MVA'94 IAPR Workshop on Machine Vision Applications Dec. 13-15, 1994, Kawasaki

OBJECT RECOGNITION USING CONIC-BASED INVARIANTS FROM MULTIPLE VIEWS

Osamu Takae Yoshinori Kuno Jun Miura Yoshiaki Shirai
Department of Mechanical Engineering for Computer-Controlled Machinery
Osaka University

2-1, Yamadaoka, Suita, Osaka 565, Japan

ABSTRACT

Geometric invariants from multiple views provide useful information for 3D object recognition. This paper presents two methods of calculating conic-based invariants to improve efficiency in recognition. When we can find a corresponding pair of conics that are derived from a circle in 3D, we can get two invariants for a point. If we obtain a corresponding point pair in addition to the conic pair, we can get three invariants for another point. We present an object recognition method using these invariants. Experimental results confirm the usefulness of the method.

1 INTRODUCTION

One of the critical issues in 3D object recognition is how to treat any changes of object appearances depending on view points. If we can obtain any invariants under imaging transformation, they can help to solve the issue. We can recognize an object by calculating its invariants and searching the model database for the object with these values. Unfortunately, it is proved that such invariants cannot be obtained for general 3D shapes from a single view image[1]. Thus, research interests on multiple view invariants have been increasing[2][3].

If we have two images in which we can obtain five corresponding pairs of points, we can obtain three 3D affine invariants. These three invariants can be interpreted as the 3D position of a fifth point in the basis derived from the other four points. Vinther and Cipolla have proposed a 3D object recognition system using such invariants[4]. Their system uses corners of objects for feature points. In this case, we cannot tell which points should be used to establish the basis vectors. Thus, we have to consider all combinations of basis vectors in model database preparation, and search for the proper basis in recognition. This makes object recognition inefficient although they use geometric hashing technique to reduce the combinatorial problem.

To solve the above problem, we propose the use of more complex features for basis setting. The number of such features in the scene is small. Thus, the number of combinations for basis setting is small as well as that for correspondence determination in a pair of images. In this paper, we use conics[5] as complex features because man-made objects have often circles and they are projected as conics in images.

We present two methods of calculating conic-based invariants. If we can find a corresponding pair of conics that are derived from a circle in 3D and obtain a corresponding point pair, we can get two invariants indicating the 3D position of the point. If we use a corresponding point pair in addition to the conic pair to establish the basis, we can get three invariants for another point.

These conic-based invariants have the following major advantages. First, correspondences between features in two images can be determined easily. Second, the model base used in the recognition process can be small. The reasons are that these invariants can be calculated from one conic and one point or one conic and two points instead of five points and that the number of conics found in images is small.

This paper presents the two methods of calculating conic-based invariants and an object recognition system using these invariants.

2 CONIC-BASED INVARIANTS

2.1 AFFINE INVARIANTS FROM MULTIPLE VIEWS

We adopts weak perspective projection as our camera model. It is a good approximation of the actual camera transformation when the distance from an object to the camera is much greater than the extent of the object. The linearity of weak perspective projection allows us to recover invariant 3D object structure from two or more images. We briefly show affine invariants from multiple views before we present conic-based invariants.

Given five 3D points $\mathbf{X}_i, i\epsilon 1,2,3$, it is possible to calculate three invariants. This can be shown by considering a basis $\mathbf{E}_i = \mathbf{X}_i - \mathbf{X}_0, i\epsilon 1,2,3$ in 3D space as in Fig.1. Any 5th point can be expressed as a linear combination of the basis vectors.

$$\mathbf{X}_4 = \mathbf{X}_0 + \alpha \mathbf{E}_1 + \beta \mathbf{E}_2 + \gamma \mathbf{E}_3 \tag{1}$$

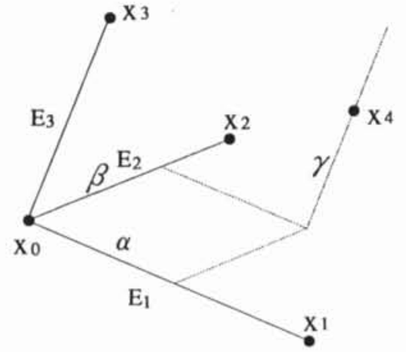


Figure 1: Affine Basis

The (α, β, γ) can be viewed as the invariant coordinates of the point $\mathbf{X}_4$ in this basis. They are related to the 3D shape of the object under 3D affine transformations. Due to the linear nature of weak perspective projection, (α, β, γ) will remain viewpoint invariant under this transformation. Thus, we can obtain the following equation for the projection points in an image.

