

A Line Feature Matching Technique Based on an Eigenvector Approach¹

Sang Ho Park

School of Electrical Engineering, Seoul National University, Seoul, Korea 151-742

Kyoung Mu Lee

Department of Electronics and Electrical Engineering, Hong-Ik University, 72-1 Sangsu-dong, Mapo-gu, Seoul 121, Korea

and

Sang Uk Lee²

School of Electrical Engineering, Seoul National University, Seoul, Korea 151-742

Received May 8, 1998; accepted September 29, 1999

In this paper, we propose a new eigenvector-based line feature matching algorithm, which is invariant to the in-plane rotation, translation, and scale. First, in order to reduce the number of possible matches, we use a preliminary correspondence test that generates a set of finite candidate models, by restricting combinations of line features in the input image. This approach resolves an inherent problem relating to ordering and correspondence in an eigenvector/modal approach. Second, we employ the *modal analysis*, in which the Gaussian weighted proximity matrices for reference and candidate models are constructed to record the relative distance and angle information between line features for each model. Then, the *modes* of the proximity matrices of the two models are compared to yield the *dissimilarity measure*, which describes the quantitative degree of the difference between the two models. Experimental results for synthetic and real images show that the proposed algorithm performs matching of the line features with affine variation fast and efficiently and provides the degree of dissimilarity in a quantitative way. © 2000 Academic Press

1. INTRODUCTION

Matching between an object and a reference model is a fundamental task in computer vision and a key step for a variety of computer vision applications, including automated

¹ This work was supported by the Agency for Defense Development.

² To whom correspondence should be addressed. Fax: +82-2-880-8201. E-mail: sanguk@sting.snu.ac.kr.

target recognition, bin picking, surveillance, and so on [1, 2]. Various techniques have been proposed to solve this important problem in the past, such as the graph matching technique [3], the continuous relaxation technique [4–7], and the modal matching approaches [8–10]. These methods differ in their descriptions for the model as well as for the matching algorithms they use.

In general, geometric features provide much more robust and useful information than do photometric or image features for extracting and describing the objects in a scene. And the higher the level of descriptions at which matching is attempted, the more likely the descriptions are to be invariant to imaging changes. However, this gain may be offset by the error and deficiencies of the algorithms that compute these descriptions [6]. Among many geometrical features, from the viewpoint at the level of description, line features are the most appropriate ones, since most manmade objects can be effectively characterized by straight-line edges. Moreover, the extraction of these line features is much easier than that of other features [11].

The structural description for a line model can be established by exploiting the properties and the mutual relations between the line features in the model. Once the structural descriptions for the reference model and the input image are established after a reliable edge processing, the matching algorithm searches for an object, whose structural description is similar to that for the reference model.

Among many existing matching algorithms employing the line model, the relaxation-based optimization techniques using the attributed relational graph description model have shown great potential in many applications, since they solve the labeling problems efficiently [4–7]. These labeling methods utilize the information on relations between the features and perform the matching by updating the probability of assigning the label features to the reference features iteratively, based on the current probability and the relational constraints on the label features.

However, these methods normally assume small geometric distortions between the reference and object models, suffering from somewhat wider shape deformation, local shape variation, and the input noise. Moreover, the overall performance is often severely influenced by the selection of the initial probability; thus the convergence is not guaranteed to the global minimum of the cost function. In addition, since the number of features extracted from the input images is often much greater than that extracted from the reference model, the relational description becomes very complicated, requiring enormous computational complexity. In fact, the relaxation-based methods do not provide a canonical description for the model.

Recently, new methods, based on the eigenvector (modal) analysis for point correspondence or contour matching, have been proposed by several researchers [8–10, 12]. The main notion behind the modal analysis is the transformation of the features into a hyper-space of the generalized axes of symmetry or modes of a shape, so that the feature–mode relationships can be readily examined.

Scott and Longuet-Higgins proposed an efficient method to associate 2-D point features for two arbitrary shapes [8]. In [8], an interimage shape matrix, called the proximity matrix, is used to record the affinity of all possible pairwise matches, and the eigenvectors of this matrix are employed to determine the correspondences between two sets of points. However, the drawbacks are that it cannot cope with relatively large rotations in the image plane and it is sometimes very difficult to compute the interimage distances. In addition, the overall

performance of [8] is observed to be significantly affected by the difference in the number of patterns in two images.

Shapiro and Brady proposed an improved method for point–feature correspondence, based on a modal shape description using the intrimage proximity matrix [9]. By incorporating the shape information into a low-level matching process, the method can cope easily with rotations and translations in the image plane. However, although the method provides appropriate results for small-scale changes, the performance becomes unreliable for large-scale changes. In addition, if the number of feature points in two images is quite different, then the performance of the method also degrades substantially.

Recently, Sclaroff and Pentland [12] proposed an algorithm based on the eigenmodes of a shape matrix for point correspondence and shape recognition. By using Galerkin interpolation [13], they obtain the shape description, which is independent of the sampling points, and extract canonical frequency-ordered eigenmodes for shape. Point correspondences are then determined by comparing their displacement in modal space, and the shape comparison can be made by evaluating the amount of deformation energy required to align them. The parametric eigenspace method which represents trained images using principal component analysis also has been applied to the correspondence of black and white images and pose measurement [14].

In general, the conventional eigenvector approaches yield relatively good results for the point–feature correspondence between two shapes of similar dimensions, provided that an appropriate ordering of features is given. However, a direct application of these methods to the matching or registration of line features causes several problems. The shape metric, which defines the relations between the line features, is not so simple as in the case of the point features. And, in registration problems, the distribution and the number of features in an input image are quite different from those in a reference pattern, due to the background objects and noise. Therefore, direct computation and comparison of the modes for the whole input pattern and reference model are meaningless. Another problem relating to the eigen mode-based matching is that the modes are not invariant to the numbering order of the features in a model, implying that, if the labeling order of the model features is changed, the modes also are altered substantially. Although this problem can be remedied somehow, by solving the correspondence problem first, there still remains a problem of sign correction between corresponding modes [9, 12]. Therefore, in general, if the reference features and input features are not consistent with each other, the comparison of the modes is not appropriate.

In this paper, we propose a new efficient eigenvector-based line feature matching or registration algorithm for target recognition, mainly for manmade objects. The proposed algorithm is fast, 2D rotation- scale- and shift-invariant, and robust to shape deformation or input noise. Let us refer to the set of all the line features, extracted from the input image, as the label image. Also, the line feature in the reference model is referred to as the reference feature and the line feature in the label image as the label feature. In order to alleviate the computational cost and enhance the robustness, we employ a two-phase matching strategy. First, in order to reduce the number of possible matches before the modal matching, we employ a preliminary compatibility test between pairs of reference features and label features, using the compatibility measure of simple binary relations between them. Then, by selecting combinations of label features which are likely to correspond to the reference features, we can construct a set of finite candidate models, which can be matched

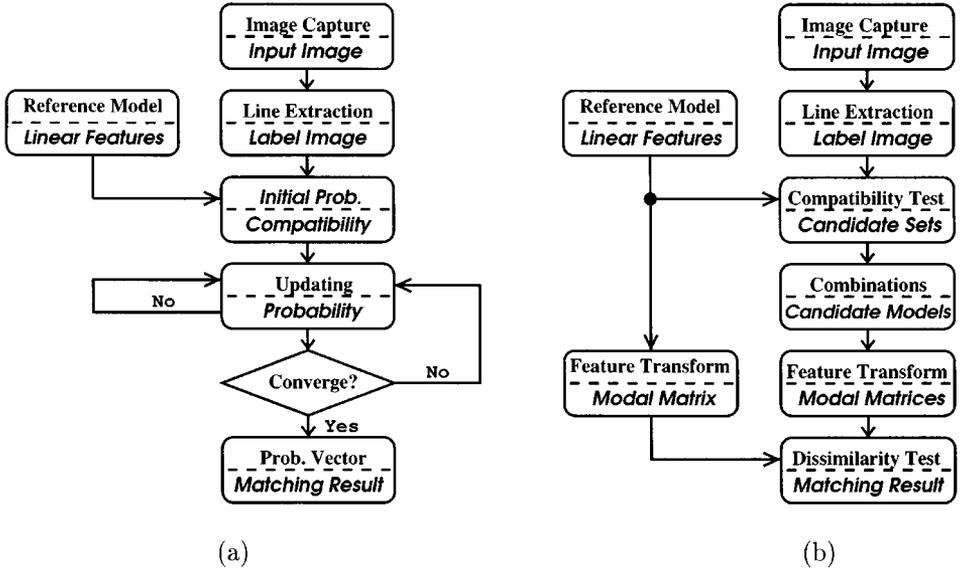


FIG. 1. Comparison of the relaxation labeling and proposed modal approach: (a) relaxation labeling; (b) proposed modal approach.

to the reference model. By constructing the candidate models, both the feature dimension and the label ordering problems are readily resolved in our approach, since each candidate feature set satisfies a one-to-one correspondence to the reference features. We show that this approach significantly simplifies the modal matching. Second, the proximity matrices for both the reference and candidate models are obtained by employing a Gaussian weighted metric which records the relative distance and angle information between line features for each model. By comparing the modes of those matrices, we can evaluate the dissimilarity between the reference and candidate models. Then, based on the sum of weighted Euclidean distances between the corresponding modes for the reference and the candidate models, the degree of dissimilarity can also be easily evaluated quantitatively.

