A New Structure of Invariant for 3D Point Sets from A Single View

Y. Zhu, L. D. Seneviratne and S. W. E. Earles
Department of Mechanical Engineering , King's College London,
Strand, London WC2R 2LS, United Kingdom

Abstract

The invariant used as an index has shown many advantages over the pose dependent methods in model-based object recognition. Although perspective and even weak perspective invariants do not exist for general three dimensional point sets from a single view, invariants do exist for structured three dimensional point sets. However, such invariants are not easy to derive. A new special structure for calculating invariants of three dimensional objects is presented. The 3D invariant structure proposed by Rothwell requires seven points that lie on the vertices of a six-sided polyhedral and is applicable to position free objects. In comparison, the proposed algorithm requires only six points on adjacent (virtual) planes that provides two sets of four coplanar points and does not require the position free condition. Hence it is applicable to a wider class of objects. The algorithm is demonstrated on images of real scenes.

1. Introduction

The invariant used as an index has shown many advantages over pose dependent methods in model-based object recognition. It is important to investigate simple and easy-to-get three dimensional special structures for computing projective invariants since invariants do not exist for unconstrained 3D point sets.

Much work has been done on planar objects [1-4] using plane collections of point sets or conics. Lamdan *et al.*[2], assuming the affine approximation to the perspective transformation, used three non-collinear points as the basis. Any affine transformation applied to the set points does not change the coordinates based on the same ordered base triplet, or on a system of the distinguish frame. Forsyth *et al.*[4] focused their research on a plane project group to recognise curved plane objects in three dimensional space. Models are generated directly from image data. Their main contribution is that curves are handled as geometric features with an identity of their own and are not approximated by line segments. An invariant fitting theorem which works for algebraic curves of any degree was introduced. Image curves are represented by invariant shape descriptors, which allow direct indexing into a model library. Pose recovery, using algebraic curves themselves instead of point sets via equiform invariants, was presented and the position and orientation of the object plane with respect to the camera is obtained, which is a typical problem of camera calibration.

More attention has been paid to 3D object recognition [5-10]. Previous 3D object recognition systems [15] rely on the pose of objects for recognition. Shapes measured in images depend not just on the shape of the object observed but on their pose and the intrinsic parameters of the camera. Some 3D invariants have been constructed when weak perspective approximation is assumed [8, 9]. This assumption is valid only when the relative distances between points in the object are much smaller than their distances to the camera.

Rothwell *et al.* [6, 7] argued that invariants do exist for structured three dimensional point sets, although not for three dimensional point sets in a general position from a single view, as proved in a number of papers [12, 13]. Projective invariants are derived for two classes of objects. The first is for points that lie on the vertices of polyhedron and invariants are computed by using an algebraic framework of constraints between points and planes. The second is for objects that are bilateral symmetric. For the first class of objects, a minimum of seven points, that lie on the vertices of a six-sided polyhedron, are required in order to recover the structure of a projectivity. The object should be position free [6, 14]. For the second class of objects, a minimum of eight points, or four points and two lines that are bilateral symmetric, are needed.

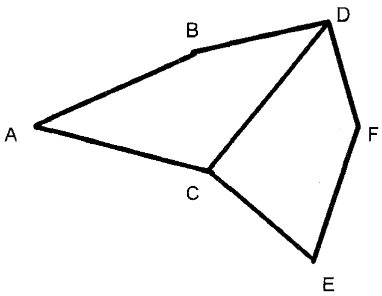
A new algorithm to compute an invariant is presented, based on a structure of six points on adjacent planes which provide two sets of four coplanar points. Extension can be easily made to a structure of seven points on three mutual adjacent planes which provides three sets of four coplanar points. The essential requirement for the method is six points instead of seven. No position free condition is needed. The condition to identify a 3D object is less restrictive, compared to Rothwell's method, and can be used more widely. The algorithm is demonstrated on images of real scenes.