$$\begin{bmatrix} u_4 - u_0 \\ v_4 - v_0 \end{bmatrix} = \begin{bmatrix} e_{1u} & e_{2u} & e_{3u} \\ e_{1v} & e_{2v} & e_{3v} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \tag{2}$$

where $\mathbf{x}_0 = (u_0, v_0)^T$,$\mathbf{x}_4 = (u_4, v_4)^T$,$\mathbf{e}_i = (e_{iu}, e_{iv})^T$ are the projected origin $\mathbf{X}_0$, point $\mathbf{X}_4$, and basis $\mathbf{E}_i$, respectively. This provides two equations with three unknowns. The problem has become underdetermined, so a single 2D image does not allow the recovery of the invariants. A second view with known point correspondences to the first view will however give an overdetermined set of equations solvable by standard least square minimization methods.

2.2 INVARIANTS BASED ON A CONIC

If we can find a corresponding pair of conics that are derived from a circle in 3D, we can get two invariants for a point.

Given a circle with the center $\mathbf{X}_0$ in 3D space, we consider the 3D points $\mathbf{X}_1$ and $\mathbf{X}_2$ on the circle that are projected to a point on the major axis of the conic and that on the minor axis, respectively, as shown in Fig.2. Then, we take the point $\mathbf{X}_3$ in such a way that the line segment $\overline{\mathbf{X}_0\mathbf{X}_3}$ is perpendicular to the circle plane and its length is the circle radius. A point $\mathbf{X}_4$ can be expressed as a linear combination of the basis vectors $\mathbf{E}_i = \mathbf{X}_i - \mathbf{X}_0$. The relation is represented by the following equation in the cylindrical coordinate system based on the circle.

$$\mathbf{X}_4 = \mathbf{X}_0 + r\cos\theta \mathbf{E}_1 + r\sin\theta \mathbf{E}_2 + z\mathbf{E}_3 \tag{3}$$

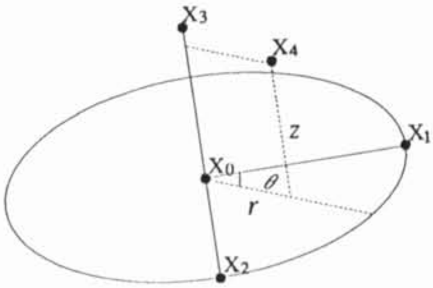


Figure 2: Conic Basis

The (r, θ, z) are the coordinates of the point $\mathbf{X}_4$ in this basis. Under weak perspective projection, we can get the following relation in a projected image.

$$\begin{bmatrix} u_4 - u_0 \\ v_4 - v_0 \end{bmatrix} = \begin{bmatrix} e_{1u} & e_{2u} & e_{3u} \\ e_{1v} & e_{2v} & e_{3v} \end{bmatrix} \begin{bmatrix} r\cos\theta \\ r\sin\theta \\ z \end{bmatrix} \tag{4}$$

where $\mathbf{u}_0 = (u_0, v_0)^T, \mathbf{u}_4 = (u_4, v_4)^T$ are the projected points $\mathbf{X}_0, \mathbf{X}_4$, and $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the projections of the vecotrs $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$.

If we have a second view, we can calculate the invariants. In this case, however, we can not determine the projected points of $\mathbf{X}_1$ and $\mathbf{X}_2$ in the second image. Thus, we introduce $\mathbf{X}'_1$ and $\mathbf{X}'_2$ instead of $\mathbf{X}_1$ and $\mathbf{X}_2$. They are determined in the second image as in the same way as for $\mathbf{X}_1$ and $\mathbf{X}_2$. Then the point $\mathbf{X}_4$ can be given by

$$\mathbf{X}_4 = \mathbf{X}_0 + r\cos\theta' \mathbf{E}'_1 + r\sin\theta' \mathbf{E}'_2 + z\mathbf{E}_3 \tag{5}$$

where $\mathbf{E}'_1 = \mathbf{X}'_1 - \mathbf{X}_0, \mathbf{E}'_2 = \mathbf{X}'_2 - \mathbf{X}_0$.