The procedures for conventional relaxation labeling and the proposed algorithm are compared in Fig. 1. Figure 1a shows the flowchart for the relaxation labeling, where the probability vector is updated until it converges. In applying relaxation labeling, usually it is difficult to obtain an appropriate initial probability and to decide when to stop updating the iterations. Figure 1b depicts the procedure of the proposed matching algorithm. The compatibility test constrains the possible combinations of label features and prevents the algorithm from performing a blind tree search. The feature transformation allows the comparison to be executed in an orthogonal modal space. In contrast to relaxation labeling, the proposed modal approach is quite stable, since it is noniterative and does not depend on the initial conditions.

2. CONSTRUCTION OF CANDIDATE SETS USING BINARY RELATIONS

The modal approach for matching is applicable only to the case when the numbers of reference and label features are equal. But, in most cases, the number of reference features

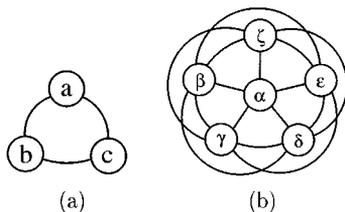


FIG. 2. Graph representation of a reference model and a label image: (a) reference model; (b) label image.

is much smaller than the number of label features. If the modal approach can be applied to these real cases, the candidate models, which satisfy one-to-one feature correspondence to the reference model, should be extracted in advance. The easiest way to extract the candidate models is to consider all the combinations of label features. However, if the number of label features is much greater than that of reference features, this scheme requires an enormous number of combinations.

Instead of considering all the possible combinations of label features, we first attempt to reduce the number of possible combinations. That is, we restrict the label features which can correspond to the reference features, based on the constraint relations between the reference features and the label features. This step significantly reduces the search range and eliminates the inherent problem relating to the conventional eigenvector approach: the alignment of eigenvectors extracted from different feature sets. Thus, it is a very important step not only in alleviating the computational cost, but also in improving the overall performance.

Figure 2 shows the schematic diagram of this approach, in which the nodes represent features, and arcs represent binary relations between them. Figures 2a and 2b represent a reference model composed of three features and a label image with six features, respectively. Now, the purposes of the proposed algorithm are to assign the most appropriate label features to reference features based on the relational constraints and to extract finite candidate feature sets among the label features which are most likely to be compared to the reference model.

2.1. Binary Relation Models for Line Features

For a given reference model, composed of M line features, the object corresponding to the reference model should be searched from the input image. Let us assume that the label image has N line features, where $N > M$. Also let us denote the ordered sets of the reference features and the label features by \mathcal{Q}_r and \mathcal{Q}_c , respectively, then we have

$$\begin{aligned}\mathcal{Q}_r &= \{a_1, a_2, \dots, a_M\} \\ \mathcal{Q}_c &= \{\alpha_1, \alpha_2, \dots, \alpha_M, \dots, \alpha_N\},\end{aligned}\quad (1)$$

where a_i and α_j are the i th reference feature and the j th label feature, respectively. For example, in Fig. 2, $\mathcal{Q}_r = \{a, b, c\}$ and $\mathcal{Q}_c = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta\}$.

By denoting the set of binary relations between two features by \mathbf{r} , we have

$$\begin{aligned}\mathbf{r}(a_m, a_n) &= \{r_i(a_m, a_n) \mid a_m, a_n \in \mathcal{Q}_r\}, \quad i = 1, \dots, N_r, \\ \mathbf{r}(\alpha_r, \alpha_s) &= \{r_i(\alpha_r, \alpha_s) \mid \alpha_r, \alpha_s \in \mathcal{Q}_c\}, \quad i = 1, \dots, N_r,\end{aligned}\quad (2)$$

where N_r is the number of binary relations. In Fig. 2, the arcs connecting two nodes represent the binary relations, which can be directional or not, depending on the properties

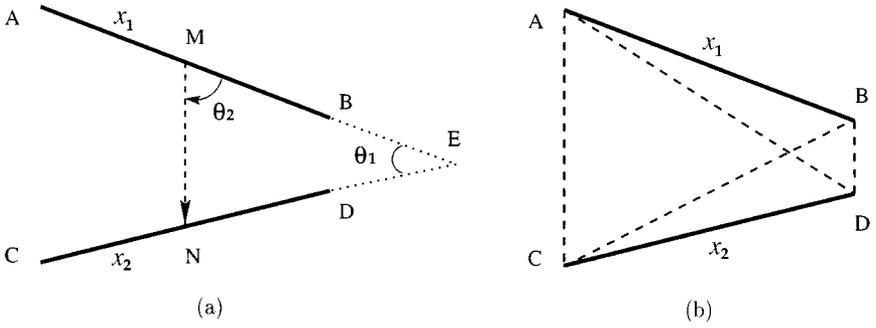


FIG. 3. Binary relations between a pair of line features: (a) angular relations; (b) distance relations.

of the relations. As shown in Fig. 3, in order to describe the mutual relations between a pair of features, in our approach, we employ the four different measures ($N_r = 4$)

$$\begin{aligned}
 r_1(x_1, x_2) &= \theta_1, \\
 r_2(x_1, x_2) &= \theta_2, \\
 r_3(x_1, x_2) &= \overline{AB}/\overline{CD} \\
 r_4(x_1, x_2) &= (\overline{AB} + \overline{CD})/d,
 \end{aligned} \tag{3}$$

where $d = (\overline{AC} + \overline{AD} + \overline{BC} + \overline{BD})/4$.

Note that r_1 is the angle which two line segments x_1 and x_2 form, r_2 is the directional angle from the midpoint of x_1 to that of x_2 , which is the angle measured from \overrightarrow{AB} to \overrightarrow{MN} in Fig. 3a. r_3 is the length of x_1 relative to that of x_2 , and r_4 is the ratio of the sum of line segments to the average distance between the endpoints of the line segments. Note that r_2 and r_3 are directional, while r_1 and r_4 are not; therefore, in general, $r_i(x_1, x_2) \neq r_i(x_2, x_1)$ for $i = 2, 3$, and $r_i(x_1, x_2) = r_i(x_2, x_1)$ for $i = 1, 4$. The above four binary relations describe the relative orientation, position, length, and distance between two segments, respectively. Figure 4 illustrates the role of each binary relation in describing the shape of a pair of lines. Figures 4b–4e are some typical variations of the pair of reference line segments in Fig. 4a while r_1 is kept fixed, including the mirrored pattern in Fig. 4b. Although some relations of each pattern may be the same as those of the reference model, they can be easily distinguished from the reference by r_2, r_3, r_4 , respectively. Figure 4e is the scaled version of the original, and in this case, all the binary relations for both patterns are the same, indicating the same pattern.