IEEE International Conference
on Robotics and Automation
0-7803-1965-6/95 $4.00 ©1995 IEEE

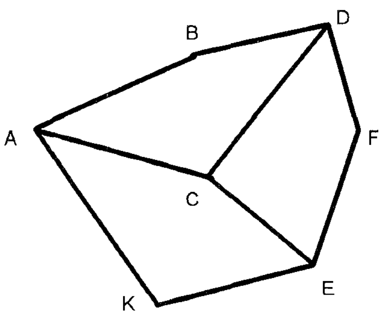
2. Invariant from A Single Perspective View

An invariant is defined in the context of a particular group of transformations for a set of geometric entities under a certain structure. To derive invariants, Forsyth *et al.* [4] approached the problem from algebraic theory, while Barrett *et al.* [10] used the techniques of determinant ratio and homogeneous equations and provided a good account of the measurement of given geometric entities.

In this section a 3D projective invariant from a single view, based on a structure with six points on adjacent (virtual) planes that provides two sets of four coplanar points, is presented. The following theorem is proved using determinant ratio techniques.



a. A structure of 6 points which provides 2 sets of 4 coplanar points



b. A structure of 7 points which provides 3 sets of 4 coplanar points

FIG. 1

Theorem. Let A, B, C, D, E, and F be six points on adjacent planes of a object providing two sets of four coplanar points (Fig. 1a). No three of the four coplanar points are collinear. Let A', B', C', D', E', and F' be their image points on the image plane. Let P_{ABC} represent the area of the triangle composed of the three vertices A, B, and C. The cross-ratio of the areas of the corresponding triangles is a projective invariant.

$$I = \frac{P_{ABD}}{P_{ABC}} \frac{P_{FEC}}{P_{FED}} = \frac{P_{A'B'D'}}{P_{A'B'C'}} \frac{P_{F'E'C'}}{P_{F'E'D'}}$$

Proof. Consider a world coordinate system *O-xyz*. Let the object points A (x_A, y_A, z_A), B (x_B, y_B, z_B) and C (x_C, y_C, z_C) lie on the plane p_1 (Fig 2). Then for the point A,

$$\frac{a}{d}x_A + \frac{b}{d}y_A + \frac{c}{d}z_A + 1 = 0,$$

where a, b, c, and d are constants.

Let the corresponding image points A' (u_A, v_A, w_A), B' (u_B, v_B, w_B), and C' (u_C, v_C, w_C) lie on the image plane p_2. Then for the point A',

$$\frac{e}{h}u_A + \frac{f}{h}v_A + \frac{g}{h}w_A + 1 = 0,$$

where e, f, g, and h are constants.

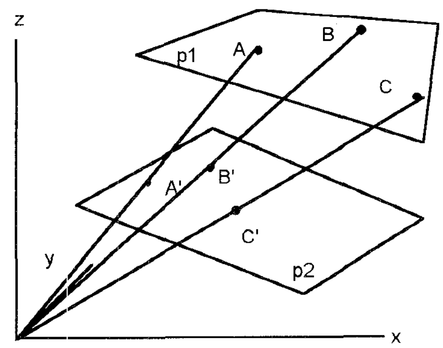


FIG. 2 Configuration of coplanar point projection.

For perspective projection,

$$\frac{x_A}{z_A} = \frac{u_A}{w_A}; \qquad \frac{y_A}{z_A} = \frac{v_A}{w_A};$$

Now, the following matrix-vector equation can be written similar to Barrett [10].

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & a/d & b/d & c/d \end{pmatrix} \frac{1}{z_A} \begin{pmatrix} 1 \\ x_A \\ y_A \\ z_A \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & e/h & f/h & g/h \end{pmatrix} \frac{1}{w_A} \begin{pmatrix} 1 \\ u_A \\ v_A \\ w_A \end{pmatrix}$$

The equation can be expanded for the three points A, B, and C, as follows:

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & a/d & b/d & c/d \end{pmatrix} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \frac{1}{z_A} \begin{pmatrix} 1 \\ x_A \\ y_A \\ z_A \end{pmatrix} \frac{1}{z_B} \begin{pmatrix} 1 \\ x_A \\ y_B \\ z_B \end{pmatrix} \frac{1}{z_C} \begin{pmatrix} 1 \\ x_C \\ y_C \\ z_C \end{pmatrix} \right) =$$

