A Local Structural Descriptor for Image Matching via Normalized Graph Laplacian Embedding

Jun Tang, Ling Shao, *Senior Member, IEEE*, Xuelong Li, *Fellow, IEEE*, and Ke Lu

Abstract—**This paper investigates graph spectral approaches to the problem of point pattern matching. Specifically, we concentrate on the issue of how to effectively use graph spectral properties to characterize point patterns in the presence of positional jitter and outliers. A novel local spectral descriptor is proposed to represent the attribute domain of feature points. For a point in a given point-set, weight graphs are constructed on its neighboring points and then their normalized Laplacian matrices are computed. According to the known spectral radius of the normalized Laplacian matrix, the distribution of the eigenvalues of these normalized Laplacian matrices is summarized as a histogram to form a descriptor. The proposed spectral descriptor is finally combined with the approximate distance order for recovering correspondences between point-sets. Extensive experiments demonstrate the effectiveness of the proposed approach and its superiority to the existing   methods.**

Index Terms—**Graph spectrum, local structural descriptor, normalized Laplacian matrix, point pattern matching (PPM).**

I. INTRODUCTION

LOCAL features are popular and important in computer vision [1]. Therefore, finding correspondences between feature point-sets is critical for many practical applica- tions, such as 3-D reconstruction [2], object recognition [3], motion tracking [4], image retrieval [5]–[7], and   action recognition [8], [9].

Graph spectral methods for point pattern matching (PPM) have received much attention over the past two decades due largely to their prominent ability of capturing the essence of point-set structure. However, their sensitivity to   structural errors caused by positional jitter and outliers is still an open issue. In this paper, we aim to improve the robustness of spectral matching methods by providing a novel spectral representation for point patterns. Almost all the available spectral matching methods concentrate on using the eigenvectors of various matrix representations. For instance, in [10]–[13], the eigenvectors are employed as the global features of point patterns. In [14] and [15], the eigenvectors are used as the correspondence indicators. The numeric approach to eigen-decomposition is a least squares method, and therefore, positional jitter inevitably brings errors to the computation of eigenvectors. In order to overcome this weakness, some iterative optimization methods were proposed to resist such negative effects [12], [13]. Nonetheless, the presence of outliers makes the sizes of graphs different, which renders the comparison of the eigenvectors difficult. Although some researchers proposed to truncate the eigenvectors to make the length of the compared eigenvectors identical, this empirical operation may result in the ineffectiveness of the spectral representation when the difference is relatively large. Moreover, there are no satisfactory solutions available to determine which components of the eigenvectors to remove so far.

To address these issues, we attempt to localize the spectral representation of point patterns. Specifically, we use spectral properties of structural graphs constructed on its neighboring sub-point-sets to describe a feature point. The advantage of such a localized representation is that the influence of positional jitter is only brought by the neighboring sub-point-sets rather than the whole point-set. And a completely different idea is adopted to tackle the issue of comparing spectral properties with different sizes. We have observed that most of the available local descriptors such as shape context [16], SIFT [17], and SURF [18] use the distribution of some properties extracted from images to describe a feature point. Inspired by these works, we turn to exploiting the distribution of spectral properties to construct a spectral descriptor. The consequent difficulty is that the numeric boundary of most spectral properties is very hard to obtain, which makes the analysis of their distribution infeasible. Fortunately, thanks to the known spectral radius of the normalized Laplacian matrix, we can use the distribution of their eigenvalues to form a histogram-based descriptor. Meanwhile, as eigenvalues contain useful information for characterizing graph structures [19], [20], we believe that the proposed descriptor holds sufficient discriminative power.

J. Tang is with the Key Laboratory of Intelligent Computing and Signal Processing, Ministry of Education, Anhui University, Hefei 230039,   China.

L. Shao is with the Department of Computer Science and Digital Technologies, Northumbria University, Newcastle upon Tyne, NE1 8ST, U.K. (e-mail: ling.shao@ieee.org).

X. Li is with the Center for OPTical IMagery Analysis and Learning (OPTIMAL), the State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China.

K. Lu is with the University of Chinese Academy of Sciences, Beijing 100049, China, and also with the Beijing Center for Mathematics and Information Interdisciplinary Sciences, Beijing, China.

The rest of this paper is organized as follows. In Section II, we give a brief review of related works. In Section III, we present the construction procedure of the proposed spectral descriptor. In Section IV, the descriptor is used with the approximate distance order to solve PPM via the technique of probabilistic relaxation. We demonstrate the effectiveness of our approach with experiments in Section V, and this paper is concluded with some future work in Section VI.

II. RELATED WORK

A number of attempts have been made for applying spectral graph theory to PPM. Umeyama [21] proposed the pioneer work in spectral matching. They developed an approximate algorithm for graph matching by using eigen-decomposition. Scott and Longuet-Higgins [10] recov- ered correspondences between point-sets by performing sin- gular value decomposition on a weight point associate matrix. Pilu [22] improved [10] by including the neigh- boring intensity correlation into the computation of asso- ciate weights. In order to overcome the weakness of the method of Scott and Longuet-Higgins for large rota- tion and scaling, Shapiro and Brady [11] presented a match- ing algorithm by comparing the ordered eigenvectors of the intraimage point proximity matrices. However, this method is extremely sensitive to the size difference between point-sets. Fairly speaking, [10] and [11] established a foun- dation for the subsequent research on spectral matching. Carcassoni and Hancock [12] embedded Shapiro and Brady's method into the expectation maximization framework, where iterative correspondences and transformation estimation are used to refine the matching results delivered by Shapiro and Brady's method. Caracassoni and Hancock [23] also reported a hierarchical matching method by using spectral clusters. Delphonte et al. [24] extended [22] by incorporating the sim- ilarity of SIFT descriptors into the computation of the prox- imity matrix. Wang and Hancock [13] used the rationale of kernel principal component analysis (PCA) to theoretically analyze Shapiro and Brady's method. They also showed that label consistency constraints can be incorporated into the definition of the point proximity matrix. Silletti et al. [25] used a variety of similarity metrics to construct the point associate matrix. Their solution to recovering correspon- dences is built upon Scott and Longuet-Higgins' method. Leordeanu and Hebert [14] formulated point matching as a problem of integer quadratic programming. In order to avoid combinational search, the leading eigenvector of the affinity matrix is computed as the correspondence indicator by means of spectral relaxation. Qiu and Hancock [26] used the Fiedler vector of the Laplacian matrix to decompose graphs into smaller subgraphs to simplify the complexity of inexact graph matching. Cour et al. [15] extended [14] by enforcing an affine constraint. Pang et al. [27] treated each candidate correspon- dence in the affinity matrix as a voter and the matching results are efficiently computed by simple addition and ranking.