Now, let us define the *compatibility measure* between a pair of reference features (a_m, a_n) and a pair of label features (α_r, α_s) as

$$\mathcal{S}(a_m, a_n; \alpha_r, \alpha_s) = 1 / \left(1 + \sum_{i=1}^{N_r} \frac{|r_i(a_m, a_n) - r_i(\alpha_r, \alpha_s)|}{\omega_i} \right), \tag{4}$$

where the ω_i 's are the weights for each binary relation, and they should be carefully chosen so that the contribution of each relation to the measure \mathcal{S} becomes approximately equal. In this paper, we employ the mean of the absolute difference of $r_i(a_m, a_n)$ and $r_i(\alpha_r, \alpha_s)$ for ω_i , given by

$$\omega_i = \frac{1}{MN(M-1)(N-1)} \sum_{m,n,m \neq n}^M \sum_{r,s,r \neq s}^N |r_i(a_m, a_n) - r_i(\alpha_r, \alpha_s)|, \quad i = 1, \dots, N_r. \tag{5}$$

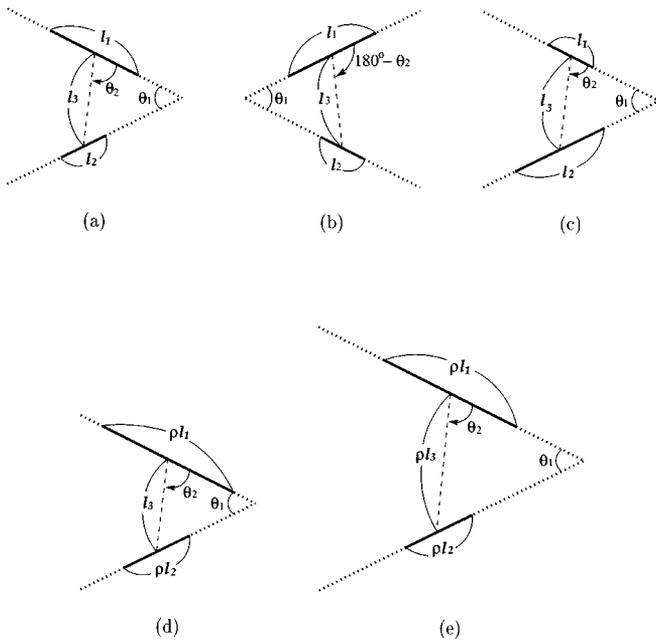


FIG. 4. The role of binary relations in discriminating possible varieties: (a) a reference; (b) different r_2 (mirrored pattern); (c) different r_3 ; (d) different r_4 ; (e) ρ scaled.

Thus, each ω_i can normalize the scale of the factor adaptively. The compatibility measure $\mathcal{S}(a_m, a_n; \alpha_r, \alpha_s)$ yields the maximum value 1.0 when the binary relations between the reference pair and the label pair are the same, while it monotonically decreases as the difference increases.

2.2. Construction of the Candidate Models from a Label Image

One might find the correspondences between the reference and label features by simply comparing the compatibility measures \mathcal{S} , using the blind tree search technique [15]. However, in general, since the number of label features N is much greater than the number of reference features M , there are a total of $N P_M$ cases to be examined. For example, for $M = 5$ and $N = 50$, there are about 254 million cases to search for the best match, which is practically impossible to implement in the real world. In this section, we present a method to construct the candidate model sets for a label image by imposing some constraints on binary relations. As a result, we can reduce the search space considerably, and the correspondence problem between the reference and candidate sets becomes trivial.

Let us construct the compatibility matrix \mathbf{R}_{mn} for each pair of reference features (a_m, a_n) , given by

$$\mathbf{R}_{mn} = \begin{bmatrix} \mathbf{g}_{mn}^T(1) \\ \mathbf{g}_{mn}^T(2) \\ \vdots \\ \mathbf{g}_{mn}^T(N) \end{bmatrix}_{N \times N}, \quad 1 \leq m, n \leq M, m \neq n, \tag{6}$$

where $\mathbf{g}_{mn}^T(i)$ denotes the i th row vector of size $1 \times N$, and the (i, j) th element of the matrix

$[\mathbf{R}_{mn}]_{ij}$, equivalently the j th element of $\mathbf{g}_{mn}^T(i)$, is defined by

$$[\mathbf{R}_{mn}]_{ij} = [\mathbf{g}_{mn}^T(i)]_j = \begin{cases} \mathcal{S}(a_m, a_n; \alpha_i, \alpha_j) & \text{if } \mathcal{S}(a_m, a_n; \alpha_i, \alpha_j) > \kappa \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

where κ is a prespecified threshold. Notice that the matrix \mathbf{R}_{mn} represents the compatibilities between a pair of reference features (a_m, a_n) and all the pairs of label features. It can be assumed that the reference pair (a_m, a_n) and the label pair (α_i, α_j) have similar relations if $\mathcal{S}(a_m, a_n; \alpha_i, \alpha_j)$ is greater than κ . In order to consider all the possible label pairs, in this paper, we choose a relatively small value $\kappa = 0.6$ for threshold.

Now, let us define the vectors ψ_m and ϕ_m for every reference feature a_m as

$$\psi_m = \begin{bmatrix} u_m(1) \\ u_m(2) \\ \vdots \\ u_m(N) \end{bmatrix}, \quad \phi_m = \begin{bmatrix} v_m(1) \\ v_m(2) \\ \vdots \\ v_m(N) \end{bmatrix}, \quad 1 \leq m \leq M, \quad (8)$$

where the i th components of both vectors are given by

$$u_m(i) = \begin{cases} 1, & \text{if all } \mathbf{g}_{mn}(i) \neq \mathbf{0}, n = 1, \dots, M \\ 0, & \text{otherwise,} \end{cases} \quad (9)$$

$$v_m(i) = \begin{cases} \sum_{n=1}^M \|\mathbf{g}_{mn}(i)\|_\infty, & \text{if all } \mathbf{g}_{mn}(i) \neq \mathbf{0}, n = 1, \dots, M \\ 0, & \text{otherwise,} \end{cases} \quad (10)$$

where $\mathbf{0}$ is the null vector. Note that $u_m(i)$ indicates whether the i th label feature α_i can be matched to the m th reference feature a_m or not, and $v_m(i)$ represents the degree of matching between them.

Then, the ordered candidate set of label features Λ_{a_m} for each reference feature a_m can be obtained by selecting the label features whose indices correspond to those of nonzero elements in ψ_m in decreasing order of the magnitude of the corresponding elements in ϕ_m . That is, by denoting the ordered index set $I_m, m = 1, \dots, M$, by

$$I_m = \{\gamma_{mi} \mid 1 \leq i \leq t_m, u_m(\gamma_{mi}) \neq 0, v_m(\gamma_{mi}) \geq v_m(\gamma_{m(i+1)})\}, \quad (11)$$

where t_m is the number of nonzero elements in ψ_m , the candidate set of label features Λ_{a_m} can be expressed as

$$\Lambda_{a_m} = \{\alpha_{\gamma_{m1}}, \alpha_{\gamma_{m2}}, \dots, \alpha_{\gamma_{mt_m}}\}, \quad m = 1, \dots, M. \quad (12)$$

Note that each candidate set contains all the possible label features which can be assigned to the corresponding reference feature in decreasing order of compatibility.

Once all the candidate sets $\Lambda_{a_1}, \Lambda_{a_2}, \dots, \Lambda_{a_M}$ are obtained, the candidate models can be constructed by collecting one label feature from each candidate set:

$$\mathcal{A} = \{(\tilde{\alpha}_1, \tilde{\alpha}_2, \dots, \tilde{\alpha}_M) \mid \tilde{\alpha}_1 \in \Lambda_{a_1}, \tilde{\alpha}_2 \in \Lambda_{a_2}, \dots, \tilde{\alpha}_M \in \Lambda_{a_M}\}. \quad (13)$$

It is worth noting that the number and order of the label features in the candidate models are

consistent with those of the reference features, so that the feature correspondence between the reference model and a candidate model is readily established.

3. EIGENVECTOR-BASED MATCHING TECHNIQUE

Once candidate models are selected from the label image, we then compare these models to the reference model. The binary relations used in the previous section provide only the information on the local relations between two features. Although the sum of $\mathcal{S}(a_m, a_n; \alpha_k, \alpha_l)$ between two models may indicate whether the two models are similar or not, it is not sufficient for a measure of global shape matching, since the entire relations are not taken into account. Thus, in our approach, we employ an eigenvector approach to evaluate the global similarity between the reference model and the candidate model.

