Detecting Bilateral Symmetry of 3D Point Sets from Affine Views

Torfi Thórhallsson
Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, U.K.
torfi@robots.ox.ac.uk

Abstract

This paper describes a novel method to detect 3D bilateral symmetry from affine structure as well as directly from affine views.

Applying the notion of skewed symmetry to 3D affine structure, the problem is reduced to detecting geometric degeneracies in sets of 3D points. We present a method of detecting such degeneracies in affine structure in the presence of measurement errors. We further extend this method to be applicable directly to point correspondences in two or more affine views, without the need of explicit 3D reconstruction.

We describe an algorithm, which applies the tests of degeneracy to symmetry detection in an efficient manner. Results are included of performance under noise and application to real images.

1 Introduction

Many man made and natural objects display symmetry to an extent which, if robustly detected, could provide cues for scene segmentation and object recognition [19]. Prior knowledge of 3D object symmetry has also been shown to aid in the recovery of 3D structure information from a single view [13, 20, 5]. Detecting 3D symmetry from a single view is a very difficult problem, however, because of the lack of constraints available. Consequently the bulk of the work published on bilateral symmetry detection falls roughly into two categories: (i) planar symmetry from a single view, and (ii) 3D symmetry from a complete metric 3D description.

We acknowledge the problems involved in detecting 3D symmetry from a single view of unconnected point features, and propose a method to detecting bilateral symmetry from two or more affine views, with the aim of providing information to aid scene segmentation and object recognition.

Before describing the approach of this paper, we give a brief review of previous work on symmetry detection. The various approaches mainly differ in the assumptions made about the viewing conditions, prior segmentation, and connectivity of features. Of the planar methods, those which detect bilaterally symmetric configurations of 2D image points have only limited applicability to images of 3D scenes. See [6] for a comprehensive bibliography.

Skewed symmetry was defined by Kanade [13] as the constraints imposed by bilateral symmetry of planar figures on the geometry of an affine view of the figure.
Image points related by skewed symmetry are connected by parallel line segments (or chords), the midpoints of which are collinear. The line defined by the midpoints is the image of the symmetry axis of the depicted object. Many different methods of detecting skewed symmetry have been proposed [8, 17, 11, 24, 3]. There are two main approaches to reducing the combinatorial complexity of the problem: global and local. Global methods use the entire figure to find the axis of symmetry [8, 11] thereby assuming a segmented image. Local methods use a measure of similarity around each image point to match symmetric point pairs. Most local methods are concerned with the matching of connected contours, making use of contour derivatives for similarity measure [17, 24, 3].

In a perspective view chords are no longer parallel, and their midpoints, as measured in the image, do not correspond to the true midpoints on the object. Glauche et al. [9] propose a method to determine the pose of a bilaterally symmetric planar object from the convex hull, again assuming a segmented scene.

Methods reported for the detection of 3D object symmetry generally assume the knowledge of true Euclidean structure, as opposed to the more easily obtained affine or projective structure. Graph-based methods [12, 7] further assume a complete polyhedral description of the object, as provided by a wire-frame model. Imperfections due to either occlusion or noisy measurement are not considered. These requirements severely limit the applicability of these methods to vision problems.

This Paper

In this paper we consider the problem of detecting bilaterally symmetric point features on a three dimensional object, whose structure is known only up to an affine transformation. Such affine structure can be robustly recovered given either two or more affine views [23, 15], two projective views under pure translation [16], or three projective views under planar motion [1]. We extend the notion of skewed symmetry to 3D affine structure, thereby reducing the problem of symmetry detection to that of detecting geometric degeneracies in sets of 3D points. We develop a method to determine such degeneracies in the presence of non-isotropic uncertainties in point locations.

The information contained in a 3D affine structure can be obtained from two affine views of the scene [23]. This fact has been used to perform computations on a 3D structure in the image domain, without explicit 3D reconstruction [18]. Proceeding in that spirit, we give an alternative method to determine the geometric degeneracies directly from point correspondences in two or more affine views.

We describe an efficient algorithm to detect sets of four or more bilaterally symmetric point pairs, contained within a larger point set. The algorithm is based on successive degeneracy testing, and works with both affine structure and point correspondences in two or more affine views. We present the results of a performance analysis under different noise conditions, as well as an example of applying the algorithm on real views.

