Model-Based Object Recognition Using Geometric Invariants of Points and Lines[1]

Bong Seop Song

School of Electrical Engineering, Seoul National University

Kyoung Mu Lee

Department of Electronics and Electrical Engineering, Hong-Ik University

and

Sang Uk Lee

School of Electrical Engineering, Seoul National University

In this paper, we derive new geometric invariants for structured 3D points and lines from single image under projective transform, and we propose a novel model-based 3D object recognition algorithm using them. Based on the matrix representation of the transformation between space features (points and lines) and the corresponding projected image features, new geometric invariants are derived via the determinant ratio technique. First, an invariant for six points on two adjacent planes is derived, which is shown to be equivalent to Zhu's result [1], but in simpler formulation. Then, two new geometric invariants for structured lines are investigated: one for five lines on two adjacent planes and the other for six lines on four planes. By using the derived invariants, a novel 3D object recognition algorithm is developed, in which a hashing technique with thresholds and multiple invariants for a model are employed to overcome the over-invariant and false alarm problems. Simulation results on real images show that the derived invariants remain stable even in a noisy environment, and the proposed 3D object recognition algorithm is quite robust and accurate. © 2001 Elsevier Science (USA)

Key Words: geometric invariants; constrained 3D points; constrained 3D lines.

---

1. INTRODUCTION

Projective invariants are known to be an effective tool for 3D object recognition. They can provide a feasible solution for the view-variant problem in object recognition, using invariant properties common to all views of an object, irrespective of viewing direction and position of a camera [2]. Exploiting these properties, the model-based recognition can be efficiently carried out, avoiding the problematic camera calibration step. So far, a lot of work has been done to extract geometric invariants from various types of image features, including points, lines, or curves, and to apply them to recognize objects. However, most of them are concerned with planar features, which are usually inappropriate to general 3D object recognition using projected images [3–6].

Recent research reveals that invariants cannot be obtained for a set of 3D points in general configuration from a single view [7]. Thus, additional information on the geometric relation between space points is required to obtain invariants for general 3D point sets.

One approach is to derive invariants from multiple uncalibrated images of an underlying object. Barrett *et al*. proposed a method to derive an invariant for six points by using stereo images [8]. Hartley has shown that once the equipolar geometry is known, the invariant could be computed from the images of the four lines in two distinct views with uncalibrated cameras [9]. Quan [10] has also derived invariants for six space points from three images, provided that the correspondences of the points are known.

While, recently, several researchers have proposed methods to construct invariants, by approximating the perspective projection model in simplified forms, such as weak perspective, para-perspective, and affine projections, provided that the camera is far away from the object compared to the depth range of the object. Weinshall proposed a hierarchical representation of 3D shape of an object which is invariant to affine transformation using multiple frames [11]. Wayner considered a structure of four points, composing three orthogonal vectors in 3D space and constructed an invariant descriptor of single 2D image under orthographic projection and scaling [12].

Another common approach to derive invariants is to impose some restrictions on the geometric structure of space features. Rothwell *et al*. have shown that the projective invariants could be derived from a single view for two classes of structured objects [13]: points that lie at the vertices of a (virtual) polyhedron and objects that have bilateral symmetry. To derive these invariants, minimum of 7 and 8 points are required, respectively. More recently, Zhu *et al*. argued that there exists an invariant for a structure with six points on two adjacent planes, constituting of two sets of four coplanar points [1]. Sugimoto also has derived an invariant from six lines on three planes in a single view, including a set of four coplanar lines and two pairs of coplanar lines [14].

In this paper, we derive new invariants for several structured sets of space lines as well as points, and we investigate the relationships with the previously known invariants. First, a new invariant for six restricted space points, which has the same geometric configuration as Zhu's approach, is derived. Unlike Zhu's approach in which an implicit point, obtained by the intersection of lines, is involved to calculate the invariant, the proposed approach provides the same invariant without any auxiliary point, yielding a much simpler form of invariant. Then, by extending the approach to line features, and investigating the properties related to the projection of lines, we derive two new invariants for structured lines in space: the invariant for five lines on adjacent planes, and that for six structured lines on four planes.

Based on the derived geometric invariants for structured features in space, a novel 3D object recognition algorithm is proposed, in which the matching is carried out based on the geometric hashing technique [15]. For the given model of a 3D object, by analyzing the geometric relations among space features, all pertinent sets of features having the constrained structures to yield the invariants are automatically extracted. Then, invariant values are computed and recorded in a table of model invariants as the index of related features. For a given input image, candidate sets of features are selected through a geometry test to confirm the structural feasibility to the model. Then, for each candidate set, an invariant is calculated to index the table of model invariants, and it is used to vote for the corresponding feature set in the model. The final result of matching is determined by the maximum scores credited by repeated votes.

Results on extensive experiments demonstrate that the proposed invariants remain quite stable in spite of pixel noise and change of viewing direction, so that the 3D object recognition can be efficiently carried out. Moreover, it is shown that by employing a geometric hashing technique, the target object in images can be correctly located even when some features of the object are missing due to occlusion and the imperfect feature extraction process.

This paper is composed of six sections. Following the Introduction, the invariant for constrained points is discussed in Section 2, and the invariants for lines are proposed in Section 3. In Section 4, the proposed object recognition technique based on the developed invariants for structured features is described. The robustness of the proposed invariants and the performance of the proposed recognition technique are examined by several experiments for real and synthetic images in Section 5, and finally the conclusions are drawn in Section 6.

2. INVARIANT FOR SIX RESTRICTED SPACE POINTS

In this section, by using a Euclidean transformation and determinant ratio technique, a new invariant for a restricted set of points in space is derived. Then, the comparison of the proposed invariant with the invariant derived by Zhu *et al*. is also provided.

2.1. Derivation of a New Invariant

In a projective image formation process, a space point $\mathbf{X} = [X, Y, Z]^T$ is mapped to an image point $\mathbf{x} = [x, y]^T$. Let us denote $\mathbf{M} = [X, Y, Z, 1]^T$ and $\mathbf{m} = [x, y, 1]^T$ as the homogeneous representation of $\mathbf{X}$ and $\mathbf{x}$, respectively. Then, the relation between these two points can be given linearly, via a 3×4 projection matrix $\mathbf{P}$, by

$$w\mathbf{m} = \mathbf{PM}, \tag{1}$$

where w is a scaling factor.

Let us consider six space points on two adjacent planes that constitute two sets of four coplanar points as shown in Fig. 1, where the points A, B, C, and D lie on one plane Γ, while points C, D, E, and F are on the other plane Ω. Through the imaging process, these points are projected to the corresponding image points $A'-F'$, respectively, given by

$$w_i \mathbf{m}_{i'} = \mathbf{PM}_i, \quad \text{for } i = \text{A, B, C, D, E, and F}. \tag{2}$$
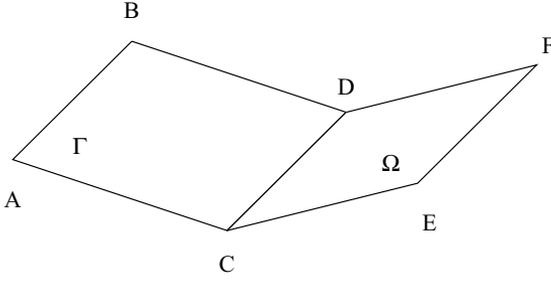
FIG. 1. Geometric configuration of six points on two adjacent planes.

Let $\mathbf{U}_\Gamma$ be a Euclidean transform of only rotation and translation, which moves the points on the plane Γ to the XY plane. Then, (2) can be rewritten as

$$w_i \mathbf{m}_{i'} = \mathbf{PU}_\Gamma^{-1} \mathbf{M}_i^\Gamma, \quad \text{for } i = A, B, C, \text{ and } D, \tag{3}$$

where $\mathbf{M}_i^\Gamma = [X_i^\Gamma, Y_i^\Gamma, 0, 1]^T$, which is the transformed coordinates of $\mathbf{M}_i$ by $\mathbf{U}_\Gamma$. Note that the third column of the matrix $\mathbf{PU}_\Gamma^{-1}$ can be ignored, since the Z coordinate of $\mathbf{M}_i^\Gamma$ is zero. Therefore, (3) can be simplified as

$$w_i \mathbf{m}_{i'} = \mathbf{T}_\Gamma \mathbf{n}_i^\Gamma, \quad \text{for } i = A, B, C, \text{ and } D, \tag{4}$$

where $\mathbf{T}_\Gamma$ is a 3×3 matrix, obtained by eliminating the third column of $\mathbf{PU}_\Gamma^{-1}$, and $\mathbf{n}_i^\Gamma = [X_i^\Gamma, Y_i^\Gamma, 1]^T$. Similarly, using the 3×3 matrix $\mathbf{T}_\Omega$, the relation between the points on the plane Ω and the image points can be expressed as

$$w_j \mathbf{m}_{j'} = \mathbf{T}_\Omega \mathbf{n}_j^\Omega, \quad \text{for } j = C, D, E, \text{ and } F.$$