Recently, using graph spectral properties to describe an object has been of particular interest as it can provide a discriminative and compact summary of graph structure.

Luo et al. [19] used spectral properties of the adjacency matrices such as eigenvalues, Cheeger constant, and eigen- mode perimeter to embed symbolic relational graphs in a pattern space and applied them to the task of object classifica- tion. Wilson et al. [28] employed eigenvectors of Laplacian matrix to construct permutation invariants to characterize graph structures. Wilson and Zhu [29] investigated combining eigenvalues of various graph matrix representations to summa- rize a graph structure. Shokoufandeh et al. [20] explored an approach to using eigenvalues of the adjacency matrix of the directed acyclic graph to construct a topological signature for encoding hierarchical image structures. Xiao et al. [30], [31] studied the relationship between the heat kernel of graph and the Laplacian spectrum. The rationale of their work lies in that the heat kernel of graph can be computed by expo- nentiating the Laplacian eigensystem over time. In [30], they developed some handcrafted invariants by using heat kernel trace and zeta function. They also proposed to embed the individual nodes of graphs into a vector space by applying multidimensional scaling to the geodesic distances between nodes, where the geodesic distances are computed from the parametrix expansion of heat kernel [31]. In [32], we proposed to use spectral context, i.e., the eigenvalues of the adjacency matrices of local graphs, to describe point patterns. However, spectral context is not a descriptor in a strict sense, and the similarity has to be evaluated by padding zeros into the vectors constructed from the eigenvalues.

In this paper, we inherit the idea of localized spectral rep- resentation in [32], which can be used to effectively deal with the presence of outliers. In particular, we study how to use the spectra of the normalized Laplacian matrix to character- ize point patterns. On the one hand, the Laplacian spectra are much more natural and important than the adjacency spec- tra and contain more information [33]. And the normalized Laplacian spectra perform the best with respect to cospec- trality according to the empirical study in [29]. On the other hand, the known spectral radius of the normalized Laplacian matrix is leveraged to construct a fixed size feature vector, which results in a well-formed descriptor and is beneficial for the potential research on feature learning [34]–[36].

III. LOCAL SPECTRAL DESCRIPTOR

A. Preliminary

Let $\mathbf{G} = (\mathbf{V}, \mathbf{E}; \mathbf{W})$ be a structural graph where $\mathbf{V} = \{v_1, v_2, \ldots, v_n\}$ is the node set and $\mathbf{E} \subseteq \mathbf{V} \times \mathbf{V}$ is the edge set and $\mathbf{W}$ is the weight function defined on $\mathbf{E}$. The adjacency matrix $\mathbf{A}$ of the graph is denoted by

$$\mathbf{A}(v_k, v_l) = \begin{cases} \mathbf{W}(v_k, v_l) & \text{if } (v_k, v_l) \in \mathbf{E} \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Let $\mathbf{D} = \text{diag}\{d_1, d_2, \ldots, d_n\}$ be the diagonal degree matrix of Graph $\mathbf{G}$, where $d_k = \sum_{(v_k, v_l) \in \mathbf{E}} \mathbf{W}(v_k, v_l)$. And the Laplacian matrix is defined as

$$\mathbf{L} = \mathbf{D} - \mathbf{A} \quad (2)$$

which is well-known as a key part in manifold learning [37], [38]. The corresponding normalized Laplacian
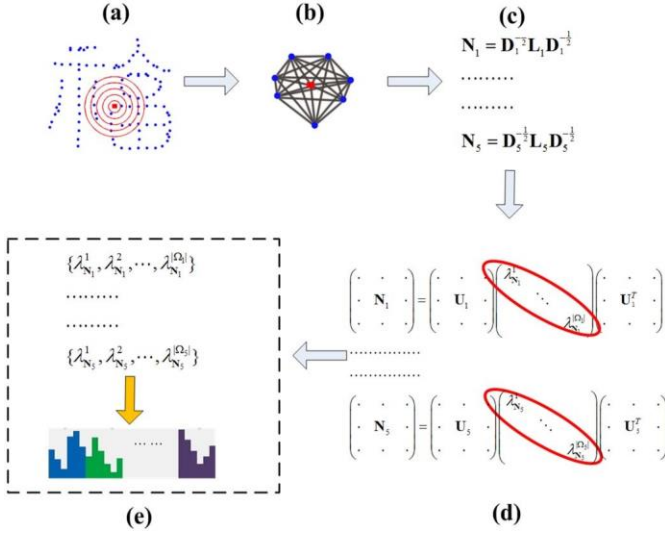
Fig. 1. Pipeline of descriptor construction. (a) Step 1: choose sub-point-sets. (b) Step 2: construct weight graphs on the sub-point-sets. (c) Step 3: compute the normalized Laplacian matrices. (d) Step 4: perform singular value decomposition (SVD) on the normalized Laplacian matrices. (e) Step 5: summarize the eigenvalues as a histogram-based descriptor.

matrix is denoted by

$$
\mathbf{N}(v_k, v_l) = \begin{cases} 1 & \text{if } k = l \\ -\dfrac{\mathbf{W}(v_k, v_l)}{\sqrt{d_k d_l}} & \text{if } (v_k, v_l) \in \mathbf{E} \\ 0 & \text{otherwise}. \end{cases} \quad (3)
$$

We can also write it as $\mathbf{N} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$. As the Laplacian matrix is positive semidefinite, all its eigenvalues are greater than or equal to zero. Its normalization form means that the largest eigenvalue is less than or equal to 2, with equality only when graph $\mathbf{G}$ is bipartite [39]. That is, all the eigenvalues of the normalized Laplacian matrix are in the range of [0, 2].

B. Descriptor Construction

Suppose that a point-set $\mathbf{X}$ is composed of M points, $\mathbf{X} = \{x_1, x_2,..., x_M\}$, and a related point-set $\mathbf{Y}$ contains N points, $\mathbf{Y} = \{y_1, y_2,..., y_N\}$.

Fig. 1 illustrates the pipeline of the construction procedure, including five steps such as sub-point-sets selection and graph construction. As we can see, the spectral descriptor is a local structural descriptor somewhat like shape context [16]. We consequently follow some settings of shape context, for example, five rings for radius. Without loss of generality, we use the notation set of point set $\mathbf{X}$ to present the details. We first compute the average closest distance d_X of point set $\mathbf{X}$ and define a set $\mathbf{R}_X = \{r_a | r_a = a \cdot d_X, a = 1, 2,..., 5\}$. Formally, d_X is computed by

$$
d_X = \sum_M \min \|x_i - x_{i'}\| \,|i' \neq i \, / M. \quad (4)
$$

To construct a histogram-based descriptor, we then uniformly divide the range of [0, 2] into K intervals. K is set to 200 throughout this paper. For a given point $x_i \in \mathbf{X}$, its spectral descriptor is computed as follows.