3.1. Modal Representation of a Line Model

The main notion behind the modal analysis is to transform the features into a hyperspace of generalized axes of symmetry or modes of a shape, so that the feature-mode relations can be readily examined. To gain an insight into the proposed approach, let us consider a method for forming the modes of a single model with M features $x_i, i = 1, 2, \dots, M$. For each binary relation r_k , a square matrix of size $M \times M$ \mathbf{P}_k , called a *metric matrix*, is constructed to describe the interaction between the features *within* the model, whose element is given by

$$[\mathbf{P}_k]_{ij} = r_k(x_i, x_j). \quad (14)$$

Then, a square *proximity matrix* $\mathbf{H}_{M \times M}$ is obtained from the sum of Gaussian-weighted exponential of (14), given by

$$[\mathbf{H}]_{ij} = \exp \left\{ - \sum_{k=1}^{Nr} [\mathbf{P}_k]_{ij}^2 / \sigma_k \right\}, \quad (15)$$

where the parameter σ_k controls the weighting for each metric matrix as well as an intrainage interaction between the features for each binary relation. Note that if the proximity matrix \mathbf{H} is symmetric, then the eigenvalues are always real numbers and the corresponding eigenvectors form an orthogonal basis [16]. Thus, in order to make the proximity matrices symmetric, among the binary relations in (4), we use only r_1 and r_4 , since they satisfy the commutative property exclusively. Also, to make the effect of each measure on the entire interaction roughly equal, the mean values of the components of each metric matrix are adaptively used for σ_k as given by

$$\sigma_k = \frac{1}{M(M-1)} \sum_{i,j,i \neq j}^M [\mathbf{P}_k]_{ij} \quad \text{for } k = 1, 4. \quad (16)$$

Let us denote the i th eigenvector or *mode* of \mathbf{H} and the corresponding eigenvalue or *frequency* by \mathbf{e}_i and λ_i , respectively; then the following relation holds:

$$\mathbf{H}\mathbf{e}_i = \lambda_i \mathbf{e}_i, \quad i = 1, \dots, M. \quad (17)$$

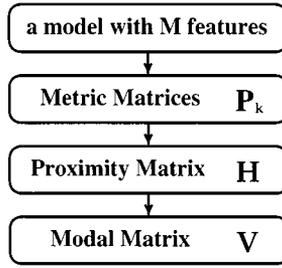


FIG. 5. Procedure for the computation of a modal matrix.

In matrix form, we have

$$\mathbf{H} = \mathbf{VDV}^T, \quad (18)$$

where \mathbf{V} is the orthogonal *modal matrix* whose columns are the eigenvectors, and \mathbf{D} is a diagonal matrix whose elements are the eigenvalues in decreasing order, respectively:

$$\mathbf{V} = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_M],$$

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_M \end{bmatrix}, \quad \text{where } \lambda_i \geq \lambda_{i+1}. \quad (19)$$

The modes provide the frequency ordered canonical description for the model and its natural deformation. In general, low-frequency modes describe global deformation, while higher frequency modes describe more localized deformation. The procedure for computing the modal matrix is summarized in Fig. 5.

3.2. Sign Correction for the Modes

Unlike Shapiro and Brady's work on point feature correspondence in modal space [9], we employ the modal approach to evaluate the similarity between the line models for matching or registration. However, one significant problem relating to the modal analysis for matching is that the shape of modes is not invariant to the ordering of the features, implying that, if the labeling order of model features is changed, the modes are also altered substantially. Therefore, unless the orderings of corresponding features for two models are consistent, a direct comparison of the modes leads to incorrect results. However, note that in our approach, as discussed in the previous section, in constructing the candidate sets, the label ordering problem as well as the dimension problem is readily resolved, since each element of the candidate set Λ_{a_m} already corresponds to the reference feature a_m .

Given the reference model \mathcal{Q}_r and a candidate model $\mathcal{A} \in \mathcal{Q}_c$, we obtain the proximity matrices \mathbf{H} and their modal matrices \mathbf{V} , given by

$$\mathbf{H}_r = \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^T$$

$$\mathbf{H}_c = \mathbf{V}_c \mathbf{D}_c \mathbf{V}_c^T. \quad (20)$$

Then, the matching can be directly carried out by simply comparing the corresponding mode vectors of \mathbf{V}_r and \mathbf{V}_c for two sets of features.

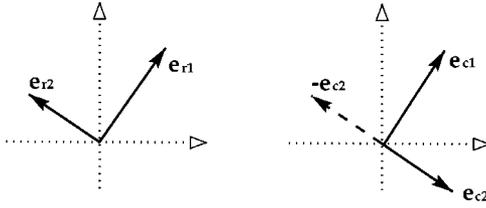


FIG. 6. A simple sign correction example.

However, there is still another problem, called the sign correction problem [9], in comparing the modes for matching. It is noted that the sign of eigenvectors which satisfy (17) is not unique; i.e., if \mathbf{e}_i is an eigenvector, then $-\mathbf{e}_i$ also can be an eigenvector. Therefore, without sign correction or alignment of the corresponding modes for the two models, a direct comparison of \mathbf{e}_{r_i} and \mathbf{e}_{c_i} is also inappropriate. Thus, before the modes for models are compared, the sign correction stage is necessary to make both sets of modes have consistent directions [9]. Figure 6 shows a typical example of a reference and a candidate model to be matched with two eigenvectors. Assume these two models are the same; however, one of the candidate modes \mathbf{e}_{c_2} is opposite to \mathbf{e}_{r_2} . In this case, unless the sign of \mathbf{e}_{r_2} is corrected to $-\mathbf{e}_{c_2}$, a simple comparison of corresponding modes implies the two models are different.

In this paper, we use the following method to correct the sign of the corresponding modes. Let us consider \mathbf{V}_r as the reference modal matrix and let the modes of \mathbf{V}_c be aligned to those of \mathbf{V}_r one by one. First, let us assume that all the eigenvalues of \mathbf{V}_r and \mathbf{V}_c are distinct. Then, since the two sets of modes are already correspondent, we have only to compare each eigenvector \mathbf{e}_{c_i} with its counterpart \mathbf{e}_{r_i} and to correct the sign of \mathbf{e}_{c_i} so that

$$\hat{\mathbf{e}}_{c_i} = \left\{ \begin{array}{ll} \mathbf{e}_{c_i} & \text{if } \|\mathbf{e}_{r_i} + \mathbf{e}_{c_i}\| > \|\mathbf{e}_{r_i} - \mathbf{e}_{c_i}\| \\ -\mathbf{e}_{c_i} & \text{otherwise} \end{array} \right\} \quad \text{for } 1 \leq i \leq M. \quad (21)$$

3.3. Matching by Modal Comparison

Since the structural description for the line model is embedded in the modes of its proximity matrix, once the modal descriptions are established, we can compare the line models easily by evaluating the similarity between their modes. Once the direction of corresponding modes is aligned for \mathbf{V}_r and \mathbf{V}_c , the similarity between the reference model and the candidate model can be measured by computing the affinity between their corresponding modes.

Although insignificant modes are discarded to alleviate the effect of the input noise in most conventional eigenvector-based approaches [17, 18], all the retained modes are involved equally for the evaluation of the similarity measure. However, for robust matching, it is argued that more weights should be given to low-order modes, since the low-order modes represent the global shape information and are relatively immune to the input noise or local variations. Thus, in our approach, the Euclidean distance between eigenvectors is weighted by the corresponding eigenvalues, so that the global shape information could be a dominant factor for matching. Thus,

$$\mathcal{D}(\mathbf{V}_r; \mathbf{V}_c) = \sum_{j=1}^M \|\lambda_{r_j} \mathbf{e}_{r_j} - \lambda_{c_j} \hat{\mathbf{e}}_{c_j}\|^2. \quad (22)$$

Notice $\mathcal{D}(\mathbf{V}_r; \mathbf{V}_c)$ is a quantitative measure for evaluating the similarity between the two models. $\mathcal{D}(\mathbf{V}_r; \mathbf{V}_c)$ is zero for perfect matching, and it increases as the difference between the reference and the candidate models increases. Therefore, we refer to $\mathcal{D}(\mathbf{V}_r; \mathbf{V}_c)$ as the *dissimilarity measure*.