The paper is organized as follows. We begin in §2 by describing the algorithm for the detection of 3D skewed symmetry in terms of degenerate structure. In §3 we give a robust method of detecting degenerate affine structure, which we extend in §4 to two or more affine views. Experimental results are given in §5, and conclusions in §6.
Figure 1: Testing a 3D affine structure for skew symmetry involves (a) testing each pair of line segments connecting any two points for 3D parallelism, (b) (optionally) testing parallel line segments for 3D collinearity of midpoints, and (c) testing any four (or more) parallel segments for coplanarity of midpoints.

2 Detecting 3D Skewed Symmetry

Bilaterally symmetric 3D objects have points arranged about a plane of symmetry. The chords connecting each pair of symmetric points are orthogonal to this plane. The midpoints of the points lie all in the plane of symmetry. In a single affine view of such objects, the chords remain parallel, but the coplanarity of midpoints does not constrain their location in the image. With the single constraint of parallel chords, the problem of detecting 3D bilateral symmetry from a single affine view is underconstrained.

The affine structure of a 3D object retains both the parallelism of chords and the coplanarity of the midpoints. The implications are that we have two constraints to test for, instead of only one as in the case of a single affine view. Furthermore we have a stronger constraint on parallelism, as the chords must be parallel in 3D, rather than just in one view. The only constraint lost in the reconstruction up to an affine transformation is the perpendicularity of chords to the plane of symmetry. By analogy to the case of an affine view of a bilaterally symmetric planar figure [13], a structure satisfying these two constraints can be termed 3D skewed symmetric.

An Algorithm for the Detection of 3D Skewed Symmetry

To test a set of points in an affine structure for 3D skewed symmetry, we check for parallelism between the line segments connecting the points and for coplanarity of the midpoints. In order for the latter test to be conclusive it must be performed on a minimum of four segments, or eight points, from the set. Applying the test exhaustively would result in an algorithm with a complexity of $O(n^8)$, where $n$ is the total number of points in the set. However the test for parallelism requires only a minimum of two segments, or four points, so applying it exhaustively requires $O(n^4)$ operations. Using the fact that only parallel segments need to be tested for coplanarity of midpoints, a more efficient algorithm for detecting skewed symmetry can be devised. We can achieve a further increase in efficiency by noting, as did Ponce [17] in the case of 2D skew symmetry, that a test of parallelism cannot succeed if line segments have distinct spatial directions. We can generate the $\frac{n^2}{2}$
possible line segments connecting \( n \) points, and distribute them into \( k^2 \) bins of a table, according to the spatial direction of each segment. The tests of parallelism and coplanarity need only be performed within each bin, which on average contains approx. \( \frac{n}{2} \) segments. In addition to the tests of parallelism and coplanarity, we test for very short line segments and for collinearity of midpoints in order to suppress trivial matches. The spatial direction of very short line segments cannot be accurately measured, giving rise to many false detections of parallelism. Similarly does a collinear point set form a planar structure with any additional point.

3 Testing for Degenerate Structure

To carry out reliable detection of 3D skewed symmetry, as described in the last section, we need to be able to detect coplanarity, collinearity, and parallelism of 3D points in a reliable way. In this section we describe such tests based on the notion of degenerate structure.

Given a number of 3D points \( \mathbf{X}_p = (X_p, Y_p, Z_p)^T, p = 1, ..., P, \) we can form the \( 3 \times P \) structure matrix \( S \) consisting of the column vectors \( \mathbf{X}_p - \mathbf{X} \), where \( \mathbf{X} \) is the centre of mass of the point set. We then have the relation between the rank of this structure matrix and the geometric configuration of the \( P \) points (or vectors) as given in the second column of Table 1. The rank of \( S \) can be obtained from

<table>
<thead>
<tr>
<th>Structure</th>
<th>Rank ( S )</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>general</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>coplanar</td>
<td>2</td>
<td>( \lambda_3 )</td>
</tr>
<tr>
<td>collinear</td>
<td>1</td>
<td>( \lambda_2 + \lambda_3 )</td>
</tr>
<tr>
<td>coincident</td>
<td>0</td>
<td>( \lambda_1 + \lambda_2 + \lambda_3 )</td>
</tr>
</tbody>
</table>

Table 1: The relation between degenerate structure and the rank of the structure matrix in the absence of errors, as well as least squared distances from the closest degenerate structure.

the singular value decomposition of \( S \), as the number of nonzero singular values equals the rank \([10]\). Equivalently we can count the non-zero eigenvalues of either of two symmetric matrices: the \( 3 \times 3 \) matrix \( SS^T \) or the \( P \times P \) matrix \( S^T S \), as the nonzero eigenvalues of either of these matrices equal the singular values of \( S \).