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & e/h & f/h & g/h \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{w_A}\begin{pmatrix} 1 \\ u_A \\ v_A \\ w_A \end{pmatrix} & \frac{1}{w_B}\begin{pmatrix} 1 \\ u_B \\ v_B \\ w_B \end{pmatrix} & \frac{1}{w_C}\begin{pmatrix} 1 \\ u_C \\ v_C \\ w_C \end{pmatrix} \\ 0 & & & \\ 0 & & & \\ 0 & & & \end{pmatrix}$$

This matrix equation is of the form

$$M \cdot Q_{ABC} = N \cdot Q_{A'B'C'} ,$$

where the matrices M and N contain the constants of the orientation of the object and image planes, and Q_{ABC} and $Q_{A'B'C'}$ contain the coordinates of object points (A, B, C) and their correspondent image points (A', B', C').

Using the determinant ratio technique, both the orientation matrices and the inverse of z coordinates of object and image points are eliminated, as outlined below:

$$\frac{\|M_1\| \cdot \|Q_{ABD}\| \frac{1}{z_A}\frac{1}{z_B}\frac{1}{z_D} \|M_2\| \cdot \|Q_{FEC}\| \frac{1}{z_F}\frac{1}{z_E}\frac{1}{z_C}}{\|M_1\| \cdot \|Q_{ABC}\| \frac{1}{z_A}\frac{1}{z_B}\frac{1}{z_C} \|M_2\| \cdot \|Q_{FED}\| \frac{1}{z_F}\frac{1}{z_E}\frac{1}{z_D}} = \frac{\|N\| \cdot \|Q_{A'B'D'}\| \frac{1}{w_A}\frac{1}{w_B}\frac{1}{w_D} \|N\| \cdot \|Q_{F'E'C'}\| \frac{1}{w_F}\frac{1}{w_E}\frac{1}{w_C}}{\|N\| \cdot \|Q_{A'B'C'}\| \frac{1}{w_A}\frac{1}{w_B}\frac{1}{w_C} \|N\| \cdot \|Q_{F'E'D'}\| \frac{1}{w_F}\frac{1}{w_E}\frac{1}{w_D}}$$

where

$$Q_{CBA} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & x_C & x_B & x_A \\ 0 & y_C & y_B & y_A \\ 0 & z_C & z_B & z_A \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & x_C & x_B & x_A \\ 0 & y_C & y_B & y_A \\ 0 & z_C & z_B & z_A \end{vmatrix} = \begin{vmatrix} x_C & x_B & x_A \\ y_C & y_B & y_A \\ z_C & z_B & z_A \end{vmatrix} .$$

Therefore

$$\frac{\|Q_{ABD}\| \cdot \|Q_{FEC}\|}{\|Q_{ABC}\| \cdot \|Q_{FED}\|} = \frac{\|Q_{A'B'D'}\| \cdot \|Q_{F'E'C'}\|}{\|Q_{A'B'C'}\| \cdot \|Q_{F'E'D'}\|}$$

The determinant of matrix Q_{CBA} is six times the volume of the tetrahedron formed by points C, B, A, and the origin of world coordinate system *O-xyz;* or six times the product of the triangular area P_{CBA} and the perpendicular distance from the plane p_1 to the origin of *O-xyz*. Further,

$$\frac{d_1 \cdot P_{ABD}}{d_1 \cdot P_{ABC}} \frac{d_2 \cdot P_{FEC}}{d_2 \cdot P_{FED}} = \frac{d \cdot P_{A'B'D'}}{d \cdot P_{A'B'C'}} \frac{d \cdot P_{F'E'C'}}{d \cdot P_{F'E'D'}}$$

and thus the invariant is given by:

$$I = \frac{P_{ABD}}{P_{ABC}} \frac{P_{FEC}}{P_{FED}} = \frac{P_{A'B'D'}}{P_{A'B'C'}} \frac{P_{F'E'C'}}{P_{F'E'D'}} \qquad (1)$$

The area of a triangle is a scalar term independent of the coordinate systems. Thus the theorem is proved.