Although it is very rare in practice, if duplicated eigenvalues exist in either model, a direct comparison of individual corresponding eigenvectors is meaningless, since the orthonormal eigenvectors for duplicated eigenvalues cannot be uniquely determined, but only the subspace spanned by these vectors can be. Therefore in this case, comparison of corresponding subspaces of the reference model and the candidate model along with associated eigenvalues is more appropriate. Note, however, that since all eigenvectors of a model are orthonormal, the orientational information (or the coordinates) of the subspace corresponding to the duplicated eigenvalues is embedded implicitly in the remaining eigenvectors. Therefore, for simplicity, without considering explicit orientational information of the subspaces, we may use the difference of eigenvalues associated with these subspaces for the dissimilarity measure between them. That is, if we assume that there exist s duplicated eigenvalues for a candidate model, $\lambda_{cp} = \lambda_{c(p+1)} = \dots = \lambda_{c(p+s-1)}$, then the dissimilarity terms corresponding to $j = p, \dots, p + s - 1$ in (22) are replaced by

$$\sum_{j=p}^{p+s-1} |\lambda_{rj} - \lambda_{cj}|^2. \quad (23)$$

4. SIMULATION RESULTS

In this section, we demonstrate the performance of the proposed algorithm on several synthetic and real imagers. Figures 7a and 7b show a synthetic reference model of a Korean character and a synthetic label image, respectively. Figure 7c is a scaled version of the reference model, and Figs. 7d–7j illustrate the candidate models extracted from Fig. 7b, using the method discussed in Section 2.2. It is noted that the candidate model in Fig. 7d is a rotated version of the reference model, while the others are deformed versions. One might notice that it is very difficult to identify the correct object from Fig. 7b, even for a human being. In Fig. 8, the dissimilarity measures \mathcal{D} for these candidate models are presented. In order to demonstrate the performance more clearly, a simple similarity measure, based on the sum of compatibility measures of binary relations, $\sum \mathcal{S}(a_m, a_n; \alpha_k, \alpha_l)$, is also shown for each candidate model. It is seen that the similarity measure yields the maximum $M - 1 = 5$ for perfect matching.

From Fig. 8, we see that the dissimilarity measures for the candidate models in Figs. 7c and 7d are zero, implying that the proposed algorithm is perfectly invariant to the translation, in-plane rotation and the scale changes. By comparing \mathcal{D} and $\sum \mathcal{S}$ for other candidate models, it is concluded that the eigenvector approach is very effective in quantitatively evaluating the similarity between the reference and candidate models, compared to the case when only the binary compatibility measures are employed. This observation is due to the fact that since the modal matrix incorporates all the mutual relations between feature elements implicitly, if several label features are quite different from the corresponding reference features, then all the modes are changed, eventually making the \mathcal{D} large. And, since $\sum \mathcal{S}$ employs only the information on the local relations independently, not the information on global relations, the total sum of \mathcal{S} does not change greatly, even if one or two labels become eccentric.

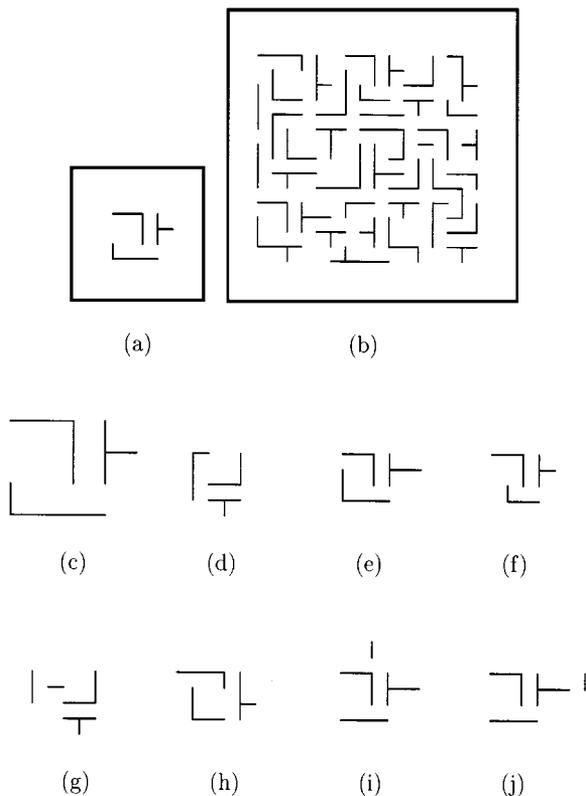


FIG. 7. Synthetic matching test: (a) reference model; (b) label image; (c) scaled version of reference model; (d)–(j) extracted candidate models from (b).

Figure 9 shows a more complex matching example on real aerial images. Figures 9a and 9b are aerial images of manmade structures acquired from different viewpoints. Figure 9c shows the reference model extracted from Fig. 9a, and Fig. 9d shows the label image extracted from Fig. 9b, respectively. We have used the Nevatia–Babu edge detector [11]

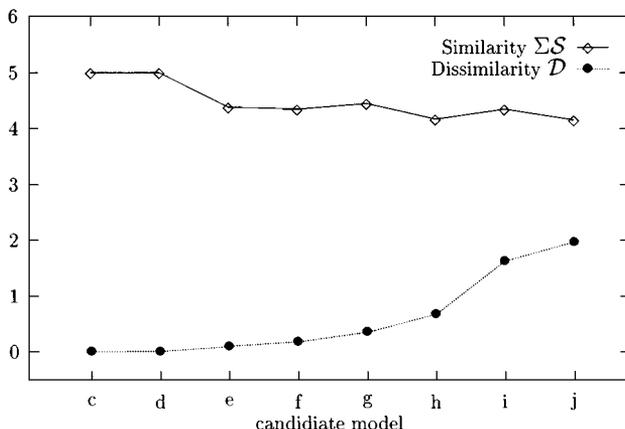


FIG. 8. The similarity $\sum S$ and the dissimilarity \mathcal{D} for each candidate model in Figs. 7c–7j.

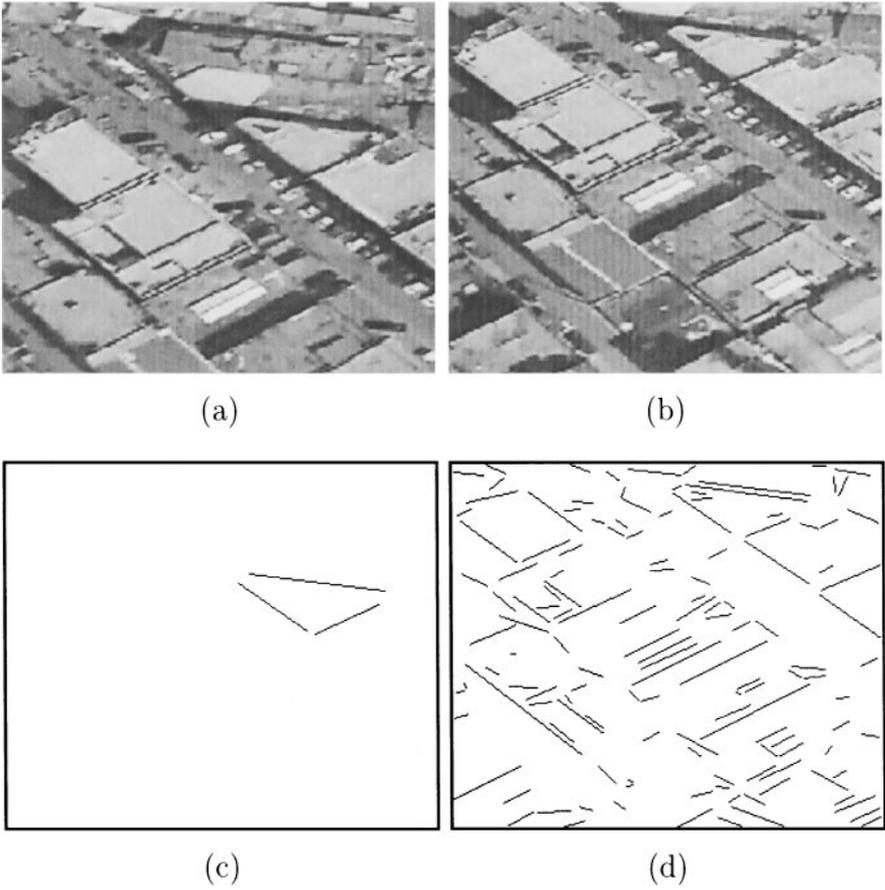


FIG. 9. Aerial image test I: (a)–(b) two aerial images of different view; (c) the reference model obtained from (a); (d) input line features extracted from (b).

followed by an edge-linking process for line feature extraction, and no further high-level postprocessing is employed. In some real scenes, due to poor contrast or near other strong features, many lines are fragmented and displaced from the actual straight line, making simple colinearization useless, as can be seen in Fig. 9d. The four strongest candidate models selected from Fig. 9d are shown in Fig. 10, together with the corresponding dissimilarity values. The correct matching is achieved with Fig. 10a, which yields the lowest dissimilarity measure.

