Grouping and Structure Recovery for Images of Objects with Finite Rotational Symmetry

Jane Liu, Joe Mundy GE Corporate Research Center Schenectady, NY 12309, USA Andrew Zisserman Department of Engineering Science University of Oxford, Oxford OX1 3PJ, UK

Abstract

It is shown that 3D objects with discrete rotational symmetry induce geometric relations in the image. Symmetry related points on the object are imaged to points which satisfy these image geometric relations. These relations are unaffected by camera calibration (interior orientation) and object pose. The relations can be utilised to group features in the image arising from the object. Furthermore, 3D structure of the object can be recovered from a single image up to a specified ambiguity. The paper illustrates these mechanisms on real image examples.

1 Introduction

A number of *classes* of 3D structures have been identified which permit 3D invariants to be measured in a single image view of the structure, e.g. rotational [6]. An important example of such classes is where the 3D object is repeated by a symmetry operation, for example bi-lateral symmetry [9]. The main theme of the work presented here is the use of invariant properties of such symmetries for image grouping [12].

As is well known, in building any automatic recognition system working with images of real scenes acquired under ambient conditions, a significant barrier to successful recognition is the extraction of feature groups which correspond to individual object boundaries. Outline curves found using state-of-the art edge detectors will have "drop outs", incomplete and incorrect topology and extraneous segments. To overcome such problems, a grouping stage is incorporated that typically involves associating curve segments based on a combination of proximity, connectivity and "heuristics". The key idea here is that a geometric class in 3D defines relations which must hold in the image between points on the perspective projection of the object. This class-based grouping can thus provide a principled basis for such heuristics.

In this paper, we explore the class of 3D objects with discrete rotational symmetry. The finite rotation group induces strong projective constraints on an image projection of the object and provides a useful extension as a recognition primitive for many man-made objects such as a hexagonal bolt or a triangular truss girder. Prior work is that of Van Gool *et al.* [11], who considered planar discrete rotational symmetries under projective transformations. This paper extends these ideas to 3D structures. We will demonstrate that it is possible to recover Euclidean properties, such as angle, directly from an uncalibrated image of an object taken from an unknown viewpoint.

2 Discrete Rotational Symmetry

Structures that repeat in a single image of a scene are equivalent to multiple views of a single instance of the structure. Thus, for example, a view of two similar cars in a car park where the cars are parked within translations of one another, is equivalent to a stereo pair of images of one such car, with the cameras related by a pure translation. The 3D shape of the car can be recovered by the familiar techniques of stereopsis. More formally,

A repeated structure is defined by a geometric structure S, and a 3D transformation T, which generates a transformed *copy* of S, i.e., S' = T(S). Both S and S' are viewed in the same perspective image.

In many cases the internal calibration parameters of the camera will be unknown. In this case a single image of a repeated structure is mathematically identical to an uncalibrated stereo pair where the two cameras are related by the transformation between S and S'. It has been shown by Faugeras [3] and Hartley *et al.* [4] that if one carries out stereo reconstruction from two uncalibrated perspective images, the reconstruction can differ from the actual 3D Euclidean geometry of the object by a 3D projective transformation. Thus, 3D projective invariants of this recovered structure have the same value as projective invariants measured on the actual Euclidean structure. Here the transformation in 3D is given by the following:

Definition 1 Discrete Rotational Symmetry *The transformation* T *which repeats the structure in 3D is* $T = \mathbb{R}$ *, where* \mathbb{R} *is an n-fold rotation matrix, i.e.* $\mathbb{R}^n = \mathbb{I}$ *, and* \mathbb{I} *is the identity matrix.*



Figure 1: a) a bolt with sixfold symmetry. The marked point in the center is the computed eigenvector \mathbf{p} corresponding to the center of rotation. The point \mathbf{p} is close to that determined by a projectively invariant construction for the center. c) the vanishing line, **l**.

3 Image Relations

3.1 Planar Relations

Corresponding symmetrical points on the object lie on planes perpendicular to the rotation axis. In the image plane, symmetric points on these planes are related by a cyclic collineation T of period n. That is, the 3D discrete rotational symmetry induces a planar symmetry of the same form, $T^n = I$, where T is the 3×3 homogeneous matrix representing the transformation. The plane projective transformation, T, has three fixed points described by its eigenvectors, p, c and \bar{c} . The eigenvector, p, is real and determines the center of rotation in the image plane. The other two eigenvectors are complex conjugates and correspond to the circular points of the planes perpendicular to the rotation axis [10]. Since these planes are parallel, their common line of intersection is the vanishing line for the planes. This line is given by:

$\mathbf{l} = \mathbf{c} \times \bar{\mathbf{c}}$

The eigenvalues of T are, up to a single scale factor, $1, e^{i\theta}, e^{-i\theta}$, where θ is the angle of rotation (i.e. $\theta = 2\pi/n$). Consequently, the rotation angle of the symmetry, which is a Euclidean invariant, can be recovered directly from the uncalibrated perspective image.