Now, by augmenting the three equations for the points A, B, and C in a single matrix form, we have

$$[w_A \mathbf{m}_{A'} \quad w_B \mathbf{m}_{B'} \quad w_C \mathbf{m}_{C'}] = \mathbf{T}_\Gamma [\mathbf{n}_A^\Gamma \quad \mathbf{n}_B^\Gamma \quad \mathbf{n}_C^\Gamma]. \tag{5}$$

By taking the determinant for both sides of (5), we obtain

$$w_A w_B w_C |\mathbf{m}_{A'} \quad \mathbf{m}_{B'} \quad \mathbf{m}_{C'}| = |\mathbf{T}_\Gamma| |\mathbf{n}_A^\Gamma \quad \mathbf{n}_B^\Gamma \quad \mathbf{n}_C^\Gamma|. \tag{6}$$

Note that the determinant of a matrix composed of three coplanar points in homogeneous coordinate is equal to twice the triangular area made of those points. Moreover, the Euclidean transform preserves the area of a planar patch as well as the distance between points. Thus, if we denote the area of the triangle composed of three vertices, i, j, and k by Q_{ijk}, then (6) becomes

$$w_A w_B w_C Q_{A'B'C'} = |\mathbf{T}_\Gamma| Q_{ABC} = i_1. \tag{7}$$

Similarly, the following three equations can be obtained for the other sets of points,

$$w_A w_B w_D Q_{A'B'D'} = |\mathbf{T}_\Gamma| Q_{ABD} = i_2 \tag{8}$$

$$w_C w_E w_F Q_{C'E'F'} = |\mathbf{T}_\Omega| Q_{CEF} = i_3 \tag{9}$$

$$w_D w_E w_F Q_{D'E'F'} = |\mathbf{T}_\Omega| Q_{DEF} = i_4. \tag{10}$$

Using Eqs. (7)–(10), we can establish the following invariant I_p for the six structured space points:

$$I_p = \frac{i_1 i_4}{i_2 i_3} = \frac{Q_{A'B'C'} Q_{D'E'F'}}{Q_{A'B'D'} Q_{C'E'F'}} = \frac{Q_{ABC} Q_{DEF}}{Q_{ABD} Q_{CEF}}. \tag{11}$$

In applying I_p for object recognition, two important facts should be noticed: One is that if three or more points among the coplanar subset of the model points are collinear, I_p could be zero or undefined, which is a degenerate configuration for I_p. Another one is the "over-invariance" problem. That is, different configurations of model features can yield the same invariant, which may result in misclassification. For example, if points A and B in Fig. 1 are translated on the plane Γ, while preserving the related areas, the invariant value will not change. Thus, to increase the reliability for recognition, multiple invariants with more feature sets should be considered.

2.2. Comparison with Zhu et al.'s Result

Zhu *et al.* [1] also have derived the invariant for six restricted space points, which have the same geometric configuration considered in this paper. However, for the derivation of the invariant, they employed an implicit point G, constructed by the intersection of several lines.

As shown in Fig. 2, the intersection of the lines AD and BC and that of the lines CF and DE are denoted by S_1 and S_2, respectively. Then, the point G is defined as the intersection of the lines FS_1 and BS_2. Based on this geometric configuration, they derived an invariant, given by

$$I_z = \frac{Q_{CBA} Q_{CED} Q_{CFG}}{Q_{CDA} Q_{CEF} Q_{CGB}}.$$

Note that the invariance of I_z can be easily verified using the same approach described in the previous section. Let us denote the virtual plane on which the four points B, C, F, and
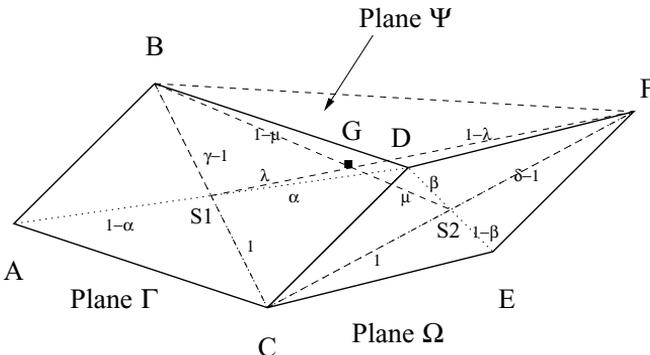
FIG. 2. A structure of six points given by Zhu *et al.*

G lie by Ψ, and the corresponding 3×3 transform matrix by $\mathbf{T}_\Psi$. Then, we can obtain the following projective relations for the six triangular areas,

$$w_C w_B w_A Q_{C'B'A'} = |\mathbf{T}_\Gamma| Q_{CBA}$$

$$w_C w_D w_A Q_{C'D'A'} = |\mathbf{T}_\Gamma| Q_{CDA}$$

$$w_C w_E w_D Q_{C'E'D'} = |\mathbf{T}_\Omega| Q_{CED}$$

$$w_C w_E w_F Q_{C'E'F'} = |\mathbf{T}_\Omega| Q_{CEF}$$

$$w_C w_F w_G Q_{C'F'G'} = |\mathbf{T}_\Psi| Q_{CFG}$$

$$w_C w_G w_B Q_{C'G'B'} = |\mathbf{T}_\Psi| Q_{CGB},$$

which simply lead to

$$\frac{Q_{C'B'A'} Q_{C'E'D'} Q_{C'F'G'}}{Q_{C'D'A'} Q_{C'E'F'} Q_{C'G'B'}} = \frac{Q_{CBA} Q_{CED} Q_{CFG}}{Q_{CDA} Q_{CEF} Q_{CGB}}.$$

Now, let us verify that the invariant I_p, derived in the previous section and I_z are equivalent to each other, by showing that

$$\frac{Q_{CBA} Q_{CED} Q_{CFG}}{Q_{CDA} Q_{CEF} Q_{CGB}} = \frac{Q_{ABC} Q_{DEF}}{Q_{ABD} Q_{CEF}}.$$

For notational simplicity, let us define some vectors and auxiliary parameters first; three vectors, $\vec{CA}, \vec{CD}$, and $\vec{CE}$ are denoted by the basis vectors $\vec{a}, \vec{d}$, and $\vec{e}$, respectively, and as shown in Fig. 2, the intersecting ratios of lines are defined by six parameters, given by

$$DS_1 : AS_1 = \alpha : (1 - \alpha); \quad DS_2 : ES_2 = \beta : (1 - \beta)$$

$$CS_1 : BS_1 = 1 : (\gamma - 1); \quad CS_2 : FS_2 = 1 : (\delta - 1)$$

$$GS_1 : GF = \lambda : (1 - \lambda); \quad GS_2 : GB = \mu : (1 - \mu).$$

Note that the three vectors $\vec{a}, \vec{d}$, and $\vec{e}$ and the four parameters α, β, γ, and δ uniquely specify the geometric configuration of the six points, while the other two parameters λ and μ are dependent ones. Then, other vectors can be expressed as the linear combination of the three basis vectors, in terms of these parameters, given by

$$\vec{CS}_1 = \alpha \vec{a} + (1 - \alpha) \vec{d}$$

$$\vec{CB} = \gamma \vec{CS}_1 = \alpha \gamma \vec{a} + (1 - \alpha) \gamma \vec{d}$$

$$\vec{CS}_2 = \beta \vec{e} + (1 - \beta) \vec{d}$$

$$\vec{CF} = \delta \vec{CS}_2 = \beta \delta \vec{e} + (1 - \beta) \delta \vec{d},$$

and $\vec{CG}$ can be written as

$$\vec{CG} = (1 - \lambda) \vec{CS}_1 + \lambda \vec{CF}$$

$$= \alpha(1 - \lambda) \vec{a} + \{(1 - \beta)\delta\lambda + (1 - \alpha)(1 - \lambda)\} \vec{d} + \beta\delta\lambda\vec{e}. \qquad (12)$$

It is noted that since $\vec{CG}$ can also be expressed as

$$\begin{aligned}
\vec{CG} &= (1-\mu)\vec{CS_2} + \mu\vec{CB} \\
&= \alpha\gamma\mu\vec{a} + \{(1-\beta)(1-\mu) + (1-\alpha)\gamma\mu\}\vec{d} + \beta(1-\mu)\vec{e},
\end{aligned} \tag{13}$$

and by comparing (12) and (13), we obtain the following relations for λ and μ,

$$1 - \lambda = \gamma\mu$$
$$1 - \mu = \delta\lambda$$
$$\lambda = \frac{\gamma - 1}{\gamma\delta - 1}$$
$$\mu = \frac{\delta - 1}{\gamma\delta - 1}.$$