More experimental results on real aerial images are illustrated in Fig. 11. Figure 11a is the reference aerial image and Fig. 11b shows the reference model extracted from it. Two test input images with different viewpoints are shown in Figs. 11c and 11d, and the label images corresponding to them are depicted in Figs. 11e and 11f, respectively, which are composed of more than 60 line segments. The final matching results are illustrated in Figs. 11g and 11h, which demonstrate that although the extracted line features are not satisfactory, the proposed algorithm works quite successfully.

Figure 12 shows the matching results for a sequence of images. Figures 12a–12d show a sequence of input images from a camera with axial and lateral motion. Figures 12e–12h show the label images extracted from Figs. 12a–12d, respectively. We extract a rectangle

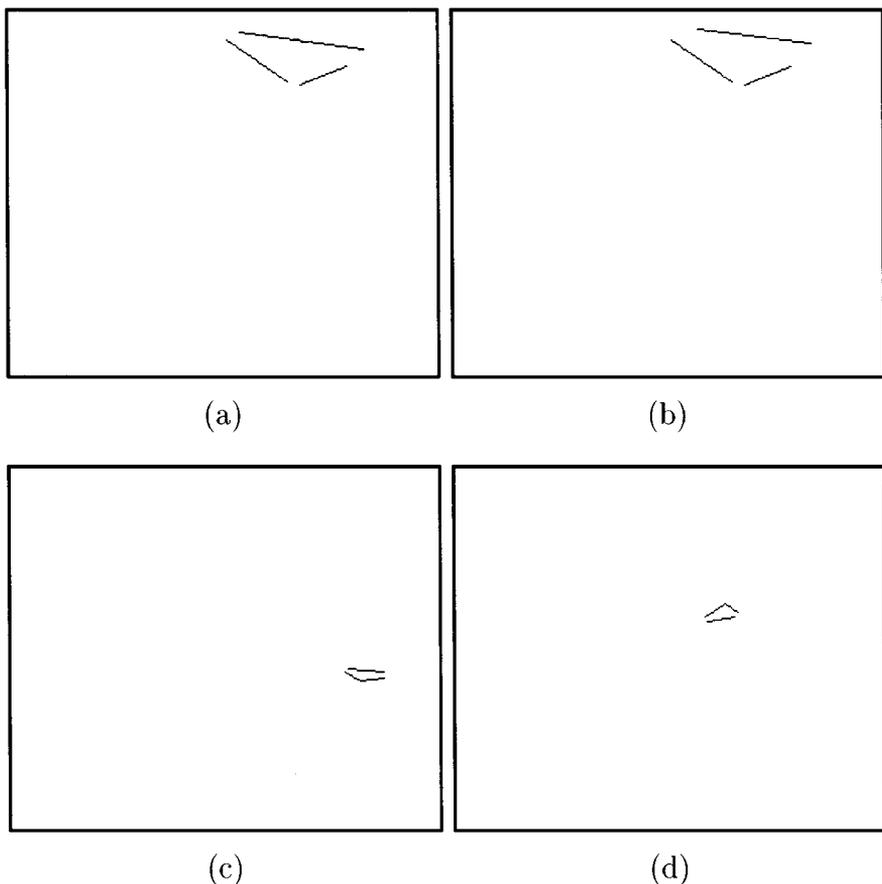


FIG. 10. Four strongest candidate models extracted from Fig. 4d and the corresponding dissimilarity values \mathcal{D} : (a) 0.0011; (b) 0.0163; (c) 0.1631; (d) 0.1678.

with four line segments and a pentagon, made up of five line segments from Fig. 12a, and use them for the reference models as shown in Figs. 12i and 12m, respectively. The final matching results for each model are shown in Figs. 12j–12l and Figs. 12n–12p, respectively.

Note that although the proposed matching algorithm is only invariant to in-plane rotation, since the binary relations r_1 and r_4 used for the construction of the proximity matrix are so, when the deformation caused by out-of-plane rotation is relatively small, it still can find the desired object successfully as the test results demonstrate.

Figure 13 shows the test results for the case where some line features in a label image are broken. Figures 13a and 13b are the two image sequences for the reference and test images, and the reference model and the label features extracted from these images are shown in Figs. 13c and 13d, respectively. The best and the second best matching results are represented in Figs. 13e and 13f, respectively. Note that, in general, the proposed algorithm may not work properly in the presence of severely broken lines, especially when the number of model features is small. However, we observe that for the case when the distortions are not severe, so that a large part of each line is still present in the scene, or when the number of features of a model is large enough so that most of the significant eigenmodes are not affected by few broken lines, the proposed algorithm works fairly robustly.

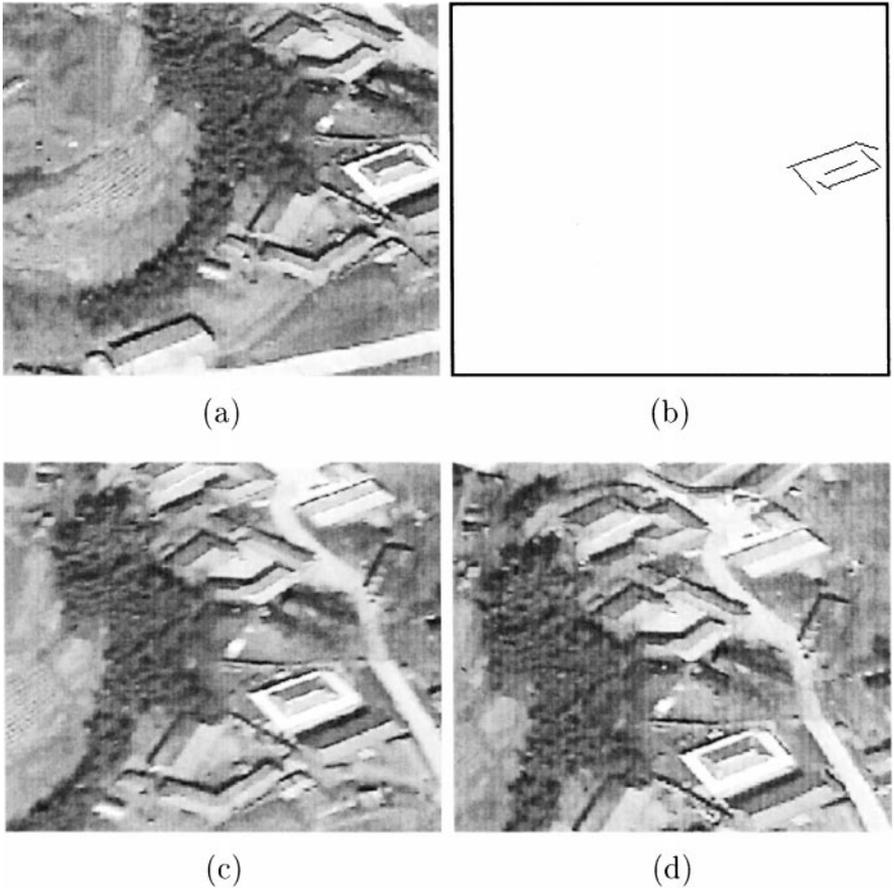


FIG. 11. Aerial image test II: (a) reference image; (b) reference model; (c)–(d) input images; (e)–(f) label images; (g)–(h) matching results.