An example of discrete rotational symmetry in the plane is shown in figure 1.

In the figure, a projective construction using triangles is shown which defines the center of rotation. That is, three symmetric points are joined to form a triangle. Three additional symmetric points define a second triangle. The intersection of the edges of the two triangles define two lines which must intersect at the center of rotation, as shown. The center constructed from the eigenvectors of T, is very close to the center constructed from these incidence relations. The rotation angle, computed from the eigenvalues of T, is 60.17 degrees.

3.2 3D objects

The previous section demonstrated the discrete rotation of planar structures. More generally, the symmetry constraints can be derived from a set of 3D points. In this approach, the symmetrically corresponding points in a single image are treated as multiple views of a single pointset. The discrete symmetry of the object constrains the form of the fundamental matrix F [2] computed from point correspondences. The fundamental matrix is defined by the epipolar constraint,

$$\mathbf{x}^{\prime +} \mathbf{F} \mathbf{x} = \mathbf{0} \tag{1}$$

The epipolar constraint expresses the fact that a point in one image defines a corresponding line in a second image which is the projection of the point in space as it moves along the ray through the center of projection of the first camera and its position in the first image. The line is given by, Fx, which is a projective correlation, interpreting the fundamental matrix as a transformation matrix. We can define an analogous epipolar geometry in a single view of a repeated structure.

When \mathbf{x}' and \mathbf{x} , are related by a discrete symmetry, there are points in the image which are fixed under the symmetry mapping. In our case, the image projection of the axis of rotation is fixed under the image transformation which maps points between periods of the discrete rotation. In addition, the line in the image which contains the common vanishing points of the planes perpendicular to the axis of rotation is also fixed.

These fixed points are specified by setting $\mathbf{x}' = \mathbf{x}$ [1]. In this case the epipolar constraint becomes a quadratic form,



Figure 2: a) The pagoda has six-fold symmetry. b) The white dots mark the set of points used to construct the rotational symmetry. Correspondences among these points are used to compute the fundamental matrix (see text). The lines drawn in black are the projected rotation (symmetry) axis, and the vanishing line for planes perpendicular to the rotation axis. These lines are computed from the fundamental matrix.

and the image location of these points is only affected by the symmetric part of F. Let $F = F_S + F_A$, where F_S is the symmetric part and F_A is the anti-symmetric part of F. Note that $\mathbf{x}^{\mathsf{T}} F_A \mathbf{x} = 0$ identically, since $F_A^{\mathsf{T}} = -F_A$. That is,

$$(\mathbf{x}^{\mathsf{T}}\mathbf{F}_{A}\mathbf{x})^{\mathsf{T}} = -(\mathbf{x}^{\mathsf{T}}\mathbf{F}_{A}\mathbf{x}) = 0$$

Therefore the anti-symmetric part contributes nothing to the quadratic form. The fundamental matrix is a rank two matrix, but in general the symmetric part, F_S , will be full rank. However, it can be shown [7] that if the rotation axis is perpendicular to the translation direction, or the translation is zero as it is here, the symmetric part of the essential matrix drops rank. It follows that F_S is also not of full rank (since the essential matrix is transformed to the fundamental matrix by full rank intrinsic parameter matrices), and the conic $\mathbf{x}^{\mathsf{T}} \mathbf{F}_S \mathbf{x} = 0$ degenerates to two distinct lines. One line, \mathbf{l}_A , is the image projection of the axis of rotation and the other line, \mathbf{l}_{∞} , is the vanishing line of the planes perpendicular to the rotation axis. The intersection of these two lines is at the null vector of \mathbf{F}_S .

In principle the fundamental matrix in this case can be computed from the correspondence of six points, since the matrix has six degrees of freedom. In general seven correspondences are required to compute one or three solutions for F. The homogeneous matrix has eight degrees of freedom (arising from the ratios of the nine elements), but in this case the fundamental matrix satisfies two (cubic) constraints. One arising from det F = 0, the other from det $F_S = 0$.

An illustration of the recovery of these symmetry elements is shown in Figure 2. A number of points are selected in the image which lie on positions which are equivalent under the discrete 3D rotation. The fundamental matrix is computed and defines the projection of the axis of symmetry in the image, as just discussed. Note that the axis is slightly perturbed from its ideal value, due to errors in the location of the original set of image points. The vanishing line of the planes perpendicular to the rotation axis is also shown.

As has been demonstrated above, in the planar relations, the 3D geometry of the structure can be reconstructed up to a similarity transformation on the planes perpendicular to the axis of rotation. However, the 3D reconstruction ambiguity is projective along the axis of rotation (i.e. a 1D projective transformation).