Now, let Q_{ABCD} be the area of the quadrilateral, made of four vertices A, B, C, and D. Then, we can represent the areas of the triangles involved in the invariants I_p and I_z, in terms of Q_{ABCD}, Q_{CDEF}, and Q_{CBF}, scaled by four parameters α, β, γ, and δ. That is,

$$Q_{ABC} = (1-\alpha)Q_{ABCD}; \quad Q_{ACD} = \frac{1}{\gamma}Q_{ABCD}; \quad Q_{ABD} = \frac{\gamma - 1}{\gamma}Q_{ABCD}$$

$$Q_{CEF} = (1-\beta)Q_{CDEF}; \quad Q_{CED} = \frac{1}{\delta}Q_{CDEF}; \quad Q_{DEF} = \frac{\delta - 1}{\delta}Q_{CDEF}$$

$$Q_{CFG} = \frac{1-\lambda}{\gamma}Q_{CBF} = \frac{\delta - 1}{\gamma\delta - 1}Q_{CBF}$$

$$Q_{CGB} = \frac{1-\mu}{\delta}Q_{CBF} = \frac{\gamma - 1}{\gamma\delta - 1}Q_{CBF}.$$

Thus, I_p is expressed as

$$\begin{aligned}
I_p &= \frac{Q_{ABC}Q_{DEF}}{Q_{ABD}Q_{CEF}} \\
&= \frac{(1-\alpha)Q_{ABCD}\frac{\delta - 1}{\delta}Q_{CDEF}}{\frac{\gamma - 1}{\gamma}Q_{ABCD}(1-\beta)Q_{CDEF}} = \frac{(1-\alpha)(\delta - 1)\gamma}{(1-\beta)(\gamma - 1)\delta},
\end{aligned}$$

and I_z can be also expressed as

$$\begin{aligned}
I_z &= \frac{Q_{CBA}Q_{CED}Q_{CFG}}{Q_{CDA}Q_{CEF}Q_{CGB}} \\
&= \frac{(1-\alpha)Q_{ABCD}\frac{1}{\delta}Q_{CDEF}\frac{\delta - 1}{\gamma\delta - 1}Q_{CBF}}{\frac{1}{\gamma}Q_{ABCD}(1-\beta)Q_{CDEF}\frac{\gamma - 1}{\gamma\delta - 1}Q_{CBF}} = \frac{(1-\alpha)(\delta - 1)\gamma}{(1-\beta)(\gamma - 1)\delta},
\end{aligned}$$

which proves the equivalence of the two invariants.

In conclusion, since the invariant, proposed by Zhu $et\ al.$, is equivalent to the invariant I_p presented in this paper, the virtual point G is redundant and unnecessary to calculate the invariant for the constrained six points.

3. INVARIANTS FOR RESTRICTED SPACE LINES

In this section, we extend the notion of the invariant over space points to restricted space lines. By investigating the properties related to the projection of lines, we derive two new invariants for structured lines.

3.1. Properties Related to Line Projection

Before deriving invariants for restricted lines, let us discuss several properties related to the projection of lines first.

Note that a line on the plane Γ can be simply represented, in terms of the transformed coordinates, as

$$\mathbf{l}^\Gamma \cdot \mathbf{n}^\Gamma = 0, \tag{14}$$

where $\mathbf{l}^\Gamma = [a^\Gamma \ b^\Gamma \ c^\Gamma]^T$, which is the coefficient vector of the planar line.

It is noted that $\alpha \mathbf{l}^\Gamma$ is equivalent to $\mathbf{l}^\Gamma$, where α is a nonzero scaling term. To discriminate the projected image line and the space line, let us denote the image line by $\mathbf{l}'$ hereafter.

Remark 1. Given two space points whose transformed coordinates are $\mathbf{n}_A^\Gamma$ and $\mathbf{n}_B^\Gamma$, let $\mathbf{l}^\Gamma$ be the line passing through them. Then, the line can be expressed by the cross product of the two point vectors, given by

$$\alpha \mathbf{l}^\Gamma = \left(\mathbf{n}_A^\Gamma \times \mathbf{n}_B^\Gamma \right), \tag{15}$$

where α is a constant.

Remark 2. The cross product of two coplanar lines $\mathbf{l}_1^\Gamma$ and $\mathbf{l}_2^\Gamma$ is equal to the crossing point $\mathbf{n}^\Gamma$ of them, or

$$\beta \mathbf{n}^\Gamma = \left(\mathbf{l}_1^\Gamma \times \mathbf{l}_2^\Gamma \right), \tag{16}$$

where β is a constant.

Note that the above definition and Remarks also hold for points and lines on the image plane. That is, $\mathbf{n}^\Gamma$ and $\mathbf{l}^\Gamma$ can be substituted for $\mathbf{m}$ and $\mathbf{l}'$, respectively, in (14), (15), and (16).

LEMMA 1. *If the projection of a point on the plane Γ, whose transformed coordinate is $\mathbf{n}^\Gamma$, is given by (4), then the line on the plane Γ, denoted by the transformed coordinate $\mathbf{l}^\Gamma$, is mapped to the line $\mathbf{l}'$ by a projection matrix $\mathbf{V}_\Gamma$, s.t.,*

$$\mathbf{l}' = \varepsilon^\Gamma \mathbf{V}_\Gamma \mathbf{l}^\Gamma,$$
$$where \quad \mathbf{V}_\Gamma = \left[\mathbf{t}_2^\Gamma \times \mathbf{t}_3^\Gamma \quad \mathbf{t}_3^\Gamma \times \mathbf{t}_1^\Gamma \quad \mathbf{t}_1^\Gamma \times \mathbf{t}_2^\Gamma \right], \tag{17}$$

$\mathbf{t}_i^\Gamma$ *is the ith column vector of the matrix $\mathbf{T}_\Gamma$, and ε^Γ is a constant.*

Proof. Assume that $\mathbf{n}_A^\Gamma$ and $\mathbf{n}_B^\Gamma$ on the line $\mathbf{l}^\Gamma$ are projected to $\mathbf{m}_A$ and $\mathbf{m}_B$, respectively. Then, from (15) and (4), the line $\mathbf{l}'$, projected from $\mathbf{l}^\Gamma$, can be expressed as

$$
\begin{aligned}
\mathbf{l}' &= \frac{1}{\alpha'}(\mathbf{m}_A \times \mathbf{m}_B) \\
&= \frac{1}{\alpha'}\left(w_A \mathbf{T}_\Gamma \mathbf{n}_A^\Gamma \times w_B \mathbf{T}_\Gamma \mathbf{n}_B^\Gamma\right) \\
&= \frac{w_A w_B}{\alpha'}\left(\mathbf{T}_\Gamma \mathbf{n}_A^\Gamma \times \mathbf{T}_\Gamma \mathbf{n}_B^\Gamma\right).
\end{aligned}
$$

Note that for a 3×3 matrix $\mathbf{T}_\Gamma$ and 3×1 vectors $\mathbf{a}$ and $\mathbf{b}$, we have

$$
\begin{aligned}
\mathbf{T}_\Gamma \mathbf{a} \times \mathbf{T}_\Gamma \mathbf{b} &= \left(a_1 \mathbf{t}_1^\Gamma + a_2 \mathbf{t}_2^\Gamma + a_3 \mathbf{t}_3^\Gamma\right) \times \left(b_1 \mathbf{t}_1^\Gamma + b_2 \mathbf{t}_2^\Gamma + b_3 \mathbf{t}_3^\Gamma\right) \\
&= (a_1 b_2 - a_2 b_1)\left(\mathbf{t}_1^\Gamma \times \mathbf{t}_2^\Gamma\right) + (a_2 b_3 - a_3 b_2)\left(\mathbf{t}_2^\Gamma \times \mathbf{t}_3^\Gamma\right) + (a_3 b_1 - a_1 b_3)\left(\mathbf{t}_3^\Gamma \times \mathbf{t}_1^\Gamma\right) \\
&= \left[\mathbf{t}_2^\Gamma \times \mathbf{t}_3^\Gamma \quad \mathbf{t}_3^\Gamma \times \mathbf{t}_1^\Gamma \quad \mathbf{t}_1^\Gamma \times \mathbf{t}_2^\Gamma\right](\mathbf{a} \times \mathbf{b}), \quad (18)
\end{aligned}
$$

where $\mathbf{t}_i^\Gamma$ is the ith column vector of the matrix $\mathbf{T}_\Gamma$, a_j is the jth element of the vector $\mathbf{a}$, and b_k is the kth element of the vector $\mathbf{b}$, respectively. Therefore, using (18), $\mathbf{l}'$ can be written as

$$
\begin{aligned}
\mathbf{l}' &= \frac{w_A w_B}{\alpha'}\left[\mathbf{t}_2^\Gamma \times \mathbf{t}_3^\Gamma \quad \mathbf{t}_3^\Gamma \times \mathbf{t}_1^\Gamma \quad \mathbf{t}_1^\Gamma \times \mathbf{t}_2^\Gamma\right]\left(\mathbf{n}_A^\Gamma \times \mathbf{n}_B^\Gamma\right) \\
&= \varepsilon^\Gamma \mathbf{V}_\Gamma \mathbf{l}^\Gamma,
\end{aligned}
$$

where $\varepsilon^\Gamma = \frac{w_A w_B \alpha}{\alpha'}$, and $\mathbf{V}_\Gamma = \left[\mathbf{t}_2^\Gamma \times \mathbf{t}_3^\Gamma \quad \mathbf{t}_3^\Gamma \times \mathbf{t}_1^\Gamma \quad \mathbf{t}_1^\Gamma \times \mathbf{t}_2^\Gamma\right]$. $\blacksquare$