The Q terms in the above equations depend on the unknown world coordinate system. What is known are the coordinates of the object points in the object coordinate system and those of image points in the image coordinate system. It is therefore necessary for the invariant to be measured independent of the world coordinate system. For the object, the calculation is carried out on three different planes. Using the determinant ratio technique and other elimination techniques, unknown information is taken out from the formula of the invariant: (a). the orientation of the object planes and the image plane, (b). the inverse of the z coordinates of object points and image points, and (c). the distance from the origin of the world coordinate system to three object planes and the image plane.

The following colliery can also be proved similar to the proof of theorem 1.

Colliery. Let A, B, C, D, E, F, and K be seven points on three mutually adjacent planes which provide three sets of four coplanar points (Fig 1b). Let point C be trihedral and not collinear with any other two points. Then the cross-ratio of areas of corresponding triangles is a projective invariant.

$$I_1 = \frac{P_{CKA}}{P_{CBK}} \frac{P_{CBD}}{P_{CAB}} \frac{P_{CBF}}{P_{CFD}} = \frac{P_{C'K'A'}}{P_{C'B'K'}} \frac{P_{C'B'D'}}{P_{C'A'B'}} \frac{P_{C'B'F'}}{P_{C'F'D'}} \qquad (2)$$

3. Application of Invariant for 3D Object Recognition

The method is first checked using the tutorial example in [6] (Fig. 3). Figure 3 shows two different perspective views of the same polyhedral object, and the image points used to compute the invariants I, as outlined in Eq. (1). The results are given in Table 1. The invariants of the model object and image 1 are identical. For Fig. 3c the coordinates of the image points are not given in [6] and so were measured to 1mm accuracy. The invariant is slightly different. If the coordinates are modified by projective transformation, the same invariant is found (Table 1: image 2*).

Table 1

	model	image 1	image 2	image 2*
I	1.2607	1.2607	1.2855	1.2607

* the coordinates of image are modified by projective transformation.

In Fig. 4, six images of a bearing support are shown, with six correspondent points marked on each image (1, 2, 3, 4, 5, 6). The images are processed by fitting straight lines to edge data. Vertex positions are found by intersecting pairs of lines (by hand). The correspondence between object points and image points is assumed. The calculated invariants for the six images are given in Table 2. They are fairly constant with change of the viewpoint.