Step 1: For each $r_a \in \mathbf{R}_X$, choose the sub-point-set K_{ia} where the distances between these points and x_i are less than r_a.

Step 2: Construct a weight graph $\mathbf{G}_{ia}$ on the sub-point-set K_{ia}, where the edge weight is defined as

$$
\mathbf{W}(k, l) = \exp\left(-\frac{\|x_k - x_l\|^2}{2\beta_X^2}\right) \quad (5)
$$

where β_X is a smoothing coefficient. According to our empirical study, β_X is not a sensitive parameter, and the value in the range of $[d_X, 2d_X]$ is applicable. We set $\beta_X = 2d_X$ throughout this paper.

Step 3: Define the adjacency matrix $\mathbf{A}_{ia}$ and the degree diagonal matrix $\mathbf{D}_{ia}$ of graph $\mathbf{G}_{ia}$, and then compute its Laplacian matrix $\mathbf{L}_{ia}$

$$
\mathbf{L}_{ia} = \mathbf{D}_{ia} - \mathbf{A}_{ia}. \quad (6)
$$

The corresponding normalized form is denoted by $\mathbf{N}_{ia} = \mathbf{D}_{ia}^{-1/2}\mathbf{L}_{ia}\mathbf{D}_{ia}^{-1/2}$.

Step 4: Perform SVD on $\mathbf{N}_{ia}$ and we have

$$
\mathbf{N}_{ia} = \mathbf{U}_{ia}X_{ia}\mathbf{U}_{ia}^T \quad (7)
$$

where $X_{ia} = \text{diag}(\lambda_{ia}^1, \lambda_{ia}^2, \ldots, \lambda_{ia}^{|K_{ia}|})$ and $\mathbf{U}_{ia} = (\mathbf{u}_{ia}^1, \mathbf{u}_{ia}^2, \ldots, \mathbf{u}_{ia}^{|K_{ia}|})$ is composed of the eigenvectors of $\mathbf{N}_{ia}$. The diagonal elements of X_{ia} are the eigenvalues of $\mathbf{N}_{ia}$.

Step 5: Compute the frequencies of falling into each pre-defined interval in the range of [0, 2] for all the obtained eigenvalues.

We finally organize the obtained frequencies as a coarse histogram, which is defined to be the spectral descriptor of point x_i. Besides, it has some appealing attributes with respect to geometric transformation invariance.

1) Spectral Descriptor is Invariant to Translation: The edge weights of the graph are defined on Euclidean distances, and hence the descriptor is invariant to translation.

2) Spectral Descriptor is Invariant to Rotation: Obviously, the histogram representation is rotation invariant. As the chosen sub-point-set is rotation invariant, we only need to prove the permutation invariance of the eigenvalues. Let $\mathbf{N}^r = \copyright \mathbf{N} \copyright^T$ be a normalized Laplacian matrix constructed on the rearranged sub-point-sets with the permutation matrix $\copyright$. Substitute $\mathbf{N} = \mathbf{U}X\mathbf{U}^T$ into $\mathbf{N}^r = \copyright \mathbf{N} \copyright^T$, we have $\mathbf{N}^r = \copyright \mathbf{U}X\mathbf{U}^T\copyright^T = (\copyright\mathbf{U})X(\copyright\mathbf{U})^T$. Since $\mathbf{N}^r$ is a real symmetric matrix, its result of SVD is unique and the diagonal values of X are the singular values of $\mathbf{N}^r$. So the spectral descriptor is rotation invariant.

3) Spectral Descriptor is Invariant to Scaling: The chosen sub-point-set is scale invariant. And hence, we can make the descriptor invariant to scaling simply by tuning the smoothing coefficient in (5). For instance, we can set $\beta_Y = \beta_X d_Y / d_X$.

In step 3, the Gaussian weighting function is used to define edge weights. In [12], several alternative weighting functions are suggested to model the proximity between graph nodes. In the experimental section, we will evaluate these alternative choices for defining edge weights.
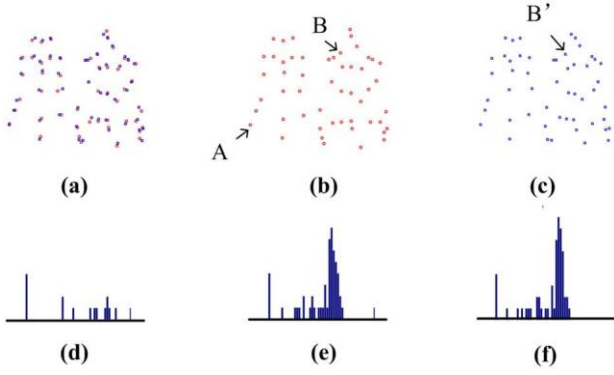
Fig. 2. Examples of the spectral descriptor on two toy point-sets. (a) Two toy point-sets. (b) Points A and B in one point-set. (c) Point B^r corresponding to point B in the other point-set. (d) Spectral descriptor for point A. (e) Spectral descriptor for point B. (f) Spectral descriptor for point B^r.

To make it clearer, we provide an example on two toy point-sets to demonstrate the discriminative power of the spectral descriptor in Fig. 2, where one point-set is produced by adding positional errors to the other. Note that the histograms for the matching points, i.e., points B and B^r, are of similar distribution. By contrast, the histogram for point A is quite different.

As the descriptors are distributions represented by histograms, the matching cost $C(x_i, y_j)$ between $x_i \in \mathbf{X}$ and $y_j \in \mathbf{Y}$ can be evaluated by the χ^2 test statistic

$$ C(x_i, y_j) = \frac{1}{2} \sum_{k=1}^{K} \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)} \tag{8} $$

where $h_i(k)$ and $h_j(k)$ denote the K-bin normalized histograms at x_i and y_j, respectively.

IV. APPLYING THE SPECTRAL DESCRIPTOR TO PPM

As there may exist many similar local structures in a point-set, it is unlikely to achieve an acceptable matching result if a local structural descriptor is individually used for solving PPM. Consequently, spatial consistency [40] is often combined with local descriptors to cope with this issue. In this paper, we leverage a pairwise constraint, the approximate distance order [32], to refine the results produced by the descriptor. The approximate distance order can be viewed as a generalization of the binary representation of neighboring relationship in [41], whose improvement over its original one has been validated in [32].