We also compare the performance of the proposed algorithm with other methods. The overall computation times for several matching algorithms are given in Eq. (24), where \mathcal{T}_{modal} , \mathcal{T}_{blind} , and \mathcal{T}_{relax} are the times required for the proposed modal approach, the blind tree search technique, and the relaxation operation, respectively, and τ_1 is the time for the extraction of candidate sets from the label image, τ_2 is the time for the evaluation of dissimilarity measure between the reference model and a candidate model, n_c is the number of candidate models, and τ_3 is the time for the convergence of a relaxation operation. Since \mathcal{T}_{modal} is dependent on n_c , it is concluded that the step for restricting the candidate models is essential to the application of the modal approach to the labeling problem. In contrast, if we compare all the combinations of M label features out of N with the reference features by a simple blind tree search technique, without using the candidate selection scheme, the total computation time \mathcal{T}_{blind} becomes proportional to the number of all combinations. Note that if the initial probability for the label features is not selected appropriately, usually the relaxation operation fails to converge. Therefore, in order to ensure the convergence, in this experiment, we use the candidate selection scheme discussed in Section 2.2 for obtaining

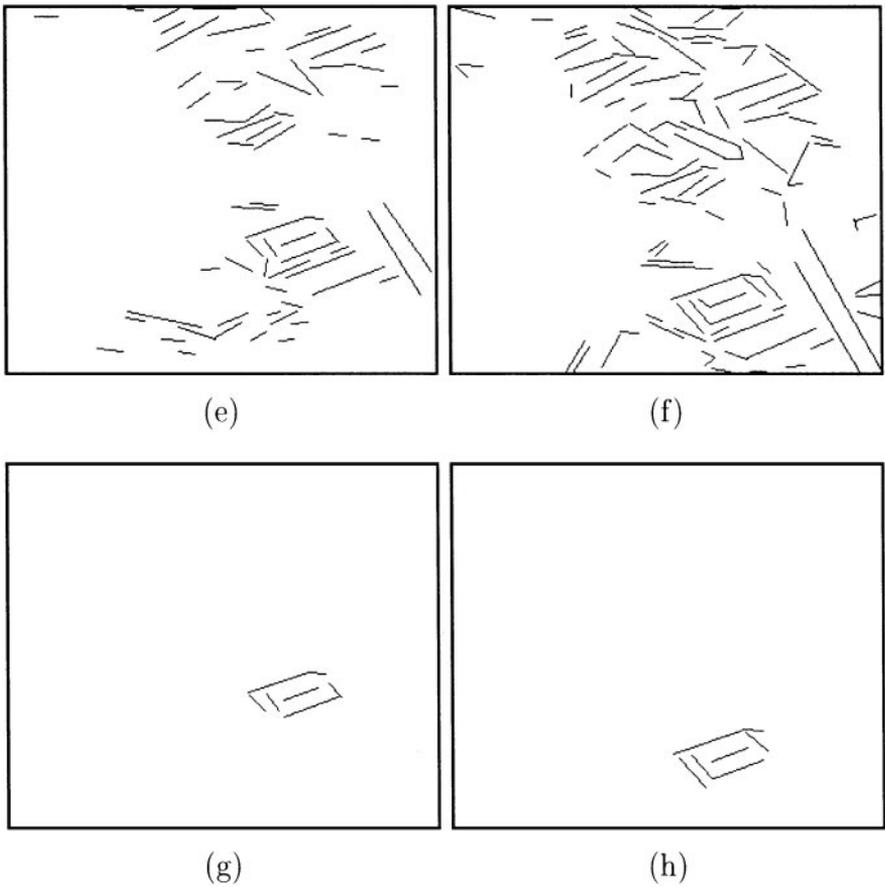


FIG. 11—Continued

the initial probability:

$$\begin{aligned}
 \mathcal{T}_{model} &= \tau_1 + n_c \cdot \tau_2 \\
 \mathcal{T}_{blind} &= N P_M \times \tau_2 \\
 \mathcal{T}_{relax} &= \tau_1 + \tau_3.
 \end{aligned}
 \tag{24}$$

For example, for the test of the third row in Fig. 12, where $M = 4$, $N = 73$, and $n_c = 10$, the elapsed times τ_1 and τ_3 are 205 and 172 ms, respectively, as given in Table 1. In this

TABLE 1
Comparison of Computation Times

	Case		
	I	II	III
Process	Candidate sets	Modal approach	Relaxation operation
Time (ms)	$\tau_1 = 205$	$\tau_2 = 172$	$\tau_3 = 18,027$

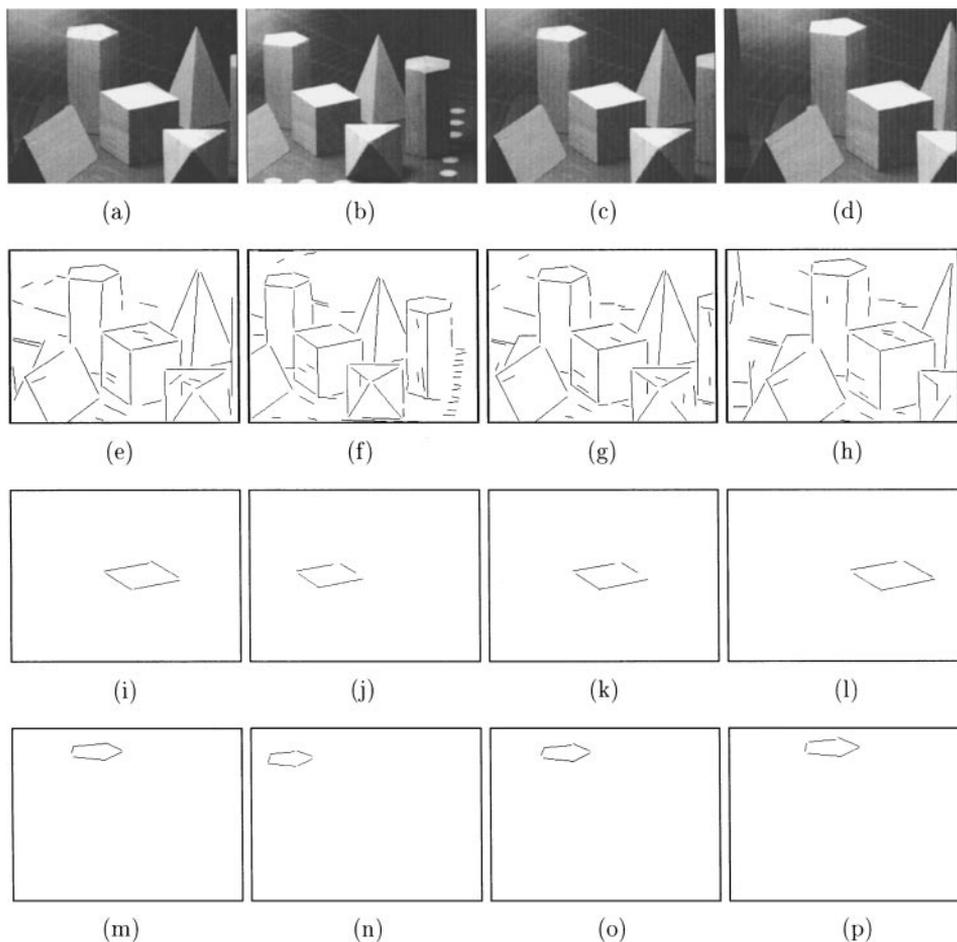


FIG. 12. Sequential image test: (a)–(d) input image sequence of blocks; (e)–(h) the line features of (a)–(d), respectively; (i) a rectangular reference model; (j)–(l) the matching results of (i) in (f)–(h), respectively; (m) a pentagon reference model; (n)–(p) the matching results of (m) in (f)–(h), respectively.

case, therefore, we have $\mathcal{T}_{model} = 1,925$ ms, while $\mathcal{T}_{blind} = 4.4$ million s. This shows clearly the computational benefit we can achieve by employing the proposed candidate selection scheme, in spite of the calculation of all the compatibility values. Also, compared to the computation time for the relaxation labeling $\mathcal{T}_{relax} = 18,232$ ms, the proposed algorithm is nearly 10 times faster.

5. CONCLUSIONS

In this paper, we proposed a line feature matching algorithm for target recognition, which is invariant to the translation, rotation, and scale changes. It was shown that the proposed algorithm performs matching of the line features in a fast, robust, and efficient way. As a result, the similarity between two feature sets was measured as the difference of the *modes*.