Contaminated Affine Structure

When the measurements of the structure are contaminated by noise, the best we can do is to measure the deviation of the structure matrix from coplanarity or collinearity.

By the minimum property of principal axes \([4]\) the least sum of squared distances of the points \( \mathbf{X}_p \) from a degenerate structure is given by the eigenvalues \( \lambda_1 \geq \lambda_2 \geq \cdots \) of \( SS^T \) or \( S^T S \) as shown in the third column of Table 1. If we could assume an isotropic distribution of errors around each point \( \mathbf{X}_p \), we would be able to decide directly from the size of these residual errors, whether the underlying structure was likely to be of a certain degenerate type. The assumption of isotropic errors is hard to justify. Affine structure is non-metric, which means we know nothing about the scale factors and shear angles relating it to the real
structure. If the errors can be attributed to image measurements however, their effect on the structure matrix can be established.

Under the assumption of Gaussian errors in the location of features in each of the views used to calculate the affine structure, the resulting uncertainty in the structure can be estimated in the form of the $3 \times 3$ covariance matrix $\Sigma_p$ associated with each 3D point $X_p$ [15]. This covariance matrix depends on the relative orientation of object and camera in each of the views determining the structure, and is generally not isotropic.

We now assume that all points $X_p$ are known to have the same\footnote{In the case of structure from affine views [33, 15] this condition is true when all the points are reconstructed from the same views.} zero mean Gaussian error distribution with a symmetric positive definite covariance matrix $\Sigma$. It is then possible to apply the linear Mahalanobis-transformation [14] to each point

$$Z_p \overset{\text{def}}{=} \Sigma^{-\frac{1}{2}}X_p,$$

where $\Sigma^{-\frac{1}{2}}$ is defined using the spectral decomposition $\Lambda \Gamma^T$ of $\Sigma$ as

$$\Sigma^{-\frac{1}{2}} \overset{\text{def}}{=} \Lambda^{-\frac{1}{2}}\Gamma^{-T}, \quad \Lambda^{-\frac{1}{2}} \overset{\text{def}}{=} \text{diag}(\lambda_i^{-\frac{1}{2}}).$$

The effect of the Mahalanobis-transformation is that the error distribution of the transformed points $Z_p$ has isotropic unit covariance. It should be noted that by applying the Mahalanobis-transformation we not only transform the error distribution, but also the structure itself. However as affine structure is only defined up to an arbitrary linear transformation, no information is lost.

When the Mahalanobis-transformation is applied to each point in the matrix $S$, we obtain the transformed structure matrix $S_z \overset{\text{def}}{=} \Sigma^{-\frac{1}{2}}S$. Detecting degeneracy in the transformed structure thus consists of finding the eigenvalues of either of the two symmetric matrices:

$$S_zS_z^T = \Sigma^{-\frac{1}{2}}S\Sigma^{-\frac{1}{2}} \quad \text{or} \quad S_z^TS_z = S^T\Sigma^{-1}S,$$

of respective dimension $3 \times 3$ and $P \times P$, the latter form being computationally more efficient for small $P$. With the errors so normalised, the sum of squared errors is $\chi^2$-distributed and can thus be interpreted directly in terms of probability of a particular type of degeneracy occurring. Selecting a suitably low probability $\alpha$ with which a true hypothesis might still be rejected, the decision on the nature of the underlying structure is reached by means of the tests given in Table 2. Using

<table>
<thead>
<tr>
<th>Test</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_3 &gt; \chi^2_{P-3, \alpha}$</td>
<td>non-coplanar</td>
</tr>
<tr>
<td>$\lambda_2 + \lambda_3 &gt; \chi^2_{P-2, \alpha}$</td>
<td>non-collinear</td>
</tr>
<tr>
<td>$\lambda_1 + \lambda_2 + \lambda_3 &gt; \chi^2_{P-1, \alpha}$</td>
<td>non-coincident</td>
</tr>
</tbody>
</table>

Table 2: Tests with which to reject the hypothesis of a degenerate structure. The number of points $P$ in the structure determines the $\text{dof}$ of the $\chi^2$ distribution as shown.

the tests in Table 2 we obtain the tests of parallelism of chords, collinearity and coplanarity of midpoints needed to detect 3D skewed symmetry, by the algorithm described in §2. The parallelism test is implemented by checking for collinearity of the vectors $V_p \overset{\text{def}}{=} X_p - X_p^t (p = 1, 2)$ defining the orientation of each chord (see Fig. 1(a)).
4 Detecting 3D Bilateral Symmetry from two or more Affine Views

We will now give a method to detect 3D bilateral symmetry directly from two (or more) affine views, without explicit 3D reconstruction. To achieve that we only need to give methods for the detection of degenerate structure from two or more views. The algorithm given in §2 then becomes applicable unchanged.