In order to tackle outliers, the dummy point, nil, is introduced. The point sets $\mathbf{X}$ and $\mathbf{Y}$ are expanded as $\{x_1, x_2,..., x_M, \text{nil}\}$ and $\{y_1, y_2,..., y_N, \text{nil}\}$, respectively. Let $\varphi : \mathbf{X} \rightarrow \mathbf{Y}$ denote the correspondences between point sets $\mathbf{X}$ and $\mathbf{Y}$, where the correspondences between inliers are one-to-one, but multiple outliers can be matched to a dummy point. The problem of establishing optimal matching is formulated as the maximization of an objective function with respect to the correspondences. The objective function is defined as the sum of the unary objective function $F_a(\mathbf{X}, \mathbf{Y}, \varphi)$ and the binary objective function $F_g(\mathbf{X}, \mathbf{Y}, \varphi)$

$$ F(\mathbf{X}, \mathbf{Y}, \varphi) = F_a(\mathbf{X}, \mathbf{Y}, \varphi) + \alpha F_g(\mathbf{X}, \mathbf{Y}, \varphi) \tag{9} $$

where the coefficient α weighs $F_g(\mathbf{X}, \mathbf{Y}, \varphi)$ and is set to 0.25 throughout this paper. $F_a(\mathbf{X}, \mathbf{Y}, \varphi)$ is defined to evaluate the feature similarities computed from the spectral descriptor between the matched point sets. The point-wise similarity is defined as

$$ \eta_{i,\varphi(i)} \equiv \eta(x_i, \varphi(x_i)) $$
$$ = \begin{cases} 0 & \varphi(x_i) = \text{nil} \\ \exp\frac{-C(x_i, \varphi(x_i))}{2\delta^2} & \text{otherwise} \end{cases} \tag{10} $$

where δ is a smoothing coefficient. δ is not a sensitive parameter and we empirically set $\delta = 1$ throughout this paper. Accordingly $F_a(\mathbf{X}, \mathbf{Y}, \varphi)$ can be given by

$$ F_a(\mathbf{X}, \mathbf{Y}, \varphi) = \sum_{i=1}^{M} \eta_{i,\varphi(i)}. \tag{11} $$

Without loss of generality, we use the notation set of point set $\mathbf{X}$ to introduce the definition of the approximate distance order. The approximate distance order of point $x_i \in \mathbf{X}$ with respect to point $x_{i^r} \in \mathbf{X} \setminus x_i$ is denoted by

$$ S(x_i, x^r) \equiv S(i, i^r) = \left\lceil \frac{|x_i - x_i^r|}{d_X} \right\rceil. \tag{12} $$

We begin with introducing the pairwise error on the approximate distance order to define $F_g(\mathbf{X}, \mathbf{Y}, \varphi)$. Let $\xi(i, i^r, j, j^r)$ denote the order difference between point pairs (x_i, x_i^r) and (y_j, y_j^r), then we have

$$ e_g(\mathbf{X}, \mathbf{Y}, \varphi) = \sum_{i=1}^{M} \sum_{i^r=1}^{M} \xi(i, i^r, \varphi(i), \varphi(i^r)) $$
$$ + \sum_{j=1}^{N} \sum_{j^r=1}^{N} \xi(j, j^r, \varphi^{-1}(j), \varphi^{-1}(j^r)). \tag{13} $$

In order to deal with the case that a point may match to a dummy point, the definition of the approximate distance order is augmented to

$$ S(i, \text{nil}) = S(\text{nil}, i) = S(\text{nil}, \text{nil}) = \infty. \tag{14} $$

Considering that the approximate distance order does not change significantly for neighboring points, we define $\xi(i, i^r, \varphi(i), \varphi(i^r))$ as follows:

$$ \xi(i, i^r, \varphi(i), \varphi(i^r)) = \begin{cases} |S(i, i^r) - S(\varphi(i), \varphi(i^r))|^2 & \text{if } S(i, i^r) \le T \\ & \text{and } S(\varphi(i), \varphi(i^r)) \le T \\ \infty & \text{otherwise} \end{cases} \tag{15} $$

where T is a threshold and we set $T = 5$ throughout this paper. $\xi(j, j^r, \varphi^{-1}(j), \varphi^{-1}(j^r))$ is defined similarly. We cast minimizing the pairwise error as maximizing the similarities by using

a Gaussian function, and then $F_g(\mathbf{X}, \mathbf{Y}, \varphi)$ is given by

$$
\begin{aligned}
F_g(\mathbf{X}, \mathbf{Y}, \varphi) \\
= \sum_{i=1}^{M} \sum_{S(i,i^r) \leq T} \sum_{S(\varphi(i),\varphi(i^r)) \leq T} \exp \left[\frac{-\xi \left[i, i^r, \varphi(i), \varphi(i^r) \right]}{2\sigma^2} \right] \\
+ \sum_{j=1}^{N} \sum_{S(j,j^r) \leq T} \sum_{S(\varphi^{-1}(j),\varphi^{-1}(j^r)) \leq T} \exp \left[\frac{-\xi \left[j, j^r, \varphi^{-1}(j), \varphi^{-1}(j^r) \right]}{2\sigma^2} \right].
\end{aligned}
$$
(16)

σ is set to 1 throughout this paper. By representing φ as a multivariant function, we can use a matrix $\mathbf{P}$ to describe the correspondences between point sets $\mathbf{X}$ and $\mathbf{Y}$. If points $x_i \in \mathbf{X}$ and $y_j \in \mathbf{Y}$ are matched, $p_{ij} = 1$; otherwise, $p_{ij} = 0$. And the extra $(M + 1)$th row and $(N + 1)$th column are added to matrix $\mathbf{P}$ to handle outliers

$$
\mathbf{P} = \begin{bmatrix} p_{11} & \cdots & p_{1N} & p_{1,\text{nil}} \\ \vdots & \ddots & \vdots & \vdots \\ p_{M1} & \cdots & p_{MN} & p_{M,\text{nil}} \\ p_{\text{nil},1} & \cdots & p_{\text{nil},N} & 0 \end{bmatrix}.
$$
(17)

In order to enable one-to-one correspondences, the following constraints are imposed on the matching matrix $\mathbf{P}$:

$$
\begin{cases} \sum_{j=1}^{N+1} p_{ij} = 1, & i = 1, 2, \ldots, M \\ \sum_{i=1}^{M+1} p_{ij} = 1, & j = 1, 2, \ldots, N. \end{cases}
$$
(18)

Hence $F(\mathbf{X}, \mathbf{Y}, \varphi)$ can be written as

$$
F(\mathbf{X}, \mathbf{Y}, \mathbf{P}) = \sum_{i=1}^{M} \sum_{j=1}^{N} \eta_{ij} p_{ij} + 2\alpha \sum_{i=1}^{M} \sum_{S(i,i^r) \leq T} \sum_{j=1}^{N} \sum_{S(j,j^r) \leq T} p_{ij} p_{i^r j^r}
$$
$$
\times \exp \left[\frac{-\xi \left[i, i^r, j, j^r \right]}{2\sigma^2} \right].
$$
(19)