In the second image, the following relation holds.

$$\begin{bmatrix} u'_4 - u'_0 \\ v'_4 - v'_0 \end{bmatrix} = \begin{bmatrix} e'_{1u} & e'_{2u} & e'_{3u} \\ e'_{1v} & e'_{2v} & e'_{3v} \end{bmatrix} \begin{bmatrix} r\cos\theta' \\ r\sin\theta' \\ z \end{bmatrix} \tag{6}$$

The equations (4) and (6) provide us with four equations with four unknowns. Thus We can recover r and z as invariants.

2.3 INVARIANTS BASED ON A CONIC AND A POINT

If we can obtain a corresponding point pair in addition to a conic pair, we can get three invariants for another point.

Suppose we have a 3D circle with the center $\mathbf{X}_0$ and a 3D point $\mathbf{X}_4$ as shown in Fig.3. In addition to the points considered in Section 2.2, we use the following points. Let $\mathbf{X}_5$ be the intersection point between the circle plane and the line passing through the point $\mathbf{X}_4$ and perpendicular to the plane. We rotate the point $\mathbf{X}_5$ by 90 degrees around the point $\mathbf{X}_0$ on the circle plane to determine the point $\mathbf{X}_6$. The points $\mathbf{X}_5$,$\mathbf{X}_6$ can be expressed by

$$\begin{aligned} \mathbf{X}_5 &= \mathbf{X}_0 + r\cos\theta\mathbf{E}_1 + r\sin\theta\mathbf{E}_2 \\ &= \mathbf{X}_0 + r\cos\theta'\mathbf{E}_1' + r\sin\theta'\mathbf{E}_2' \qquad (7) \\ \mathbf{X}_6 &= \mathbf{X}_0 - r\sin\theta\mathbf{E}_1 + r\cos\theta\mathbf{E}_2 \\ &= \mathbf{X}_0 - r\sin\theta'\mathbf{E}_1' + r\cos\theta'\mathbf{E}_2' \qquad (8) \end{aligned}$$

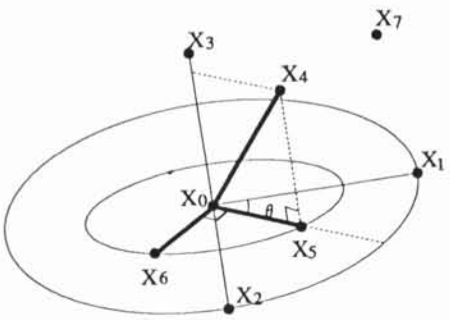


Figure 3: Conic and Point Basis

We use these points $\mathbf{X}_0, \mathbf{X}_4, \mathbf{X}_5, \mathbf{X}_6$ to set the basis. A point $\mathbf{X}_7$ can be expressed as a linear combination of $\mathbf{E}_4$,$\mathbf{E}_5$,$\mathbf{E}_6$.

$$\mathbf{X}_7 = \mathbf{X}_0 + \alpha\mathbf{E}_4 + \beta\mathbf{E}_5 + \gamma\mathbf{E}_6 \qquad (9)$$

where $\mathbf{E}_i = \mathbf{X}_i - \mathbf{X}_0, i\epsilon 4,5,6$.

We can obtain a matrix form for the first view.

$$\begin{bmatrix} u_7 - u_0 \\ v_7 - v_0 \end{bmatrix} = \begin{bmatrix} e_{4u} & e_{5u} & e_{6u} \\ e_{4v} & e_{5v} & e_{6v} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \qquad (10)$$

Since a second view gives the same kind of equation, we can calculate three invariants (α, β, γ).

3 OBJECT MODEL ACQUISITION AND RECOGNITION

We have developed an object recognition system in which the above two conic-based invariant computation methods are properly used. The recognition process consists of two stages: a candidate selection stage and a verification stage. In the first stage, we use the invariants based on a conic because the process needs to establish only one corresponding conic pair. Object candidates are chosen from the model base using the invariants. In the verification stage, these candidates are verified by the invariants based on a conic and a point because they represent complete 3D relationships.

We have experimented the system using nine objects as shown in Fig.4. First, a collection of conics and corner points are extracted from two images as shown in Fig.6. Then, the correspondences between features in the two images are hypothesized and the invariants based on a conic are calculated. The candidate model with the most matching invariant is selected in the model base. If there is no matching model, a different basis will be tried. When the candidate is found, it is verified whether or not the features are found as predicted from the invariants based on a conic and a point and the model data as shown in Fig.7. Fig.8 shows the final recognition result. The system can recognize the nine objects correctly in 90 experiments. The time required is about 20 seconds for feature extraction and 1 second for recognition on a SPARCstation10.