It is possible to form similar constructions to that shown in Figure 1 in order to define distinguished points on the axis of rotation. These distinguished points can then be used as an invariant description of the object, along with the known discrete rotation angle. Of course, many correspondences will be eliminated by the self-occlusion inherent in a 3D object.

4 Grouping

Up to this point we have described the image relationships, and demonstrated that they can be recovered from real images. We now discuss the issue of what grouping strategies to adopt in order to best utilise these relationships. This grouping task has many similarities with the problem of establishing a correspondence between image features and model features in a model based recognition system [5].

In model based recognition, two alternative strategies are used to establish a transformation between the 3D object model and the image. The first strategy is called hough or pose clustering, where each image-model correspondence votes for transformation parameters. The successful transformation is the one with the most votes. The second strategy, alignment, hypothesises sufficient image-model correspondences to compute the transformation, perhaps by exhaustive combination of a sufficient set of features. For both strategies, the resulting transformation is used to map the rest of the model onto the image, and support for the model hypothesis is measured against additional image features.

As described in section 3.2, six point correspondences are required to determine the fundamental matrix, with symmetric part rank 2, that models the image relations. Consequently, employing the first strategy of voting on a matrix has a high complexity of $O(n^6)$, where n is the number of suitable points.

A strategy with potentially far smaller cost is to follow the alignment approach: groupings are formed of six features which are likely to be symmetry related points, and a putative F computed. The support is then measured by counting how many other features obey this epipolar geometry, i.e. if x and x' support the matrix, then the distance between the point x' and the epipolar line Fx must be below a threshold. It is important that the minimum number of correspondences are used to estimate the fundamental matrix, as this reduces the chance that the estimate might be contaminated by outliers (mis-matches).

5 Conclusions

We have demonstrated that discrete rotational symmetry provides strong image constraints for feature grouping. The rotational symmetries also allow the recovery of 3D structure up to a particular ambiguity from a single perspective image. The similarity invariants of the objects cross-section will provide an effective indexing function for object classification. These results extend the bilateral symmetry example of [9] to the rotational case, and provide a framework for an implementation of feature grouping algorithms for recognition.

Acknowledgements

We are grateful for discussions with Paul Beardsley and Richard Hartley. Ellen Walker of RPI provided support and guidance for Jane Liu. Financial support was provided by ESPRIT Project 6448 'VIVA' and General Electric.

References

- Beardsley P., Zisserman A., "Affine Calibration of Mobile Vehicles" Proc. of Workshop on Geometrical Modeling and Invariants for Computer Vision, Xidian University Press, 1995.
- [2] Faugeras, O., Luong, Q.T., and Maybank, S.J., "Camera Self-Calibration: Theory and Experiments," *Proc. of ECCV*, 321-334, 1992.
- [3] Faugeras, O.D., "What can be Seen in Three Dimensions with an Uncalibrated Stereo Rig?" *Proc. ECCV*, LNCS 588, Springer-Verlag, p.563-578, 1992.

- [4] Hartley, R.I., Gupta, R. and Chang, T. "Stereo from Uncalibrated Cameras," *Proc. CVPR92*, p.761-764, 1992.
- [5] Heller A. and Mundy, J.L., "Benchmark Evaluation of a Model-Based Object Recognition System," Proc. 3rd International Conference on Computer Vision, p. 268, 1990. Reprinted in Computer Vision: Advances and Applications, Kasturi, R. and Jain, R. editors, IEEE Computer Society Press, 1991.
- [6] Liu J.S., Mundy J.L., Forsyth D.A., Zisserman A. and Rothwell C.A., "Efficient Recognition of Rotationally Symmetric Surfaces and Straight Homogeneous Generalized Cylinders", *CVPR*, 1993.
- [7] Maybank, S.J., *Theory of Reconstruction from Image Motion*. Springer Series in Information Science vol 28. Springer-Verlag: Berlin, 1993.
- [8] Mundy J.L. and Zisserman A.P., *Geometric Invariance in Computer Vision*. MIT Press, 1992.
- [9] Rothwell, C.A., Forsyth, D.A., Zisserman, A. and Mundy, J.L. "Extracting Projective Structure from Single Perspective Views of 3D Point Sets", *ICCV*, 1993.
- [10] Semple, J.G. & Kneebone, G.T., *Algebraic Projective Geometry*. Oxford: Clarendon Press, 1953 (reprinted 1979).
- [11] Van Gool L., Moons T. and Proesmans M. "Groups, fixed sets, symmetries, and invariants, Part II: fixed discrete sets of points", Techn. Report KUL/ESAT/MI2/9504, 1995.
- [12] Zisserman A., Mundy J.L., Forsyth D.A., Liu J., Pillow N., Rothwell C.A. and Utcke S., "Class-Based Grouping in Perspective Images", *Proc. ICCV*, 1995.