As $p_{ij} \in \{0, 1\}$, finding the optimal $\mathbf{P}$ is a typical NP-hard problem. In order to avoid exhaustive combinational search, we use the well-known technique of probabilistic relaxation [42] to find a local optimum by relaxing $p_{ij} \in [0, 1]$. The gradient g_{ij} with respect to the objective function in (19) is

$$
g_{ij} = \eta_{ij} + 4\alpha \sum_{S(i,i^r) \leq T} \sum_{S(j,j^r) \leq T} p_{i^r j^r} \exp \left[\frac{-\xi \left[i, i^r, j, j^r \right]}{2\sigma^2} \right].
$$
(20)

So the matching probability at the rth iteration is updated by

$$
p_{ij}^r := p_{ij}^{r-1} g_{ij}^{r-1} \bigg/ \sum_{k=1}^{N} p_{ik}^{r-1} g_{ik}^{r-1}.
$$
(21)

As only a one-way normalization constraint is enforced in (21), many to one match may exist. In order to enforce one-to-one match, Zheng and Doermann [41] proposed to convert the updated probability matrix to a generalized doubly stochastic matrix through the iterative process of alternated row and column normalization. The doubly stochastic matrix [43] is a
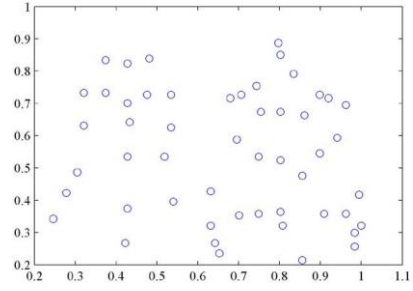

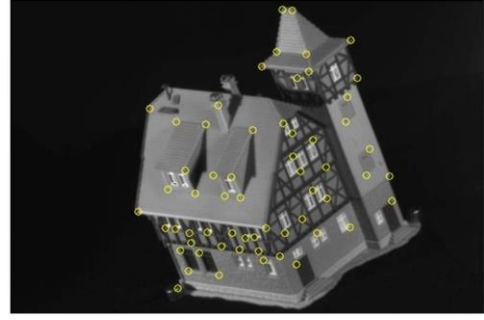Fig. 3. Model point-set for the experiments on the synthetic data.

Fig. 4. Example of labeled point-set in the CMU house images.

nonnegative square matrix with each row and column summing to one and Zheng and Doermann extended it to nonsquare matrix. Here, we follow [41] to achieve one-to-one correspondences. Probabilistic relaxation requires an initial estimation for the update procedure and the final solution depends heavily on the estimation of the initial solution. In this paper, we use the feature similarities computed from the spectral descriptor to initialize the probabilistic matrix $\mathbf{P}$. Similarly, we convert the initial probabilistic matrix $\mathbf{P}$ to a generalized doubly stochastic matrix to impose the constraint of one-to-one match. The probabilities for a point matching to a dummy point, $p_{i,\text{nil}}$ and $p_{\text{nil},j}$, are set to 0.2 empirically. We perform 200 rounds of iterative update in the following experiments, which is sufficient to produce a convergent solution according to our empirical study. The final correspondences are determined by the threshold $p_{ij} \geq 0.6$.

V. EXPERIMENTS

In this section, we present our experimental evaluation of the proposed method. Our experiments are conducted on both synthetic data and real-world images. In these experiments, we pay much attention to the algorithm performance in the presence of significant positional jitter and outliers.

A. Experimental Data

The experimental data are described as follows.
1) Synthetic Data: We perform a model-data matching to analyze the influence of positional jitter and outliers quantitatively. As shown in Fig. 3, the model point-set $\mathbf{X}$ containing 52 points is produced on the basis of the dataset provided by Chui and Rangarajan [44], which can be viewed as discrete points sampled from a shape. A synthetic point-set
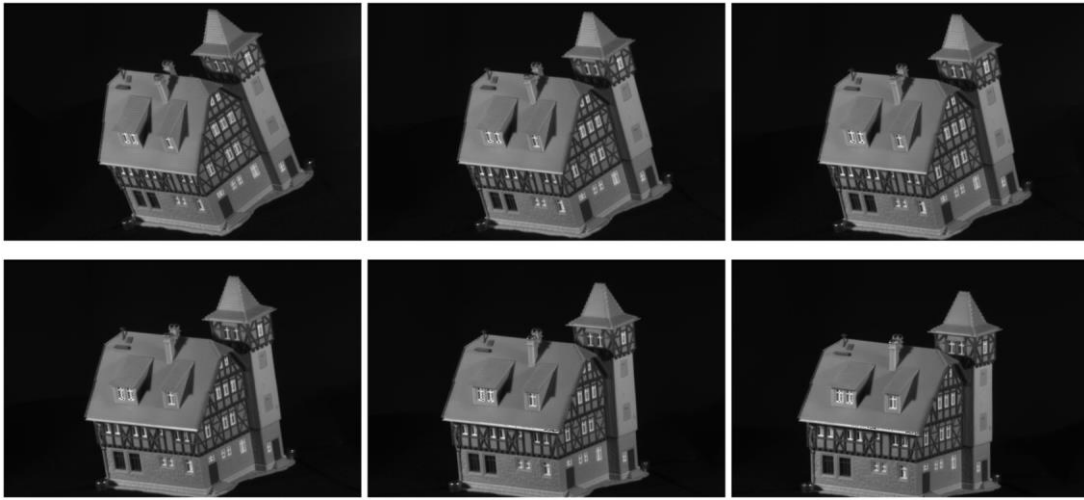
Fig. 5.   Tested CMU house images.

$\mathbf{Y}^r = s\mathbf{R}\mathbf{X} + \mathbf{t}$ is generated for evaluating these algorithms. $\mathbf{t} = [t_x, t_y]$ is the translation vector and

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{22}$$

is the 2-D rotation matrix and s is the scaling parameter. The transform parameter values are randomly chosen from the range $-\pi \le \theta \le \pi$, $1 \le t_x, t_y \le 5$ and $0.5 \le s \le 1.0$. Synthetic positional jitter is produced by adding Gaussian noise to the position of point-set $\mathbf{Y}^r$, where the mean is set to $[0\ 0]^T$ and the standard deviation is defined using the fraction of the average closest distance of point-set $\mathbf{Y}^r$, i.e., $[\lambda d_Y\, 0;\ 0\ \lambda d_Y]$. λ denotes the fraction. The controlled outliers are generated by adding a given number of random points in the scope of $\mathbf{Y}^r$. As the synthetic data are randomly generated, the experimental results are obtained by performing 100 independent trials.