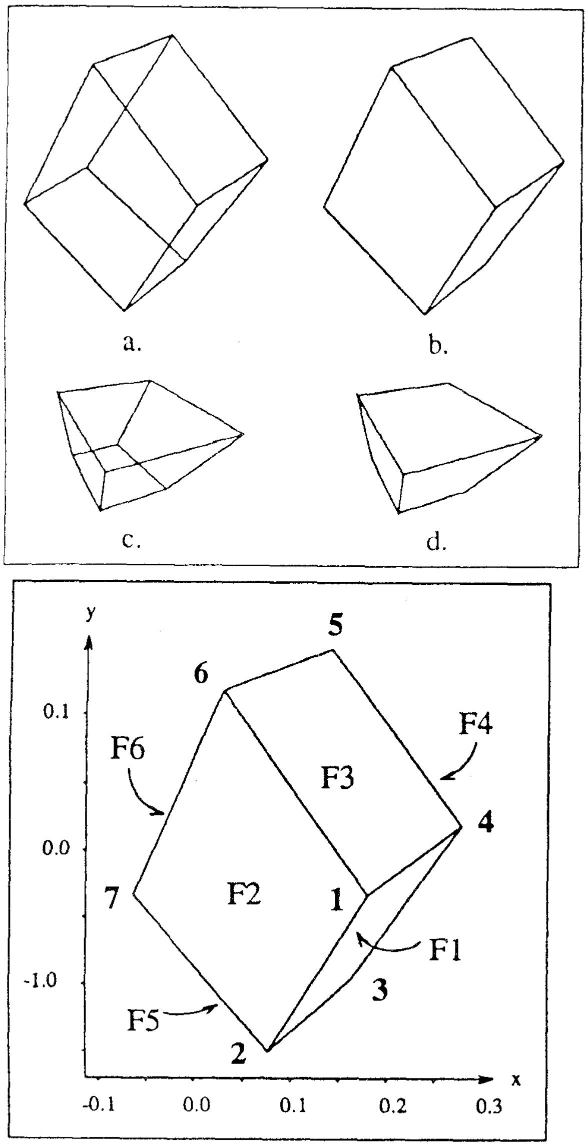


FIG. 3 Figure (a) and (b) show one perspective view of the polyhedron with and without hidden lines. Figure (c) and (d) show a second view. The points of (b) used to compute the invariants are given in the lower figure.(Adapted from Rothwell *et. al.* [6])

It can be proved that if three lines on two adjacent planes formed by the six points intersect at one point (three parallel lines can be considered to intersect at infinity), then the invariants are equal to unity. For the purpose of discrimination, this type of structure should be avoided. If point 1 on the top surface in Fig. 4a is replaced by point 7 on the base surface to form a new set of six points, the new invariants are listed in Table 3.

Table 2

	I
model	1.0000
image 1	1.0136
image 2	0.9529
image 3	1.0035
image 4	0.9624
image 5	0.9546
image 6	0.9614

Table 3

	I'
model	0.3482
image 1	0.3158
image 2	0.3423
image 3	0.3544
image 4	0.3486
image 5	0.3564
image 6	0.3135

Table 4

	model	image 1	image 2	image 2*
I_1	0.9745	0.9745	1.0597	0.9745

* the coordinates of image are modified by projective transformation.

Compared to Rothwell's structure, if seven points that lie on the vertices of a six-sided polyhedral are available, then there are three cases of six points on adjacent planes of a object providing two sets of four coplanar points. Therefore three projective invariants can be derived using the presented method. The Colliery of the last section provides another invariant for this structure. The invariants of the tutorial example are listed in Table 4 using Eq. (2).

Once the object is positively identified by matching invariants, a three dimensional projective transformation can be obtained by the method described in [11]. The transformation has 11 essential parameters since the overall scale of the matrix does not matter in homogeneous coordinates. Six points in the image plane, having 12 degrees of freedom, can provide one invariant subject to 11 essential parameters for projective transformation. Therefore six pairs of corresponding points between model and scene are enough for determining the projective transformation. If there are more than six points, the system is over constrained, but the pose of the object can be calculated using the least square method to reduce any point position inaccuracies. If the coordinates of the principal point of the camera in the image plane are known, the positions and orientation of the object observed can be recovered with respect to the camera coordinate system, which is important in robot applications. Interestingly, four coplanar points can provide a unique solution for determining the position and orientation of the object with respect to the camera coordinate system, if the coordinates of the principal point of camera in the image plane and the focal length of camera are known [16].

The transformed model, projected into the image to show registration between them and the correspondence of the