Degenerate Structure in two or more Views

Given \( F \geq 2 \) affine views of a set of \( P \) scene points, whose correspondence between the views is known, we denote the image of the scene point \( \mathbf{X}_p \) in view \( f \) by \( \mathbf{x}_{fp} = (x_{fp}, y_{fp})^T \). We form the form the \( 2F \times P \) measurement matrix

\[
W \overset{\text{def}}{=} \begin{bmatrix}
\mathbf{x}_{11} - \overline{\mathbf{x}}_1 & \cdots & \mathbf{x}_{1P} - \overline{\mathbf{x}}_1 \\
\vdots & \ddots & \vdots \\
\mathbf{x}_{F1} - \overline{\mathbf{x}}_F & \cdots & \mathbf{x}_{FP} - \overline{\mathbf{x}}_F
\end{bmatrix},
\]

where \( \overline{\mathbf{x}}_f \) denotes the centre of mass of the image points in view \( f \). Tomasi and Kanade [23] showed \(^2\) that the measurement matrix \( W \) can be factored into two matrices \( R \) and \( S \),

\[
W = RS,
\]

where \( R \) is a \( 2F \times 3 \) matrix representing the camera orientation and \( S \) is the \( 3 \times P \) structure matrix from §3. Tomasi and Kanade showed by this that the matrix \( W \) can at most have rank 3. Using eqn. (2) as a model of the imaging process, it can easily be shown\(^3\) that if one of the matrices on the right hand side, e.g. the matrix \( S \), is rank deficient the rank of \( W \) is reduced accordingly. More precisely:

**Extended Rank Theorem** If \( F \geq 2 \) and the matrix \( R \) is of full rank, then the rank of the matrix \( W \) equals the rank of the matrix \( S \).

It should be noted that if the camera orientation matrix \( R \) is rank deficient, \( W \) and \( S \) are not necessarily of the same rank. This happens when the motion between views is a pure translation or a rotation about the optic axis [21].

It follows from the extended rank theorem that the relation between the rank of \( S \) and the spatial configuration of 3D points, as given in Table 1, also holds when \( S \) is replaced by the matrix \( W \). Similarly it can be shown that the least squared distances of the image points \( \mathbf{x}_{fp} \) from degenerate structures is as given in Table 1, whereby the eigenvalues are those of either of the symmetric matrices \( WW^T \) or \( W^T W \).

Using the fact that the image plane is metric, it may be quite reasonable to make stronger assumptions about the distribution of image errors. If we assume equal and uncorrelated errors in the \( x \) and \( y \) coordinates of all views, the Mahalanobis transformation matrix \( \Sigma^{-\frac{1}{2}} \) takes the particularly simple form of a scaled identity matrix. In that case we only need to find the eigenvalues, \( \lambda_1 \geq \lambda_2 \geq \cdots \), of either of the two symmetric matrices \( WW^T \) or \( W^T W \) and determine

\(^2\)The measurement matrix, as defined here, is equal to the registered measurement matrix defined in [23] up to row ordering. This difference does not affect the cited results.

\(^3\)A proof is given in [22].
Figure 2: (a) The effect of different choices of the rejection probability, $\alpha$, from a large range. Shown is the average size of the largest set of 'chords' detected, when 10 chords of bilateral symmetry are present (top) and absent (bottom). The image noise level is $\sigma = 1.0$ pixels. (b) ‘Operating curves’ summarising the performance of the detector at different noise levels, from bottom $\sigma = 0.5$, 1.0, and 2.0 pixels. The curves are traced out by varying $\alpha$ as in (a). The horizontal axis shows the average number of chords missed, when ten chords are present. The vertical axis shows the average number of false chords in the largest detected set, when no chords are present. All sets contain 40 points.

The underlying structure by means of the tests presented in Table 3, which now incorporates the uniform standard deviation $\sigma$ of the image error distribution.