2) CMU House: Following [13], [23], and [26], we also evaluate the algorithm performance on the canonical CMU house sequence. As shown in Fig. 5, six images are chosen for the experimental evaluation. The first frame, which is employed as the template image, is tested against the other frames. To enable an accurate simulation, we labeled 60 corresponding feature points across these frames to simulate corner points. An example is demonstrated in Fig. 4. We randomly select a given number of corresponding pairs as inliers and then remove some of the remaining points in each image to produce outliers. The number of feature points and inliers in each test pairs are summarized in Table I. As both the inliers and the outliers are randomly produced, we run the evaluated algorithms on each test pair 20 times. Obviously these experiments are much more challenging than those on the synthetic data due to that both of the matched images contain outliers and there exists significant perspective distortion between them.

3) Image Pairs: The experiments on some image pairs are performed to further investigate the flexibility of the proposed algorithm. The evaluated image pairs are shown in Fig. 6, some of which are chosen from the ALOI image library [45] and the

TABLE I
NUMBER OF LABELED POINTS AND INLIERS IN THE IMAGE PAIRS OF CMU HOUSE

Image Pairs	Number of Labeled Points	Number of Inliers
1-2	56-58	54
1-3	52-57	47
1-4	50-55	45
1-5	54-57	51
1-6	55-55	50

TABLE II
NUMBER OF FEATURE POINTS AND INLIERS IN THE IMAGE PAIRS

Image Pairs	Number of Detected Points	Number of Inliers
A	48-51	45
B	53-56	51
C	57-52	47
D	47-45	36

Caltech-256 dataset [46]. The feature points are detected using the Harris corner detector. The number of detected points and the manually checked inliers are reported in Table   II.

B. Results and Discussion

1) Weighting Function and Parameter Value: In this part, we first discuss the effect of using different weighting function, and then study some parameter settings in our algorithm. In order to achieve the quantitative analysis in an easy way, we conduct these evaluations on the synthetic data.

As mentioned in Section III, we can use some alternative weighting functions to define edge weights. Here, we investigate the performance of using the Gaussian weighting function, the sigmoidal weighting function and the increasing weighting function [12]. Fig. 7 plots the average matching accuracy over 100 random trials as a function of the standard deviation of the added Gaussian positional errors. The number of outliers is five in this evaluation. We can observe that the best performance is returned by the Gaussian weighting function. Hence, we choose the Gaussian weighting function throughout this paper.
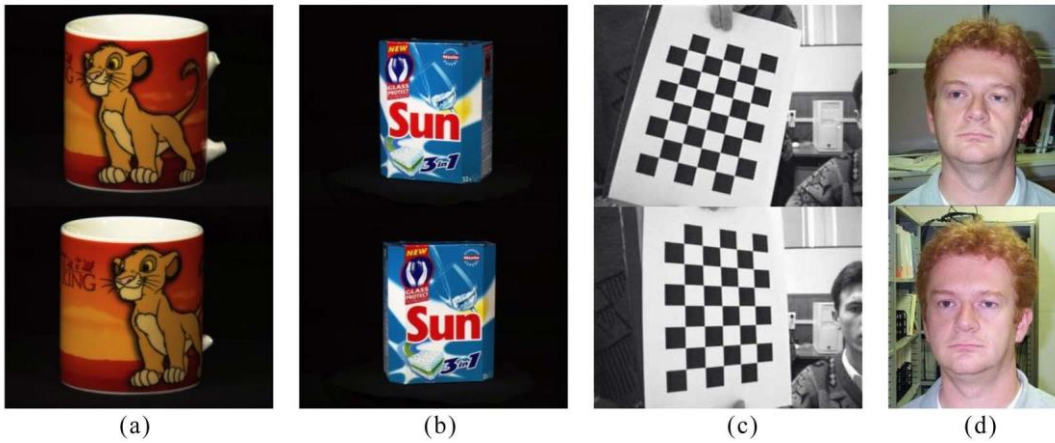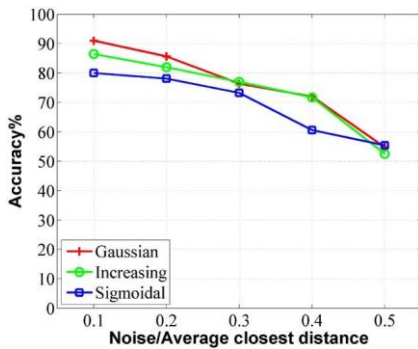
Fig. 6.   Evaluated image pairs.

Fig. 7.   Evaluation of different weighting functions on the synthetic data.
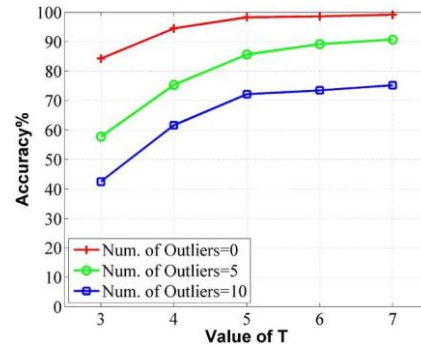

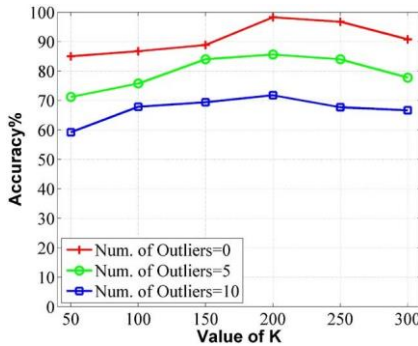Fig. 9.   Evaluation of different values of T on the synthetic data.

Fig. 8.   Evaluation of different values of K on the synthetic data.

Next, we investigate the choice of the value of K. Fig. 8 shows the average matching accuracy as a function of the value of K. The standard deviation of the added Gaussian positional errors is set to $[0.2d_Y\ 0;\ 0\ 0.2d_Y]$. We can observe that K is not a sensitive parameter and the best performance is obtained in the range of $[150, 250]$. The performance degrades to some extent when the value of K is beyond this range. The reason is that small value leads to the loss of the discriminative power of the descriptor while large value makes the descriptor sensitive to positional errors and outliers.

We finally study the effect of varying the value of T. Fig. 9 plots the average matching accuracy as a function of the value of T. We can observe that a useful margin of improvement on

accuracy can be obtained if we tune the value of T. As we focus on the performance of the descriptor and the computation of g_{ij} suffers from considerable time overhead when the value of T becomes larger, we set $T = 5$ that can yield moderately better results throughout this paper.

2) Matching: In this part, we comprehensively compare our algorithm with four methods: the method of Leordeanu and Hebert [14], the method of Cour *et al.* [15], the method of Silletti *et al.* [25], and the method of Pang *et al.* [27]. For the method of Silletti *et al.*, we do not include the metric with respect to pixel information to make a fair comparison for these methods.