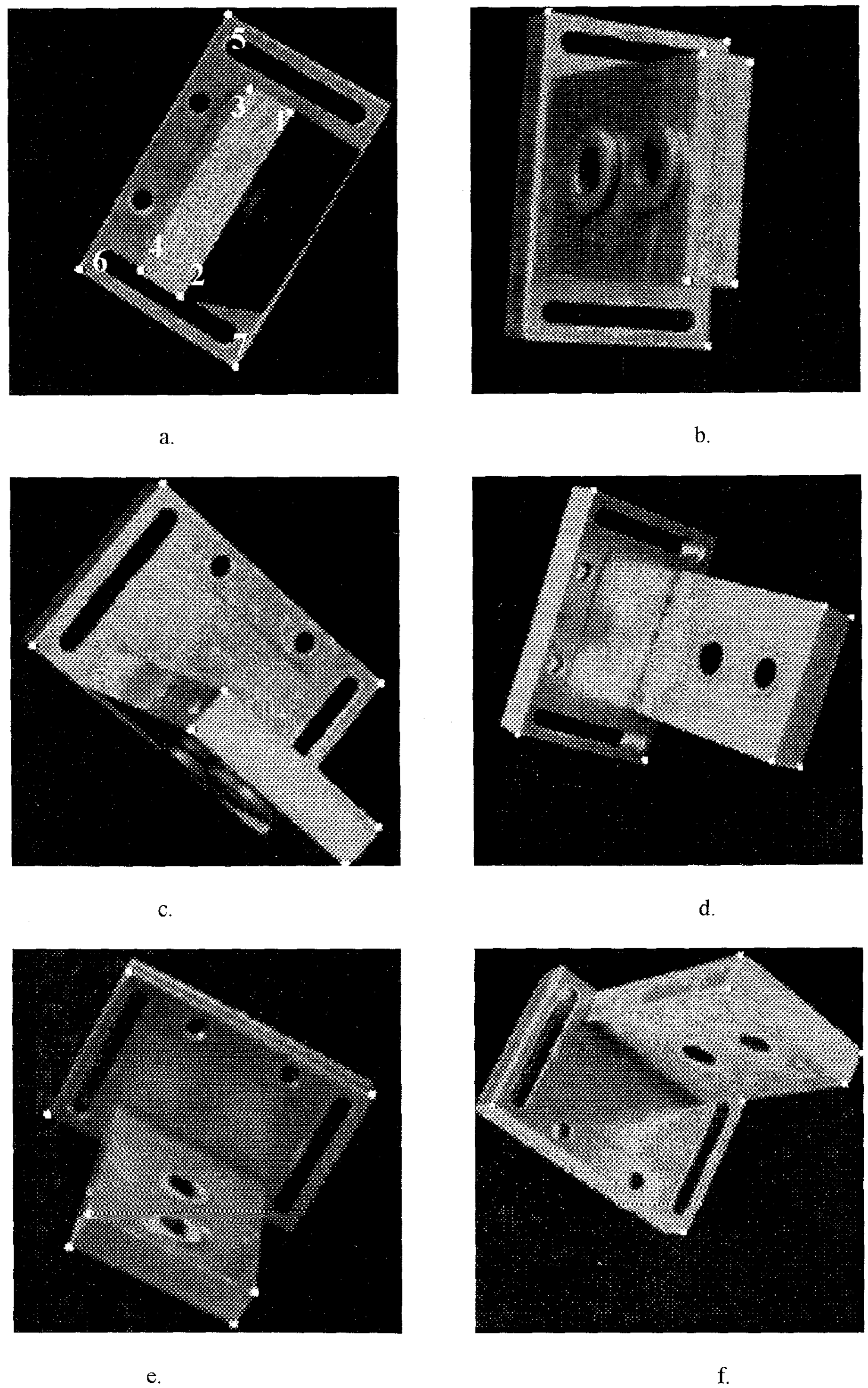


FIG 4. Six images of a bearing support from different points of view and the points used to compute the invariants

model to the image features excluding reference points, can be used as verification (Fig. 5).

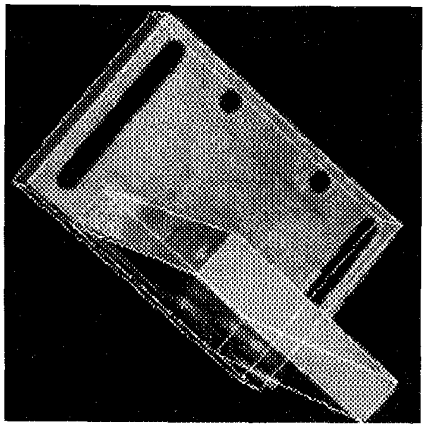


FIG. 5 Registration of 3D models onto the images

4. Conclusion

The presented algorithm is more general compared to Rothwell's structure. Even if seven points are available, which provide three sets of four coplanar points and three invariants, the requirements are still less restrictive compared to Rothwell's. With reference to Fig. 1b, only point C is necessary to be trihedral. It is not a necessary condition for points A, B, D, E, F and K. Also no position free condition is needed. The main idea is that a 3D object can be considered as the composition of a set of surfaces, both curved (nonplanar) and planar. Objects which are composed of planar surfaces or point sets on the virtual planes can be recognised by means of an invariant, which is a collection of areas of planar triangles. The invariant is derived by determinant ratio and other elimination techniques.