Remark 3. It is clear that the point and line projection matrices $\mathbf{T}_\Gamma$ and $\mathbf{V}_\Gamma$ satisfy the following relationship, given by

$$
\mathbf{V}_\Gamma^T \mathbf{T}_\Gamma = |\mathbf{T}_\Gamma|\mathbf{I}, \quad (19)
$$

where $\mathbf{I}$ is the 3×3 identity matrix.

LEMMA 2. *Consider three space lines on two adjacent planes Γ and Ω, constituting two pairs of coplanar lines with a shared line, as shown in Fig. 3, where $\mathbf{l}_i^j$ denotes the ith line on jth plane. Let $\mathbf{l}_i'$ be the projected image line of $\mathbf{l}_i^j$. Then, the following relationship holds,*

$$
D'_{123} = \frac{\varepsilon_1^\Gamma \beta'_{23} w_{23}}{\beta_{23}}|\mathbf{T}_\Gamma|D_{123}, \quad (20)
$$

where

$$
\begin{aligned}
D_{123} &= \left|\mathbf{l}_1^\Gamma \quad \mathbf{l}_2^\Gamma \quad \mathbf{l}_3^\Omega\right| \\
D'_{123} &= \left|\mathbf{l}_1' \quad \mathbf{l}_2' \quad \mathbf{l}_3'\right|.
\end{aligned}
$$
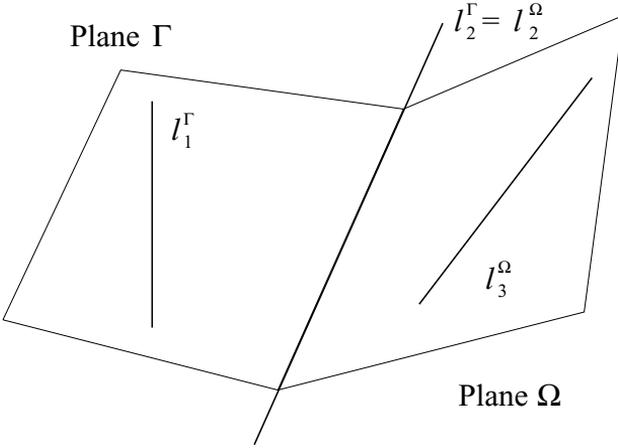
FIG. 3. Configuration of three lines on two adjacent planes.

Proof. Note that the determinant D'_{123} can be expressed as

$$D'_{123} = \mathbf{l}'^T_1 (\mathbf{l}'_2 \times \mathbf{l}'_3).$$

And, by using (2), (16), (17), and (19), this can be rewritten as

$$
\begin{aligned}
D'_{123} &= \varepsilon^{\Gamma}_1 \left(\mathbf{V}_{\Gamma} \mathbf{l}^{\Gamma}_1 \right)^T (\mathbf{l}'_2 \times \mathbf{l}'_3) \\
&= \varepsilon^{\Gamma}_1 \beta'_{23} \left(\mathbf{V}_{\Gamma} \mathbf{l}^{\Gamma}_1 \right)^T \mathbf{m}_{23} \\
&= \varepsilon^{\Gamma}_1 \beta'_{23} w_{23} \left(\mathbf{l}^{\Gamma}_1 \right)^T \mathbf{V}^T_{\Gamma} \mathbf{T}_{\Gamma} \mathbf{n}^{\Omega}_{23} \\
&= \frac{\varepsilon^{\Gamma}_1 \beta'_{23} w_{23}}{\beta_{23}} |\mathbf{T}_{\Gamma}| \left(\mathbf{l}^{\Gamma}_1 \right)^T \left(\mathbf{l}^{\Omega}_2 \times \mathbf{l}^{\Omega}_3 \right) \\
&= \frac{\varepsilon^{\Gamma}_1 \beta'_{23} w_{23}}{\beta_{23}} |\mathbf{T}_{\Gamma}| D_{123},
\end{aligned}
$$

where $\mathbf{m}_{23}$ is a virtual point, formed by $\mathbf{l}'_2$ and $\mathbf{l}'_3$. ∎

3.2. Derivation of Two New Invariants for Restricted Lines

From the relationship among three lines on two adjacent planes, given in (20), two new invariants for structured lines can be readily derived.

THEOREM 1. *For a set of five structured lines on two adjacent planes, which consists of two set of three coplanar lines as shown in Fig. 4, there exists an invariant I_{l1}, given by*

$$I_{l1} = \frac{D'_{123} D'_{543}}{D'_{143} D'_{523}}.$$

Proof. From (20), we have the following relationships for each D'_{ijk}.

$$D'_{123} = \frac{\varepsilon^{\Gamma}_1 \beta'_{23} w_{23}}{\beta_{23}} |\mathbf{T}_{\Gamma}| D_{123}$$
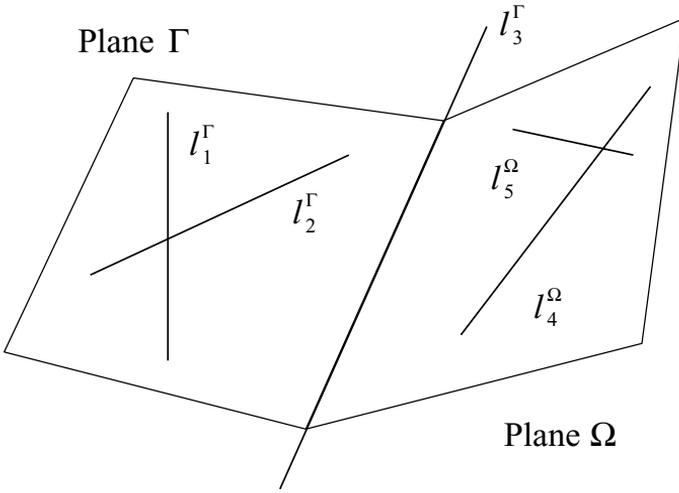
FIG. 4. Configuration of five structured lines on two adjacent planes.

$$D'_{543} = \frac{\varepsilon_5^{\Omega} \beta'_{43} w_{43}}{\beta_{43}} |\mathbf{T}_{\Omega}| D_{543}$$

$$D'_{143} = \frac{\varepsilon_1^{\Gamma} \beta'_{43} w_{43}}{\beta_{43}} |\mathbf{T}_{\Gamma}| D_{143}$$

$$D'_{523} = \frac{\varepsilon_5^{\Omega} \beta'_{23} w_{23}}{\beta_{23}} |\mathbf{T}_{\Omega}| D_{523},$$

yielding

$$I_{l1} = \frac{D'_{123} D'_{543}}{D'_{143} D'_{523}} = \frac{D_{123} D_{543}}{D_{143} D_{523}}.$$

∎

Note that even when the five lines are on a single plane in space, the invariance of I_{l1} is still preserved, which is the well-known invariant of five coplanar lines (FCL) [2].