We first investigate the experimental results on the synthetic data. The experimental settings follow those in the evaluation of different weighting functions. The values of the control parameter in the compatibility matrix for the method of Leordeanu and Hebert and the method of Cour *et al.* are tuned as 0.5 and 0.5, respectively. We test the cases of 0, 5, and 10 outliers, respectively (see Fig. 10). It is known that the method of Leordeanu and Hebert [14] is very sensitive to scaling. As we focus on the algorithm performance for positional jitter and outliers, we fix $s = 1$ for this method when generating synthetic data.

Overall, the proposed matching algorithm performs the best in this evaluation. With the increment of noise and outliers, the accuracies of the baselines decrease dramatically.
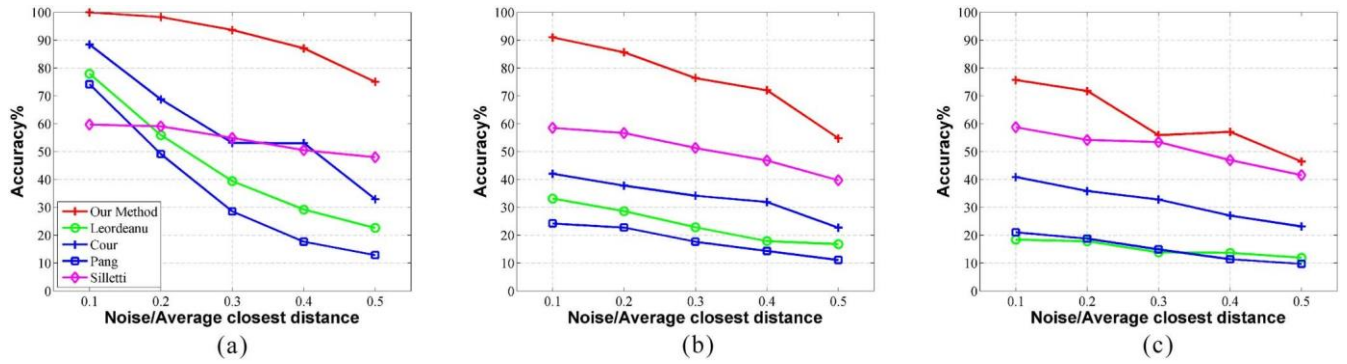
Fig. 10. Effect of varying positional jitter on matching accuracy with a given number of outliers. (a) Number of outliers = 0. (b) Number of outliers = 5. (c) Number of outliers = 10.
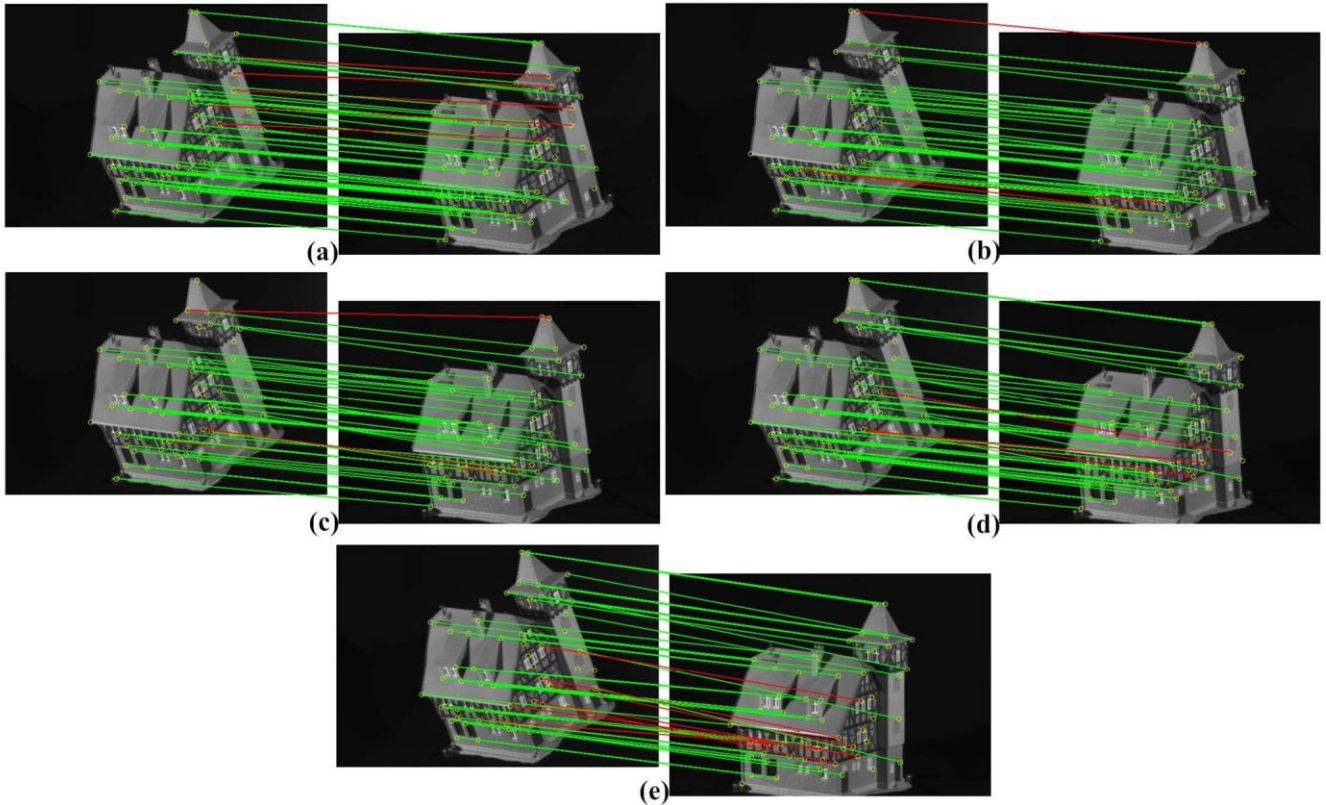

Fig. 11. Some matching results on the CMU house images. (a) First image versus the second image. (b) First image versus the third image. (c) First image versus the fourth image. (d) First image versus the fifth image. (e) First image versus the sixth image.

It is interesting to see that the results returned by the method of Silletti *et al.* are almost irrelevant to outliers, although its overall performance is not that good. This phenomenon is due largely to that these experiments fall into the categories of model-data matching and the used point associate matrix happens to handle this situation well. Although our matching results are finally obtained by using iterative optimization, the solution via probabilistic relaxation is indeed closely related to the initial estimation given by the proposed descriptor. Therefore, from these experiments we can validate the robustness and the discriminative power of the proposed descriptor.

