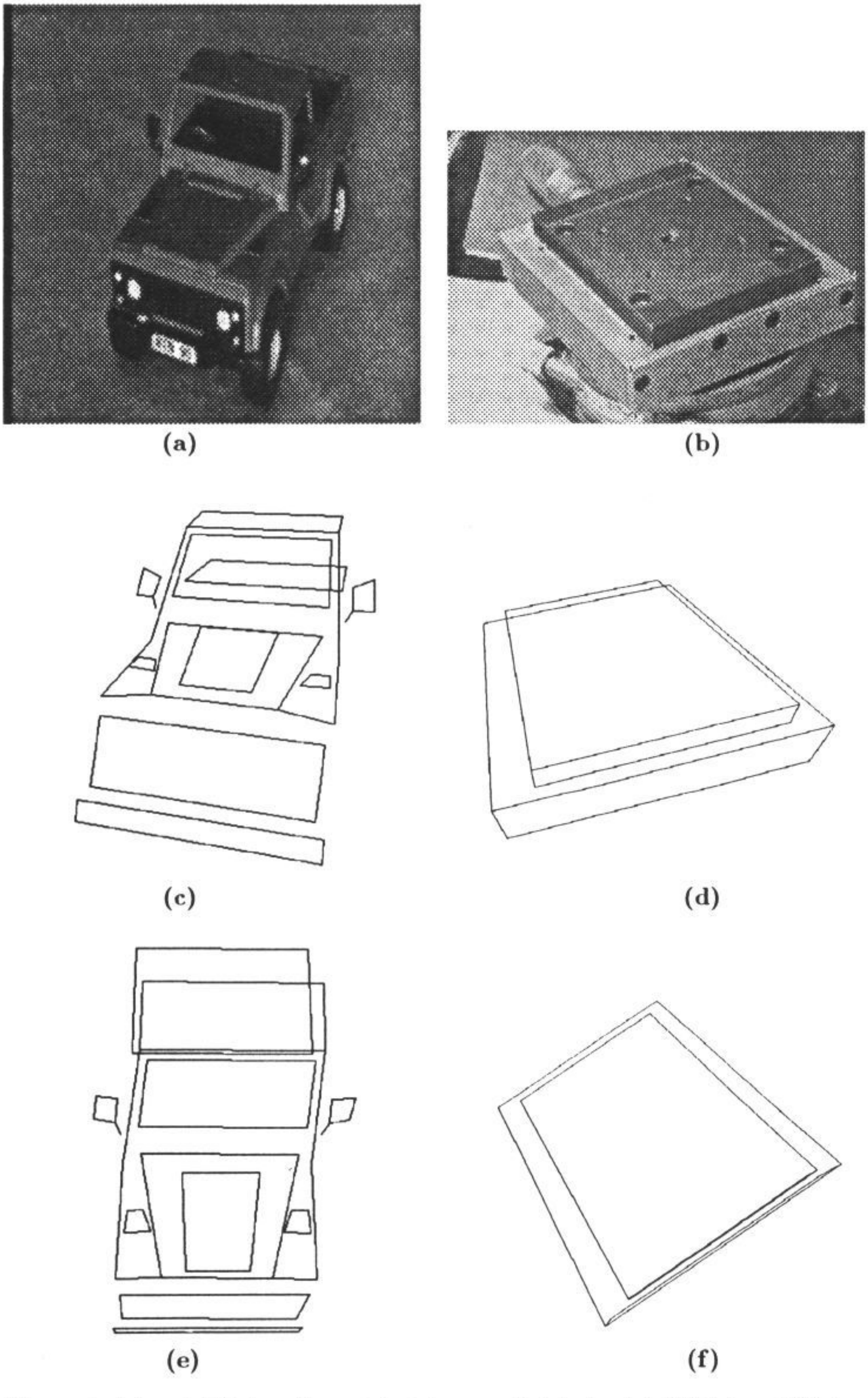
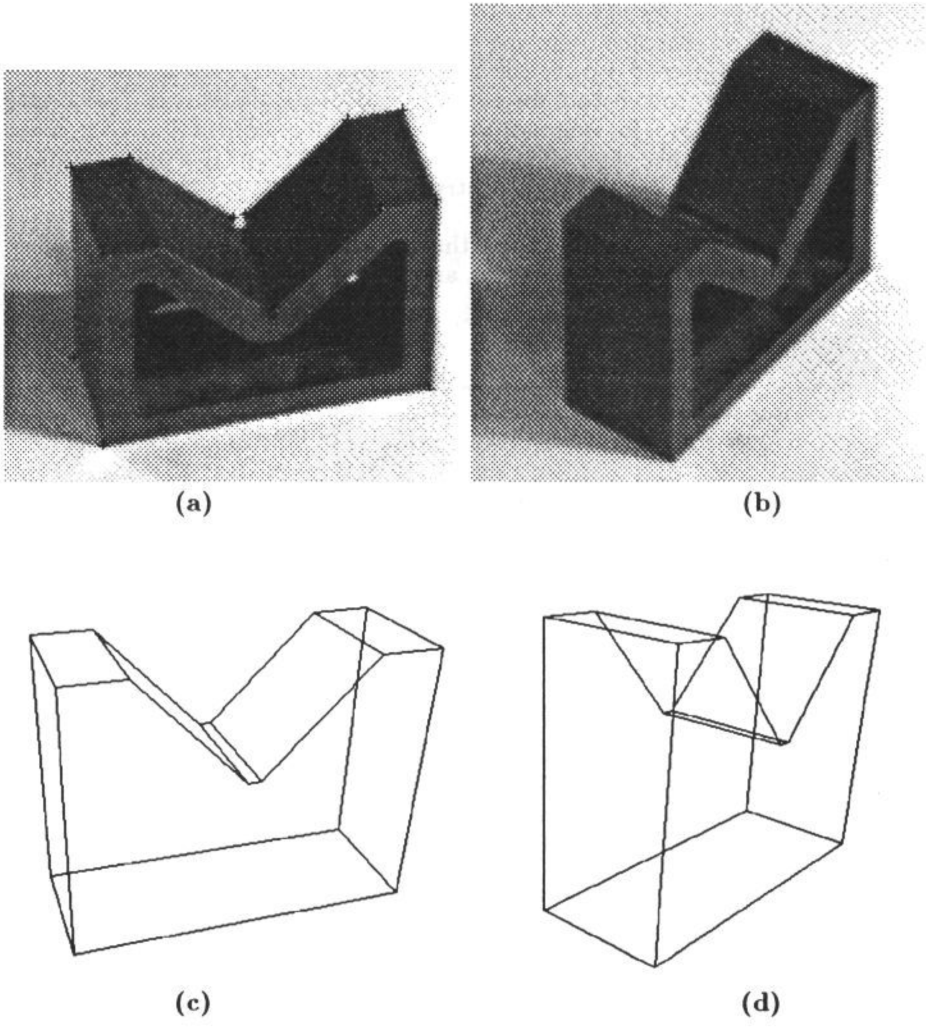