THEOREM 2. *For a set of six structured lines, which is composed of three coplanar lines and three pairs of planar lines as depicted in Fig. 5, there exists an invariant I_{l2}, given by*

$$I_{l2} = \frac{D'_{124} D'_{235} D'_{316}}{D'_{125} D'_{236} D'_{314}}.$$

Proof. Using (20), the following relationships for each D'_{ijk} can be obtained.

$$D'_{124} = \frac{\varepsilon_4^{\Lambda} \beta'_{12} w_{12}}{\beta_{12}} |\mathbf{T}_{\Lambda}| D_{124}$$

$$D'_{235} = \frac{\varepsilon_5^{\Omega} \beta'_{23} w_{23}}{\beta_{23}} |\mathbf{T}_{\Omega}| D_{235}$$
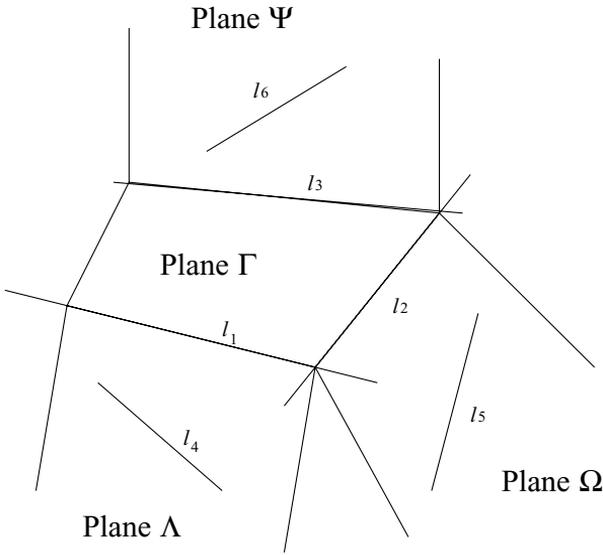
FIG. 5. Configuration of six structured lines.

$$D'_{316} = \frac{\varepsilon_6^\Psi \beta'_{31} w_{31}}{\beta_{31}} |\mathbf{T}_\Psi| D_{316}$$

$$D'_{125} = \frac{\varepsilon_5^\Omega \beta'_{12} w_{12}}{\beta_{12}} |\mathbf{T}_\Omega| D_{125}$$

$$D'_{236} = \frac{\varepsilon_6^\Psi \beta'_{23} w_{23}}{\beta_{23}} |\mathbf{T}_\Psi| D_{236}$$

$$D'_{314} = \frac{\varepsilon_4^\Lambda \beta'_{31} w_{31}}{\beta_{31}} |\mathbf{T}_\Lambda| D_{314}.$$

Therefore, we have

$$I_{l2} = \frac{D'_{124} D'_{235} D'_{316}}{D'_{125} D'_{236} D'_{314}} = \frac{D_{124} D_{235} D_{316}}{D_{125} D_{236} D_{314}}.$$

∎

Notice that the determinant D_{ijk} involves three lines on two adjacent planes, while the determinant Q_{ijk} for I_p involves three coplanar points. In a geometric sense, D_{ijk} can be interpreted as a scaled distance from one line l_i to the cross point of the other two coplanar lines, l_j and l_k, while the scale factor varies depending on the scale factors of the line parameters. While, by eliminating the scale factors using the determinant-ratio method, the derived invariants, I_{l1} and I_{l2} can be geometrically interpreted as the ratio of related distances.

Similarly to the point invariant case, degeneracy configurations as well as over-invariance problem may occur for the line invariants. Note that three concurrent or parallel lines will bring about the degeneracy conditions for I_{l1} and I_{l2}. And, since the proposed line invariants represent the distance ratio in geometric sense, there will exist many different configurations
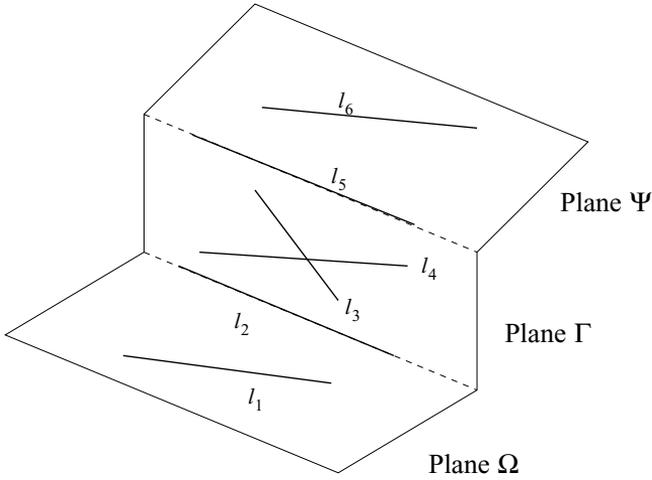
FIG. 6. Configuration of six lines given by Sugimoto.

which yield the same invariant value. Thus, for practical applications, the line invariant-based recognition algorithm should be designed to resolve these problems.

3.3. Comparison with the Invariant by Sugimoto

In [14], Sugimoto has also shown that there exists an invariant for six constrained space lines whose configuration includes a set of four coplanar lines and two pairs of coplanar lines on three mutually adjacent planes, as depicted in Fig. 6:

$$I_s = \frac{D'_{123}D'_{564}}{D'_{124}D'_{563}}.$$

By using the relationship among three lines on two adjacent planes, described in Lemma 2, this invariant can be easily verified as follows:

Since we have

$$D'_{123} = \frac{\varepsilon_3^\Gamma \beta'_{12} w_{12}}{\beta_{12}} |\mathbf{T}_\Gamma| D_{123}$$

$$D'_{564} = \frac{\varepsilon_4^\Gamma \beta'_{56} w_{56}}{\beta_{56}} |\mathbf{T}_\Gamma| D_{564}$$

$$D'_{124} = \frac{\varepsilon_4^\Gamma \beta'_{12} w_{12}}{\beta_{12}} |\mathbf{T}_\Gamma| D_{124}$$

$$D'_{563} = \frac{\varepsilon_3^\Gamma \beta'_{56} w_{56}}{\beta_{56}} |\mathbf{T}_\Gamma| D_{563},$$

the following invariant relation for cross-ratio holds

$$\frac{D'_{123} D'_{564}}{D'_{124} D'_{563}} = \frac{D_{123} D_{564}}{D_{124} D_{563}}.$$

Note that although the same number of lines are involved for the invariant I_s as for I_{l2} derived in this paper, the geometric configurations are quite different. Thus, depending on the geometry of the features which constitute the object model, an adequate invariant can be selectively employed for a specific application.

4. PROPOSED 3D OBJECT RECOGNITION ALGORITHM

In this section, we propose a novel 3D object recognition algorithm, employing the invariants for constrained 3D points and lines derived in the previous sections. Although the invariant value itself can serve as an effective key to object recognition, its limitations and related problems should also be taken into account for real applications. Note that the image formation and feature extraction processes inevitably cause positional measurement errors, yielding unreliable invariant values. And also, as was discussed in the previous sections, the over-invariance and degeneracy problems may occur. To cope with these problems, the proposed recognition algorithm employs several featuring components: table of model invariants, geometry test, and geometric hashing combined with thresholding, which will be explained in detail, respectively. Figure 7 shows the overall block diagram of the proposed algorithm.

The target object model is assumed to be composed of a group of features, such as dominant points or line segments, with the inherent information about the coplanarity among the features. For a given model, by analyzing the geometric relations among the features, all pertinent sets of features, satisfying the constrained structures are automatically extracted,
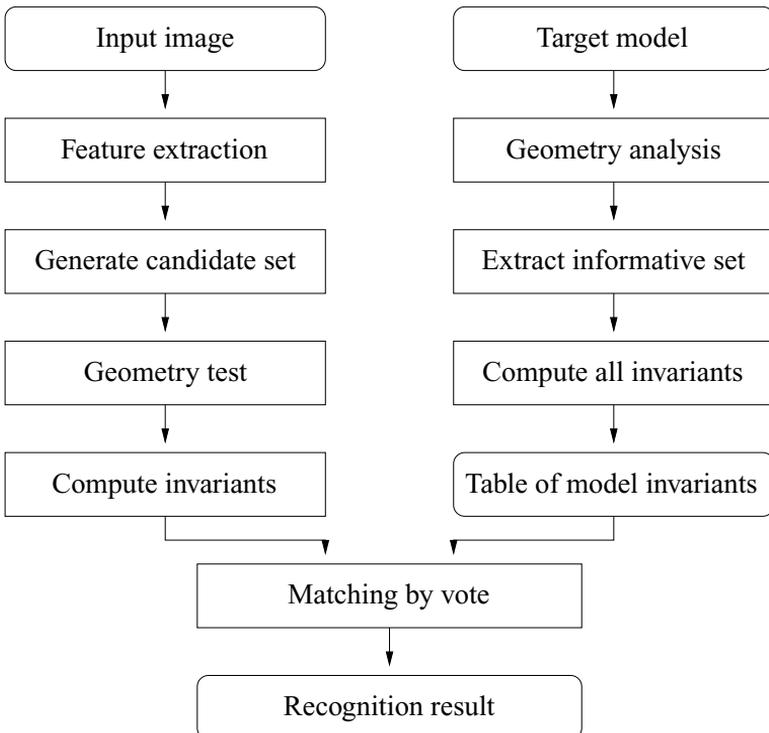
FIG. 7. The overall block diagram of the proposed algorithm.

which will be referred as *informative sets*. For example, let us consider a target object model, which has the following structural information:

plane index	lines
1	. . .
$\vdots$	
i	8 9 13
$\vdots$	
j	10 13 15
$\vdots$	
k	10 16
$\vdots$	
l	5 15

In this model, the set $\{8, 9, 13, 10, 15\}$ constitutes an *informative set* for I_{l1} and the set $\{10, 13, 15, 9, 16, 5\}$ composes another *informative set* to yield the invariant I_{l2}. Remember that the *informative set* should not include the degenerate configurations discussed in Sections 2.1 and 3.2. For each *informative set*, all possible invariant values obtained by exchanging the order of index are computed and recorded in the table of model invariants as the indices of features. The table is composed of segmented bins in terms of the invariant value. Note that the tolerance of error for invariant value is determined by the size of the bin. In order to maintain the same maximum tolerance rate Th_i for each invariant value, we choose the size of bin to be proportional to the invariant value. This *informative set* construction process is carried out in off-line mode, and thus it accelerates the overall matching process.