<table>
<thead>
<tr>
<th>Test</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_3 &gt; \sigma^2 \chi^2_{p-3,\alpha}$</td>
<td>non-coplanar</td>
</tr>
<tr>
<td>$\lambda_2 + \lambda_3 &gt; \sigma^2 \chi^2_{p-2,\alpha}$</td>
<td>non-collinear</td>
</tr>
<tr>
<td>$\lambda_1 + \lambda_2 + \lambda_3 &gt; \sigma^2 \chi^2_{p-1,\alpha}$</td>
<td>non-coincident</td>
</tr>
</tbody>
</table>

Table 3: Tests with which to reject a hypothesis of a degenerate structure directly from two or more views. $P$ is the number of feature correspondences.

5 Experimental Results

We have implemented both versions of the algorithm described in §2, using affine structure (§3) or pairs of views directly (§4). The experiments presented here were conducted using the two view version. For that version a 2D direction array was used, indexed by the orientation of a line segment in each of the two views.

5.1 Noise Behaviour

In order to evaluate the degradation of the proposed method of symmetry detection with image noise, we conducted experiments using a set of synthetically generated views, with noise added to the feature position.
The experiment was conducted using synthetic pairs of views of point sets. Each set contained 40 points, half of which formed bilaterally symmetric chords. The shape of each point set as well as the parameters of the affine cameras were chosen randomly for each pair of views. The size of the object in each image was typically between 100 and 300 pixels in diameter. To the projected point position we added isotropic Gaussian noise of standard deviation $\sigma = 0.5, 1.0, \text{ and } 2.0$ pixels. For each noise level we ran the symmetry detector with several different settings of the user selected rejection probability, $\alpha$ (Table 3). For each setting of these parameters we determined the average number of detected and missing chords over 20 different image pairs. We repeated the experiment using non-symmetric sets to determine the average number of ‘chords’ in the largest false plane of symmetry returned by the detector. The results are summarised in Fig. 2.

In this experiment the average number of subsets tested for compliance with skewed symmetry varied from around 1,400 to over 10,000 with $\alpha$ and $\sigma$. In comparison is the total number of different sets of 10 or less correspondences on the order of $10^{22}$.

### 5.2 Example of Real Views

In this example we use two views of a model chair. Grey value corner features were matched automatically between views using the matcher of Beardsley et. al. [2]. The 21 features generate 210 candidate chords, which were painted into a $30 \times 30$ direction array. The contents of the three most populated bins are shown in Fig. 3.

The tests for degenerate structure were initially conducted with image feature standard error estimated at $\sigma = 1.0$ pixels and a rejection probability $\alpha = 10^{-1}$. This value turned out to be too high for the test of parallelism, as the horizontal lines on the front and back of the chair were not classified as parallel. This may be the effect of a slight perspective distortion in the images. Consequently the rejection probability for the test of parallelism alone was lowered to $\alpha = 10^{-3}$.

Running the algorithm with these parameter values resulted in only four of the 104 bins originally occupied containing possibly symmetric point sets. Only one
Figure 4: Results from the application of the tests of 3D parallelism of chords and coplanarity of midpoints to the segments in each bin in the direction array (Fig. 3). After applying the test, as described in the text, only five bins contained sets of candidate chords consistent with 3D skewed symmetry. Only one of the bins contains sets of more than four candidate chords. Shown here are the two largest sets: (a) containing 6 ‘chords’ and (b) containing 5 ‘chords’.

of these bins contained sets with more than the minimum four line segments. This bin contained four sets of five candidate chords. Disabling the collinearity check, included to eliminate trivial although possibly correct sets of segments, increased the size of the largest set to six chords, with the addition of only a few sets of four chords. The resulting sets are shown in Fig. 4. When the rejection probability of the coplanarity test was raised to $\alpha = 0.2$, the smaller set (Fig. 4 (b)) was rejected as non-coplanar, maintaining the set of six candidate chords (Fig. 4 (a)) as the only set of more than four segments.

The total number of sets examined by the algorithm to reach the set of six chords amongst the 21 data points was 592 in this example. The total number of different sets of four to six chords connecting any of the 21 data points is close to $10^6$. The running time was about 5 seconds on a Sun Sparc IPX.

6 Conclusions

We have presented a method of detecting 3D skewed symmetry of unconnected point features affine structure or equivalently from two or more affine views. The method takes into account non-isotropic measurement noise, and is robust to outliers, mismatches, and occlusions as long as at least four chords are visible and correctly matched. The combinatorial complexity is reduced by successive testing of degenerate structure. The complexity is considerable however, as no assumptions are made of either prior segmentation or connectivity of features.

In future work we will investigate the use of connected features and forms of verification for increased efficiency, the possibilities of relaxing the requirement of affine views, as well as the application to scene segmentation.

References