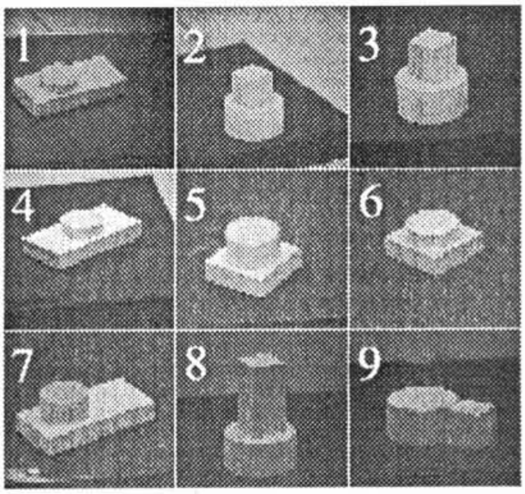
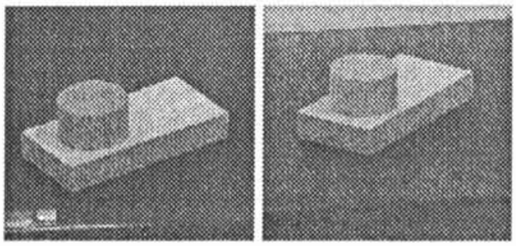
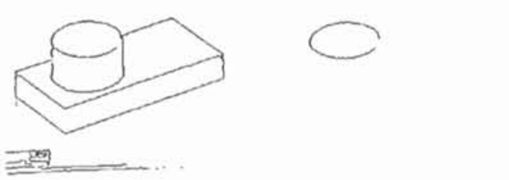


Figure 4: Test Objects

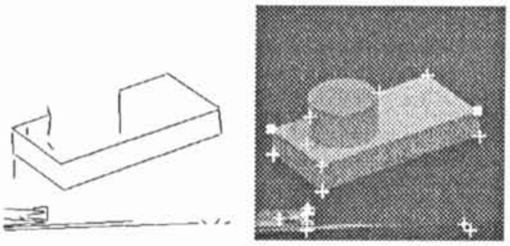


(a) First Image (b) Second Image

Figure 5: Multiple Views of the Object

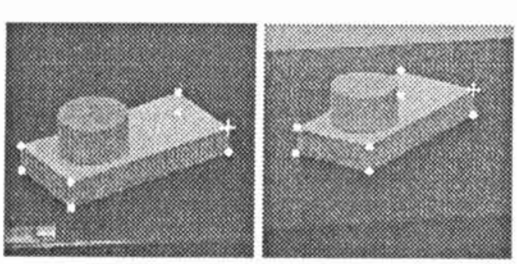


(a) Edge Image (b) Conic Extraction



(c) Edge Segment (d) Corner Extraction

Figure 6: Feature Extraction



(a) First Image (b) Second Image

Figure 7: Verification

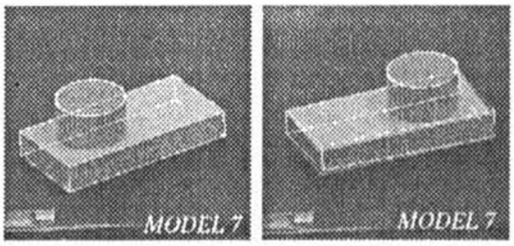


Figure 8: Recognition Result

4 CONCLUSION

We have proposed two methods of calculating conic-based invariants. When we can find a corresponding pair of conics that are derived from a circle in 3D, we can get two invariants for a point. If we obtain a corresponding point pair in addition to the conic pair, we can get three invariants for another point. We have developed an object recognition system using these invariants. Experimental results confirm the usefulness of the method.

REFERENCE

[1] E.B.Barrett, M.H.Brill, N.N.Haag and P.M.Payton. Invariant linear method in photogrammetry and model-matching. *Geometric Invariance in Computer Vision*, chapter 14, pages 277-292. MIT Press, 1992.

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

[3] J.J.Koenderink and A.J.van Doorn. Affine structure from motion. J.Opt.Soc.America, 8(2):377-385, 1991.

[4] S.Vinther and R.Cipolla. Towards 3D object model acquisition and recognition using 3D affine invariants. Technical Report CUED/F-INFENG /TR136, University of Cambridge, 1993.

[5] K.Kanatani, *Geometric Computation for Machine Vision*, chapter 8, pages 228-277. Oxford Science Publications, 1993.