When an input image is applied to the recognition system, features are extracted, making a pool of input features. From this pool, candidate sets of features are selected, followed by a geometry test, which examines if the candidate sets have adequate geometric structure to yield the invariants. Note that although the exact 3D structure of a set of candidate features cannot be inferred due to projection, some useful geometrical relations among space features can be still found in 2D images. Figure 8 shows an example of geometry test on the *informative set* for I_{l1}. In this case, we can compute an invariant value of I_{l1} from the five lines, $l_1, \ldots, l_5$ shown in Fig. 8. Note that without any constraint, all permuted indexing of five lines (5!) should be compared with the model. However, in the image formed by projecting these lines, it is obvious that the following properties are still preserved:

- l_3 divides the set into two subsets of lines, $\{l_1, l_2\}$ and $\{l_4, l_5\}$.
- the circulating order of lines in each subset is preserved, like as $l_1 \rightarrow l_2 \rightarrow l_3$ and $l_4 \rightarrow l_5 \rightarrow l_3$.

Thus, by discarding feature sets which do not satisfy these geometric constraints, the number of candidate sets for I_{l1} is greatly reduced. Similarly, we can find and use adequate invariant structural properties for the *informative sets* of I_p and I_{l2}, to not only eliminate illegal candidate sets but also rearrange legal sets.

Following the geometry test, invariant value is computed for each candidate set in the input image, which is then used as the index to the table of the model invariants. Voting
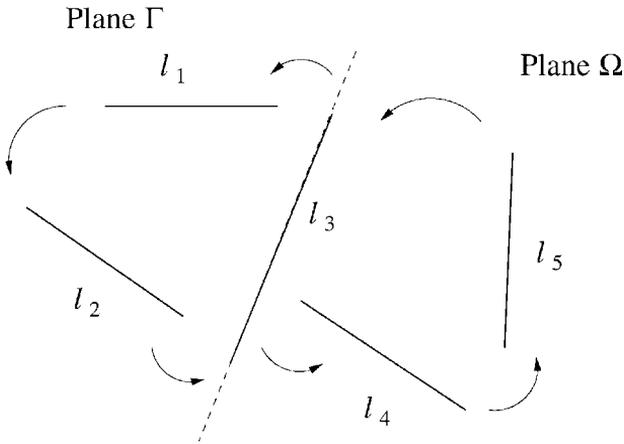
FIG. 8. An example of geometry test for I_{l1}.

is performed according to the correspondences among the candidate sets and the model *informative sets* recorded in the indexed bin.

By repeating this process including the geometry test, computation of invariants, and voting of all candidate sets, the correspondences among the model features and the input features are easily found based on the maximum scores credited by votes, and finally the matching is accomplished.

Notice that some *informative sets* of the 3D model would be unobservable or missing in a scene due to the change of viewing direction, occlusion, noise, and imperfect feature extraction, which in turn cause false alarms in matching. To cope with this problem, in
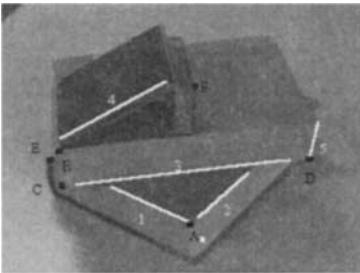
(a) View 1
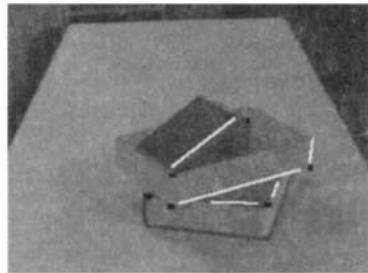
(b) View 2
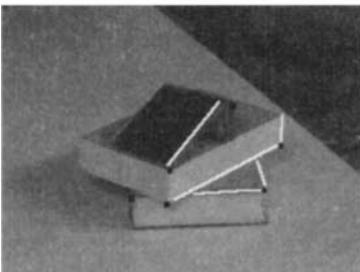
(c) View 3
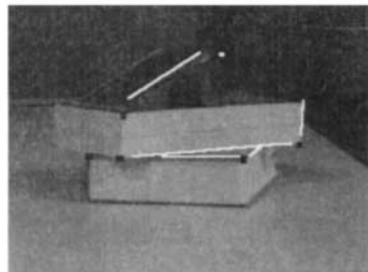
(d) View 4

FIG. 9. Test images for I_p and I_{l1}.
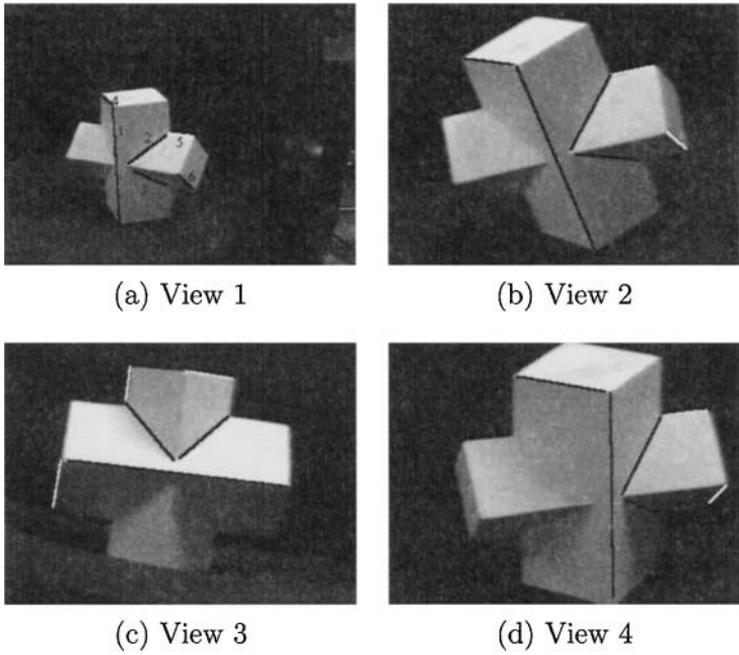
(a) View 1   (b) View 2

(c) View 3   (d) View 4

FIG. 10. Test images for I_{l2}.
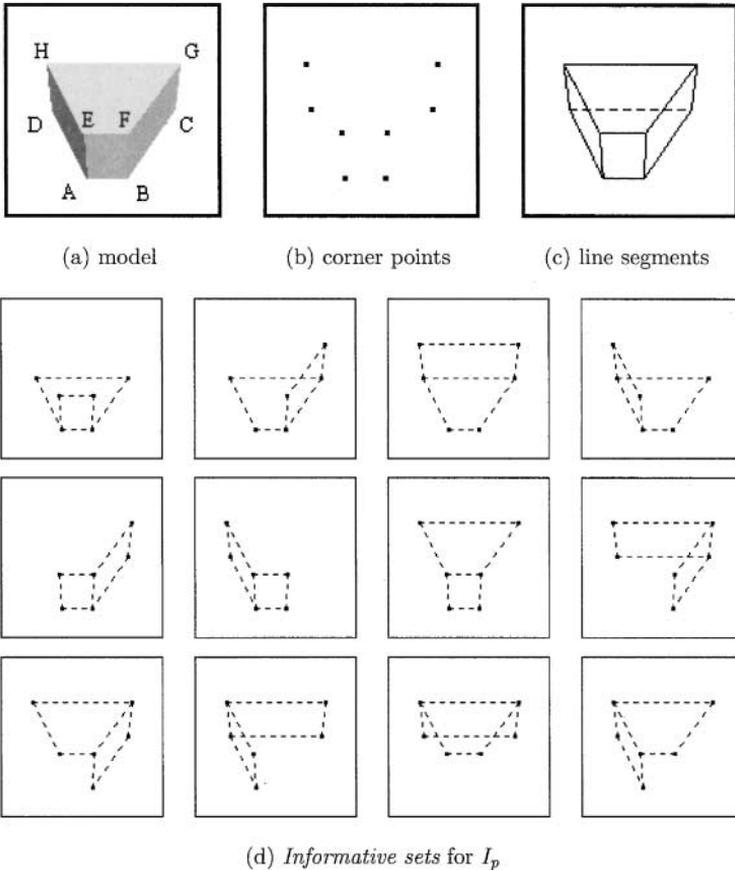
(a) model   (b) corner points   (c) line segments