In nature, an invariant based on a group of features is a local description of an object. Such invariants can only be used in cases where features are completely available. Requiring 6 points instead of 7, there are more chances of obtaining the object features necessary to compute invariants, thus the algorithm presented can be used more widely than Rothwell's. Furthermore, the proposed structure has already found applications in solving the fundamental problems of computer vision under the condition of projective geometry [6].

Further effort is needed to investigate how errors introduced by the sensor and the feature extraction scheme affect the invariant.

Acknowledgements

Y. Zhu is supported by an ORS Award and a K. C. Wong Scholarship.

References

1. I. Weiss, "Geometric invariants and object recognition," *International Journal of Computer Vision,* Vol. 10, No. 3 , pp. 207-231, 1993.
2. Y. Lamdan, J. T. Schwartz and H. J. Wolfson, "Object recognition by affine invariant matching," *Proc. CVPR88,* pp. 335-344, 1988.
3. Y. Lamdan, H. J. Wolfson, "Geometric hashing: A general and efficient model-based recognition scheme," *Proc. Int. Conf. Comput. Vision,* pp. 238-249, 1988.
4. D. A. Forsyth, J. L. Mundy, A. P. Zisserman , C. Coelho, A. Heller and C. A. Rothwell, "Invariant descriptors for 3-D object recognition and pose," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. **PAMI-13**, No. 10, pp. 971-991, Oct. 1991.
5. R. Mohr , L. Morin and E. Grosso, " Relative positioning with uncalibrated cameras," *Geometrical Invariance in Computer Vision,* J. L. Mundy and A. Zisserman editors, MIT Press, 1992.
6. C.A. Rothwell , D. A. Forsyth, A. Zisserman and J. L. Mundy, " Extracting projective information from single views of 3D point sets," *TR OUEL 1973/93,* Department of Engineering Science, Oxford University, Oxford, 1993.
7. C.A. Rothwell , D. A. Forsyth, A. Zisserman and J. L. Mundy, "Extracting projective structure from single perspective views of 3D point sets," *Proc. 4th Int. Conf. Comput. Vision,* pp. 573-582, Berlin, Germany, 1993. (the Best Paper Award received)
8. D. Weinshall, "Model-based invariants for 3-D vision," *International Journal of Computer Vision,* Vol. 10, No. 1 , pp. 27-42, 1993.
9. P. C. Wayner, "Efficiently using invariant theory for model-base matching," *Proc. Int. Conf. Comput. Vision Pattern Recog.*, pp. 473-478, 1991.
10. E. B. Barrett, P. M. Payton, N. N. Haag ang M. H. Brill, "General methods for determining projective invariancts in imagery," *CVGIP: Image Understanding,* Vol. 53, No. 1, pp. 46-65, Jan. 1991.
11. J. L. Mundy and A. Zisserman, "Projective geometry for machine vision," *Geometrical Invariance in Computer Vision*, J. L. Mundy and A. Zisserman editors, MIT Press, 1992.
12. J. B. Burns, R. S. Weiss and E. M. Riseman, "View variation of point-set and line-segment feature," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. **PAMI-15**, No. 1, pp. 51- 68, Jan. 1993.
13. D. T. Clemens and D. W. Jacobs, "Space and time bounds on indexing 3-D models from 2-D images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. **PAMI-13**, No. 10, pp. 1007- 1017, Oct. 1991.
14. K. Sugihara, *Machine Interpretation of Line Drawings,* MIT Press, 1986.
15. D. G. Lowe, "The Viewpoint consistency constraint," *International Journal of Computer Vision,* Vol. 1, No. 1 , pp. 57-72, 1987.
16. R. Horaud, B. Conio, and O. Leboulleux, "An analytic solution for the perspective 4-point problem," *CVGIP*, Vol. 44, pp. 33-44, 1989.