Figure 4: (a) and (b) show the original images of objects. (c), (d) (e) and (f) show the reconstructed 3D objects rotated and perspectively projected.



(a) (b)

(c) (d)

Figure 5: (a) and (b) show the original images a block with two orthogonal bilateral symmetries at two different orientations. (c) and (d) their respective reconstructed 3D shapes, rotated and perspectively projected.

epipolar structure is known, if a symmetric objects contains space curves, the projected curves can be matched pointwise to recover 3D points along their length. To date we have examined two of the most common symmetry cases, but the method can easily be extended to, for example, objects with three-fold rotational symmetry. Also, structure for symmetric objects can be recovered under the more general perspective, rather than affine, projection. Invariants have been obtained in this way for the case of a single symmetry [14].

**Acknowledgements** We would like to thank SERC, Oxford University, Esprit Project and VIVA for financial support, and Philip McLauchlan for assistance with the software.

# References


[1] P.A. Beardsley, D. Sinclair, and A. Zisserman. Ego-motion from six points. Insight meeting, Catholic University Leuven, February 1992.

[2] H. Blum. Biological shape and visual science (part i). *J. theor. Biol.*, 38:205–287, 1973.

[3] J.M. Brady and H. Asada. Smooth Local Symmetries and their Implementation. *Int. Journal of Robotics Research*, 3(3), 1984.

[4] S. Demey, A. Zisserman, and P. Beardsley. Affine and projective structure from motion. *BMVC92*, pages 49–58, 1992.

[5] O.D. Faugeras, Q.T. Luong, and S.J. Maybank. Camera self-calibration: Theory and experiments. *CVPR*, pages 761–764, 1992.

[6] G.H. Golub and C.F. Van Loan. *Matrix Computations 2nd ed.* Johns Hopkins University Press, USA, 1989.

[7] T. Kanade. Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence*, 17:409–460, 1981.

[8] J.J. Koenderink and A.J. Van Doorn. Affine structure from motion. *Journal of Optical Society of America*, 8(2):377–385, 1991.

[9] Y. Lamdan, J.T. Schwartz, and H.J. Wolfson. Visual hyperacuity: representation and computation of very high precision position information. *CVGIP*, 33:99–115, 1986.

[10] R. Mohr. Projective geometry and computer vision. *Handbook of Pattern Recognition and Computer Vision, Chen, Pau and Wang editors*, 1992.

[11] D.P. Mukherjee, A. Zisserman, and Brady J.M.' Shape from symmetry - detecting and exploiting symmetry in affine images. Technical Report OUEL 1988/93, Oxford University, 1993.

[12] J.L. Mundy and editors Zisserman, A. *Geometric Invariance in Computer Vision*. MIT Press, 1992.

[13] L. Quan and R. Mohr. Towards structure from motion for linear features through. *IEEE Workshop on Visual Motion*, 1991.

[14] C.A. Rothwell, D.A. Forsyth, A. Zisserman, and J.L. Mundy. Extracting projective structure from single views of 3d point sets. *International Conf. on Computer Vision*, 1993.

[15] C.A. Rothwell, A. Zisserman, D.A. Forsyth, and J.L. Mundy. Using projective invariants for constant time library indexing in model based vision. *BMVC*, pages 62–70, 1991.

[16] S. Scott G. L., Turner and A. Zisserman. Using a mixed wave diffusion process to elicit the symmetry set. *Proc. of 4th Alvey Vision Conference*, pages 221–228, 1988.

[17] C. Tomasi. *Shape and motion from image streams: a factorisation method.* PhD thesis, Carnegie Mellon University, May 1991.

[18] S. Ullman and R. Basri. Recognition by linear combination of models. *PAMI*, 13(10):992–1006, 1991.

[19] L. Van Gool, P. Kempenaers, and A. Oosterlinck. Recognition and semi-differential invariants. *CVPR*, pages 454–460, 1991.

[20] P.C. Wayner. Efficiently using invariant theory for model-based matching. *CVPR*, pages 473–478, 1991.