Although the performance of the modal approach is known to be sensitive to the number and order of the features, in our approach, the preliminary correspondence test resolved this

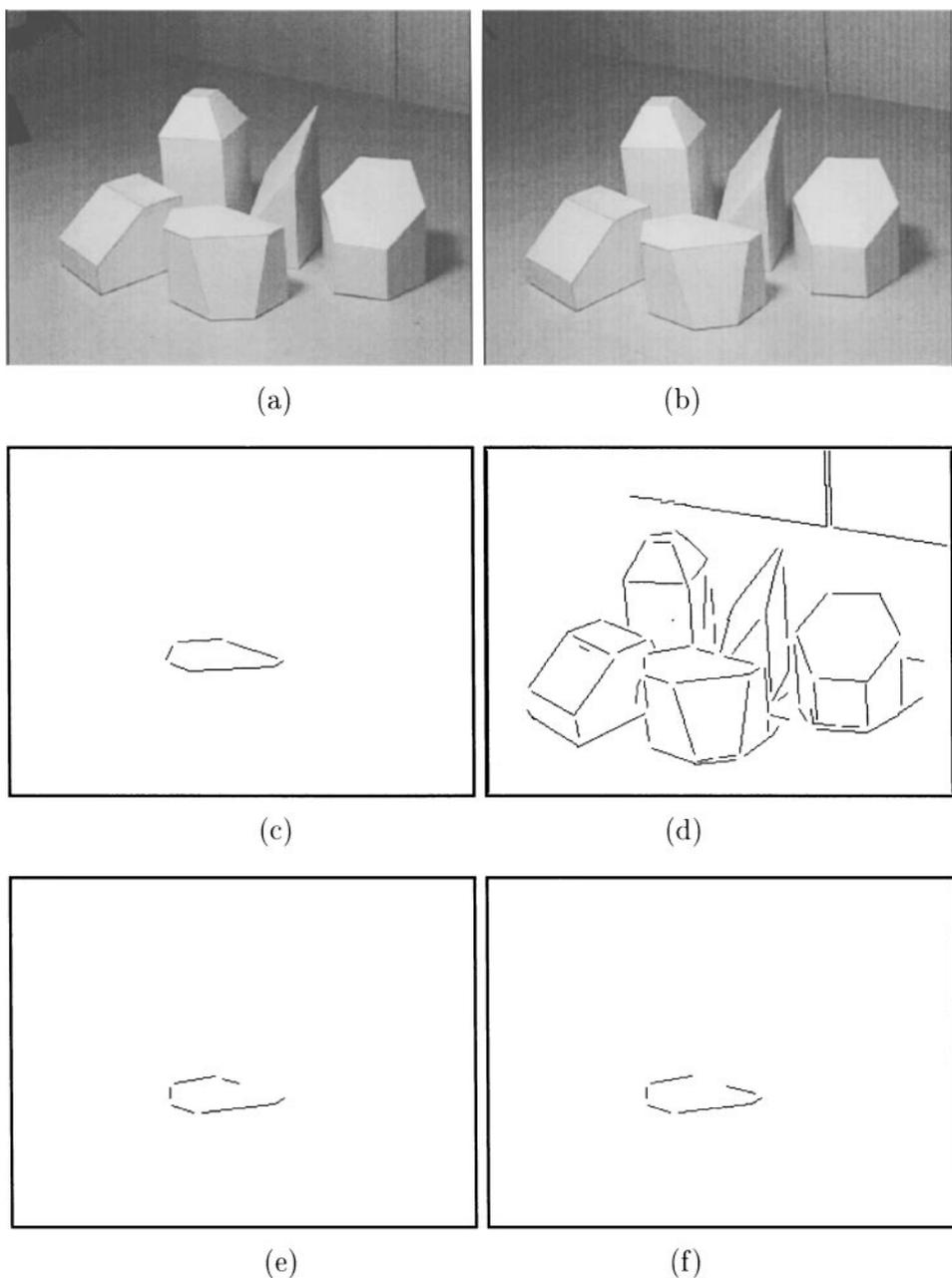


FIG. 13. Test for broken line features: (a)–(b) input images; (c) the reference model from (a); (d) the label features from (b); (e) the best matching result; (f) the second best matching.

problem. That is, a two-phase algorithm alleviates the computational load, while improving the overall performance.

The use of in-train image relations makes the proposed algorithm invariant to translation and rotation, and adaptive selection of Gaussian weight makes the proposed algorithm scale-invariant. Emphasis on important modes by their eigenvalues improves the robustness to small skews, local variation, and input noise.

Experimental results on synthetic imagery showed that the proposed algorithm performs matching of the line features on a global and quantitative basis, which bears a resemblance to human perception. The results on real images also showed that matching of the line features with some affine variation is carried out fast and efficiently, regardless of viewpoint changes.

In order to improve the performance and the flexibility of the proposed eigenvector-based matching technique, further research will be focused on the feasibility of partial matching in which the number of features of the target object is less than that of the reference model due to occlusion or noise.

APPENDIX: NOMENCLATURE

M	the number of features in model
N	the number of features in scene
\mathcal{Q}_r	the set of features in model
\mathcal{Q}_c	the set of features in scene
$\mathbf{r}(a_m, a_n)$	the set of binary relations in model
$\mathbf{r}(\alpha_r, \alpha_s)$	the set of binary relations in scene
$\mathcal{S}(a_m, a_n, \alpha_r, \alpha_s)$	compatibility measure between (a_m, a_n) and (α_r, α_s)
\mathbf{R}_{mn}	compatibility matrix
ψ_m	index vector
u_m	index number
ϕ_m	support vector
v_m	support number
Λ_{a_m}	candidate set for reference feature a_m
I_m	ordered index set
\mathcal{A}	candidate model
\mathbf{P}_k	k -th metric matrix
\mathbf{H}	proximity matrix
σ_k	k -th Gaussian weight
λ_i	i -th eigenvalue of proximity matrix
\mathbf{e}_i	i -th eigenvector of proximity matrix
\mathbf{V}	modal matrix of proximity matrix
\mathbf{D}	eigenvalue matrix of proximity matrix
$\mathcal{D}(\mathbf{V}_r; \mathbf{V}_c)$	dissimilarity measure between \mathbf{V}_r and \mathbf{V}_c .

REFERENCES

1. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.
2. R. J. Schalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches*, Wiley, New York, 1992.
3. S. Umeyama, An eigendecomposition approach to weighted graph matching problems, *IEEE Trans. Pattern Anal. Mach. Intell.* **10**(5), 1988, 696–703.
4. A. Rosenfeld, R. A. Hummel, and S.W. Zucker, Scene labeling by relaxation operations, *IEEE Trans. Systems Man Cybernet.* **6**(6), 1976, 420–453.
5. R. A. Hummel and S. W. Zucker, On the foundation of relaxation labeling process, *IEEE Trans. Pattern Anal. Mach. Intell.* **5**(3), 1983, 267–286.

6. G. Medioni and R. Nevatia, Matching images using linear features, *IEEE Trans. Pattern Anal. Mach. Intell.* **6**(6), 1984, 675–685.
7. S. Z. Li, Matching: Invariant to translations, rotations and scale changes, *Pattern Recog.* **25**(6), 1992, 583–594.
8. G. Scott and H. Longuet-Higgins, An algorithm for associating the features of two patterns, in *Proc. R. Soc. London B* **244**, 1991, 21–26.
9. L. S. Shapiro and J. M. Brady, Feature-based correspondence: An eigenvector approach, *Image Vision Comput.* **10**(5), 1992, 283–288.
10. M. Pilu, A direct method for stereo correspondence based on singular value decomposition, in *Proc. Computer Vision Pattern Recognition, 1997*, pp. 261–266.
11. R. Nevatia and K. R. Babu, Linear feature extraction and description, *Comput. Graphics Image Process.* **13**, 1980, 257–269.
12. S. Sclaroff and P. Pentland, Modal matching for correspondence and recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* **17**(6), 1995, 545–561.
13. K. Bathe, *Finite Element Procedures in Engineering Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1982.
14. J. Krumm, Eigenfeatures for planar pose measurement of partially occluded objects, in *Proc. Computer Vision Pattern Recognition, 1996*, pp. 55–60.
15. W. Grimson and T. Lozono-Perez, Localizing overlapping parts by searching the interpretation tree, *IEEE Trans. Pattern Anal. Mach. Intell.* **9**(4), 1987, 469–482.
16. G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed., Johns Hopkins Press, Baltimore, 1989.
17. M. Uenohara and T. Kanade, Use of Fourier and Karhunen–Loeve decomposition for fast pattern matching with a large set of templates, *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(8), 1997, 891–898.
18. K. Ohba and K. Ikeuchi, Detectability, uniqueness, and reliability of eigen windows for stable verification of partially occluded objects, *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(9), 1997, 1043–1048.