(d) *Informative sets* for I_p

FIG. 11. Extraction of *informative sets* for point features.
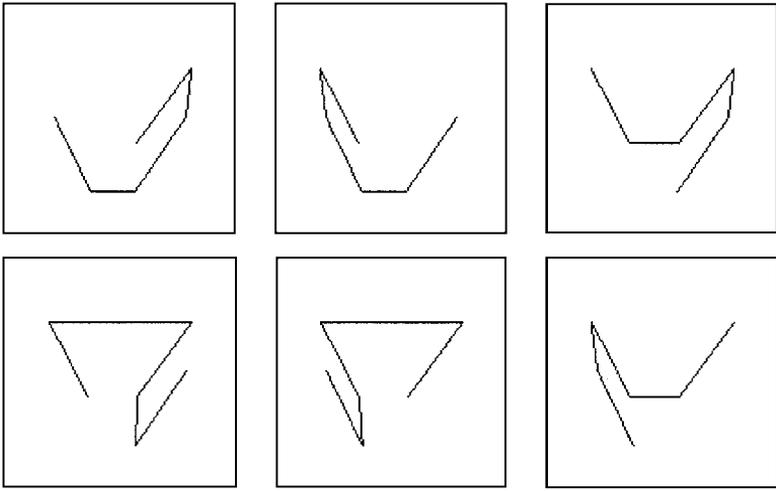
(a) *Informative sets* for I_{l1}
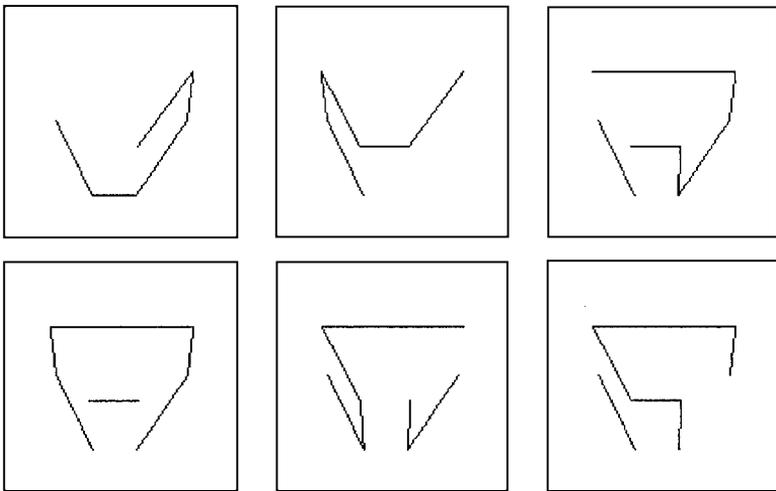
(b) *Informative sets* for I_{l2} (6 ones among 28 sets)

FIG. 12. Extraction of *informative sets* for line segments.

this work, we define a predefined threshold Th_s as the minimum ratio of voting score to the maximum expected score for each feature, and we utilize it in geometric hashing process to confirm the correspondences. By employing the geometric hashing technique with thresholds in matching, not only the problem of false alarm is efficiently solved, but also the computational complexity is greatly reduced.

5. EXPERIMENTS

We have carried out several simulations on both synthetic and real scenes to investigate the robustness of the proposed invariants. First, the numerical stability of I_p and I_{l1} for the features on two adjacent planes were examined. Figure 9 shows the four test images, which were obtained by taking images of the same object in different viewing conditions. The six
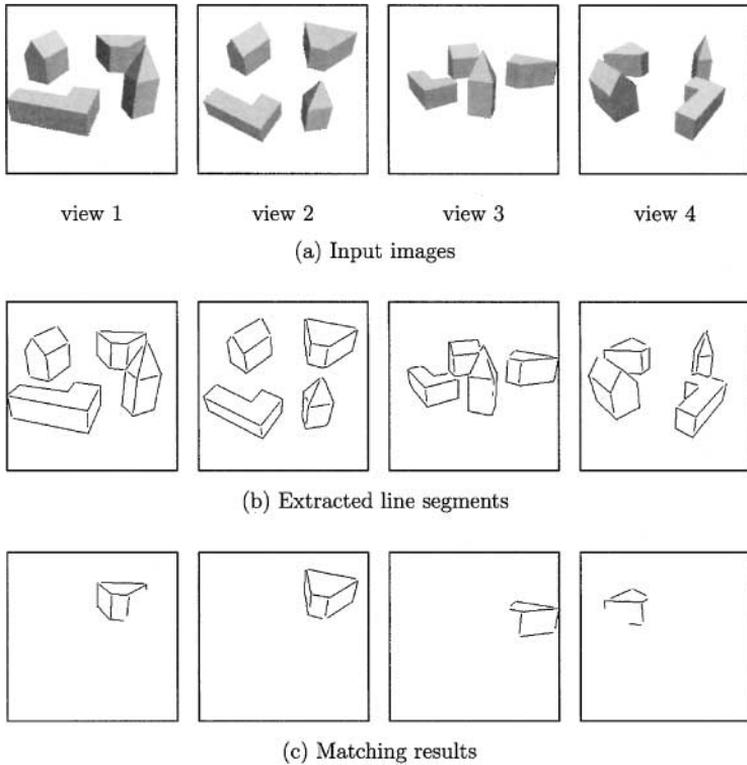
FIG. 13. Experiments on the synthetic images.

points, denoted by A–F were used to calculate I_p, and the five white lines, marked as 1–5, were used for I_{l1}, respectively. For each image, both invariants were calculated and listed in Table 1. The second row in Table 1 indicates the measured value of I_p for each image, and the third row shows the deviation from the average in percentage. Similarly, the fourth and fifth rows represent the calculated values of I_{l1}, and the errors to the mean, respectively. Note that although the positions of feature points and lines were disturbed due to imperfect extraction process, the deviation from the average value for each invariant were observed to be less than 2%. This implies that the invariants I_p and I_{l1} remain constant, in spite of the noise as well as substantial changes in viewpoint.

Tests for the invariant I_{l2} were also performed with four different images captured in different imaging conditions as shown in Fig. 10. I_{l2} was calculated using the six specified structured lines for each image. The results of the experiment are given in Table 2, in which

TABLE 1
The Measured Values of I_p and I_{l1}

View	view1	view2	view3	view4	Average
I_p	0.627	0.639	0.625	0.619	0.627
Error to the average (%)	0.0	1.9	0.3	1.5	0.9
I_{l1}	0.849	0.842	0.845	0.861	0.849
Error to the average (%)	0.0	0.8	0.2	1.4	0.6
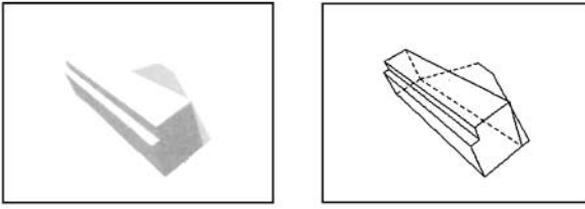
view 1
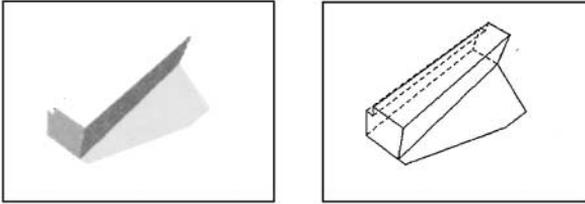
view 2

(a) Target model
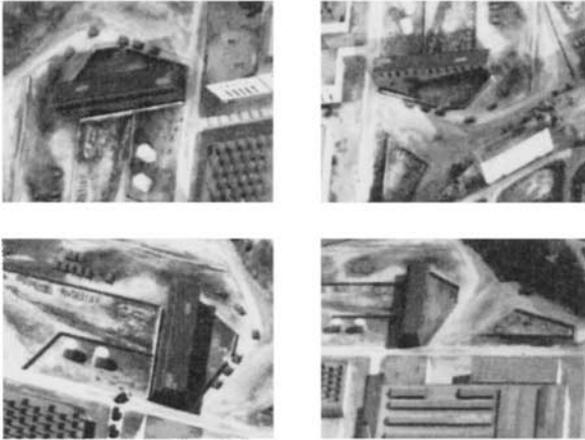
(b) Input images

FIG. 14. Experiments on the aerial images.

the second and third rows indicate the calculated value of I_{l2}, and the deviation from the average, respectively. Note that the deviation errors are less than 1%, demonstrating the stability of I_{l2}, even in the presence of the pixel noise and view variation. Thus, from these results we conclude that the proposed invariants can be applied effectively to 3D object recognition in real environment.

TABLE 2
The Calculated Value of I_{l2}

View	view1	view2	view3	view4	Average
I_{l2}	0.680	0.671	0.669	0.672	0.673
Error to the average (%)	1.0	0.3	0.6	0.1	0.5
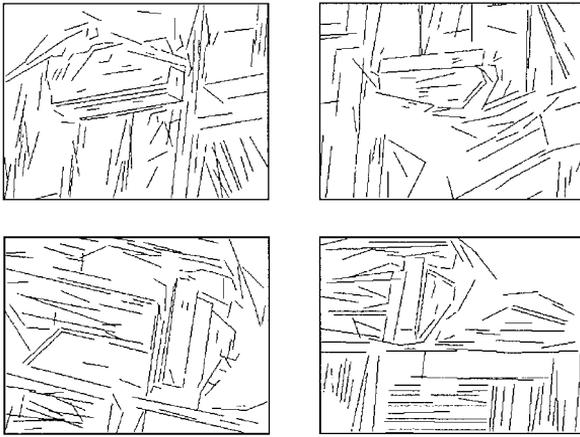
(c) Extracted line segments
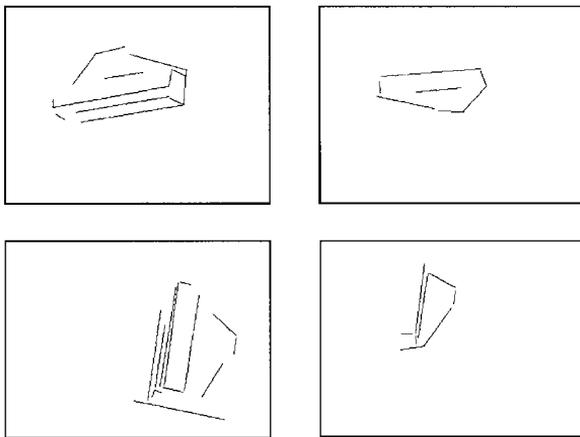
(d) Matched output

FIG. 14—*Continued*

Then, the performance of the proposed object recognition algorithm was demonstrated on synthetic images. In order to complete the recognition process, the table of model invariants should be constructed in advance by extracting *informative sets*. Figures 11 and 12 show the examples of extracted *informative sets* for the target model. A polyhedron was considered to be a target model as shown in Fig. 11a, of which the corner points and the line segments were provided as shown in Figs. 11b and 11c, respectively. From the corner points of the model, 12 *informative sets* for I_p, satisfying the structural requirements, were extracted, as depicted in Fig. 11d. Similarly, 6 *informative sets* for I_{l1} were extracted as shown in Fig. 12a. For I_{l2}, 28 *informative sets* were obtained, among which 6 are depicted in Fig. 12b. For each *informative set* of the model, an invariant was computed and saved in the table of model invariants.

Figure 13 demonstrates the recognition process based on line features. Four input images of the target model in Fig. 11c with three other objects were obtained by varying viewpoint as shown in Fig. 13a. For each input image, line segments were extracted by a line feature extractor [16] as shown in Fig. 13b. Notice that according to the viewing direction, a different set of features was obtained for each image. And, due to the imperfect performance of the line

extractor, the line segments were disturbed substantially. Moreover, owing to the occlusion by other objects, some features are missing as in views 1 and 4. Matching results using I_{l1} and I_{l2} are shown in Fig. 13c. In this experiment, the threshold values were chosen as $Th_i = 0.1$ and $Th_s = 0.2$. In spite of positional errors and missing features, it is shown that the target object is correctly recognized in all the images.

Simulations of the proposed object recognition algorithm on real images also have been performed. The threshold values were chosen as the same as in the synthetic image test. In Fig. 14a, the target model is provided with the characterizing line features. Four input test images with different viewpoints are given in Fig. 14b. The line segments extracted from each input image and final matching results are shown in Figs. 14c and 14d, respectively. Comparing these results, although the recognized set of features are somewhat different according to the viewing direction and image condition, the proposed recognition algorithm has located and identified the target object correctly in all cases.

6. CONCLUSIONS

In this paper, we have presented new projective invariants for structured space points and lines. First, based on the projection matrix representation and determinant ratio technique, we showed that the auxiliary point involved in determining the invariant for six points on two adjacent planes, proposed by Zhu *et al.* [1], is redundant, and we derived a new simpler formulation for this invariant. Then, by investigating the properties related to line projection, two new invariants for structured lines were proposed: one is an invariant for five lines on two adjacent planes and the other is an invariant for six lines on four planes. Based on the proposed invariants, a novel 3D object recognition algorithm has been suggested, employing a geometric hashing technique. Experimental results on various test images with different views showed that the proposed invariants remain numerically stable, even in a noisy environment. Moreover, the proposed object recognition algorithm was shown to perform accurately and robustly, even when some features are disturbed by noise or missed by occlusion.

APPENDIX: NOMENCLATURE

$\mathbf{X}$	3D space point
$\mathbf{x}$	2D image point
$\mathbf{P}$	Projection matrix
$\mathbf{M}$	3D space point in the homogeneous coordinates
$\mathbf{m}$	2D image point in the homogeneous coordinates
ω	Scaling factor
$\Gamma, \Sigma, \Omega, \Psi$	Planes
$\mathbf{U}_\Gamma$	Euclidean transform matrix of points on Γ onto the XY plane
$\mathbf{T}_\Gamma$	Projection matrix of the transformed points on Γ to the image plane
$\mathbf{V}_\Gamma$	Projection matrix of the transformed lines on Γ to the image plane
$\mathbf{t}_i^\Gamma$	ith column of $\mathbf{T}_\Gamma$
$\mathbf{n}^\Gamma$	Transformed point on Γ onto the XY plane
$\mathbf{l}^\Gamma$	Transformed line on Γ onto the XY plane
Q_{ijk}	Area of the triangle with vertices, i, j, and k
$\alpha, \beta, \gamma, \delta, \lambda, \mu$	Geometric parameters
I_p	Point invariant

I_z Point invariant derived by Zhu *et al*.
I_{l1} First line invariant
I_{l2} Second line invariant
I_s Line invariant derived by Sugimoto

REFERENCES

1. Y. Zhu, L. D. Seneviratne, and S. W. E. Earles, New algorithm for calculating an invariant of 3D point sets from a single view, *Image Vision Comput*. **14**, 1996, 179–188.

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

3. 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*. **13**(10), 1991, 971–991.

4. S. Carlsson, Projectively invariant decomposition and recognition of planar shapes, *Internat. J. Comput. Vision* **17**(2), 1996, 193–209.

5. I. Weiss, Noise-resistant invariants of curves, *IEEE Trans. Pattern Anal. Machine Intell*. **15**(9), 1993, 943–948.

6. I. Weiss, Geometric invariants and object recognition, *Internat. J. Comput. Vision* **10**(3), 1993, 207–231.

7. J. B. Burns, R. S. Weiss, and E. M. Riseman, View variation of point-set adn line-segment features, *IEEE Trans. Pattern Anal. Machine Intell*. **15**(1), 1993, 51–68.

8. E. B. Barrett, P. M. Payton, N. N. Haag, and M. H. Brill, General methods for determining projective invariants in imagery, *CVGIP: Image Understand*. **53**, 1991, 46–65.

9. R. Hartley, Invariants of lines in space, in *Proceedings of Image Understanding Workshop,* 1993, pp. 737–744.

10. L. Quan, Invariants of six points and projective reconstruction from three uncalibrated images, *IEEE Trans. Pattern Anal. Machine Intell*. **17**(1), 1995, 34–46.

11. D. Weinshall, Model-based invariants for 3D vision, *Internat. J. Comput. Vision* **10**(1), 1993, 27–42.

12. P. C. Wayner, Efficiently using invariant theory for model-based matching, in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,* 1991, pp. 473–478.

13. C. A. Rothwell, D. A. Forsyth, A. Zisserman, and J. L. Mundy, Extracting projective structure from single perspective views of 3D point sets, in *Proceedings of IEEE International Conference on Computer Vision*, Berlin, Germany, 1993, pp. 573–582.

14. Akihiro Sugimoto, Geometric invariant of noncoplanar lines in a single view, in *Proceedings of IEEE International Conference on Pattern Recognition,* 1994, pp. 190–195.

15. Y. Lamdan, J. T. Schwartz, and H. J. Wolfwon, Affine invariant model-based object recognition, *IEEE Trans. Robotics Automat*. **5**, 1990, 578–589.

16. R. Nevatia and K. R. Babu, Linear feature extraction and description, *Comput. Graphics Image Process*. **13**, 1980, 257–269.