The run time of each algorithm for the evaluation on the case of 0 outlier is reported in Table III. For the method of

TABLE III
RUN TIME (S) ON THE SYNTHETIC DATA

Methods	Run time
Our method	11.5
Leordeanu [14]	106.3
Cour [15]	17.3
Silletti [25]	1.7
Pang [27]	0.9

Cour *et al.*, we use the open-source code implemented with C++. All the other methods are implemented with MATLAB. All our run time results are obtained on a Win7 E5200 2.5G Hz PC. The direct comparison is unfair due to the different programming languages, but we can still learn that Pang's
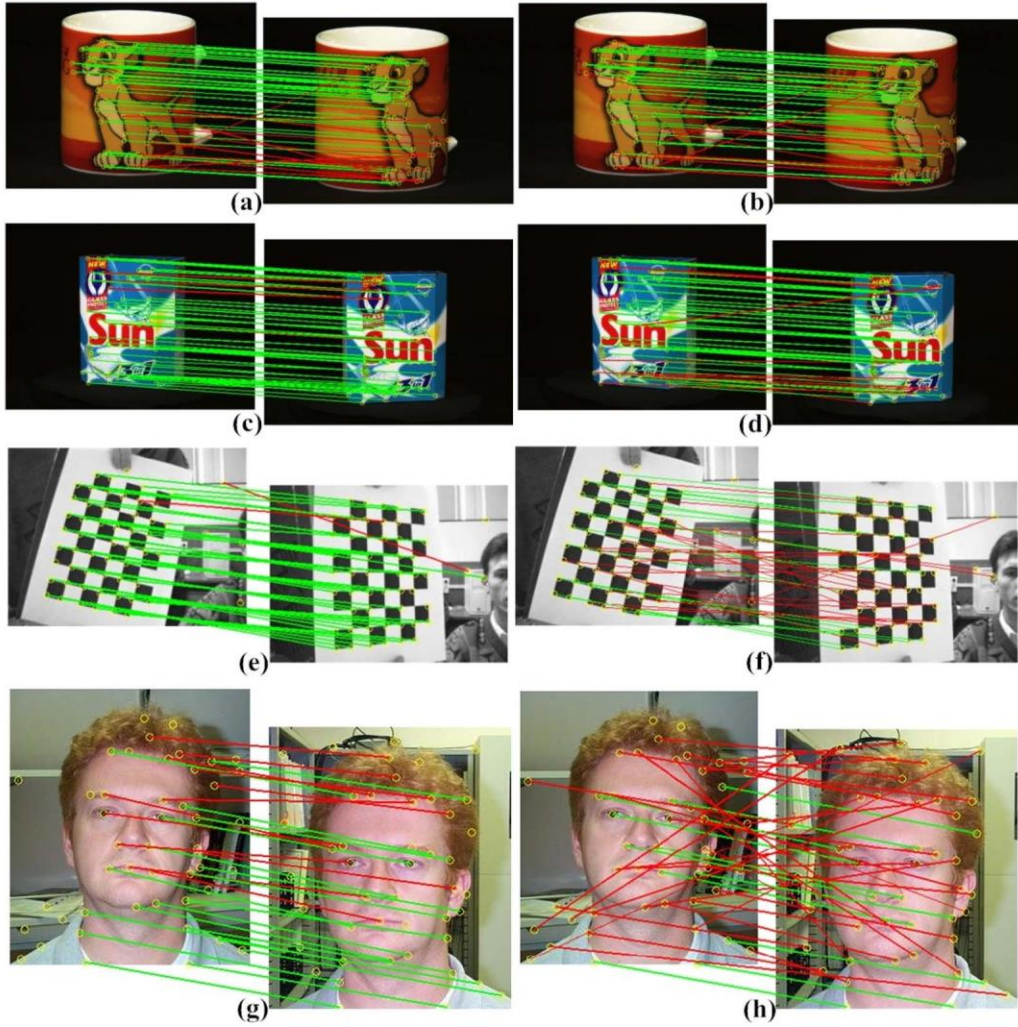
Fig. 12.   Some matching results on the image pairs. (a) Result on pair A obtained from our method. (b) Result on pair A obtained from Leordeanu's method. (c) Result on pair B obtained from our method. (d) Result on pair B obtained from Cour's method. (e) Result on pair C obtained from our method. (f) Result on pair C obtained from Silletti's method. (g) Result on pair D obtained from our method. (h) Result on pair D obtained from Pang's method.

TABLE IV
AVERAGE ACCURACIES% ON THE IMAGE PAIRS OF THE CMU HOUSE SEQUENCE

Image Pairs	Our method	Leordeanu [14]	Cour [15]	Silletti [25]	Pang [27]
1-2	91.4	71.8	84.1	65.6	46.7
1-3	83.6	45.9	62.7	59.3	28.9
1-4	71.8	14.8	41.1	49.8	16.2
1-5	84.9	15.7	20.3	35.2	14.4
1-6	65.9	8.5	15.2	39.4	10.2

TABLE V
EXPERIMENTAL RESULTS ON THE IMAGE PAIRS (NUMBER OF CORRECT PAIRS/NUMBER OF FOUND PAIRS   AND ACCURACY%)

Image Pairs	Our method		Leordeanu [14]		Cour [15]		Silletti [25]		Pang [27]	
A	37/48	77.8	31/48	64.6	36/48	75.0	30/48	62.5	25/48	52.1
B	45/53	84.9	41/53	77.4	43/53	81.1	28/48	58.3	30/48	62.5
C	41/47	87.2	9/52	17.3	34/52	65.4	20/44	45.5	5/52	9.6
D	19/31	61.3	12/45	26.7	17/45	37.8	11/39	28.2	11/45	24.4

method is the most efficient. The method of Leordeanu and Hebert, the method of Silletti *et al.*, and our method contain a relatively time-consuming procedure of matrix decomposition with cubic complexity. The difference of run time between them is mainly related to the sizes of the operated matrices.

The method of Silletti *et al.* performs decomposition on an associate matrix between point-sets. The method of Leordeanu and Hebert handles a compatible matrix with   considerable size that enumerates all the possible correspondence candidates, making it unsuitable for large point-sets. Although   the
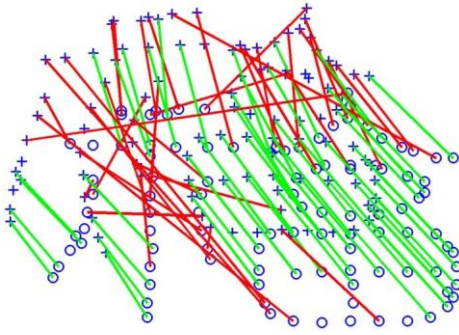
Fig. 13. Matching result obtained from our method on two toy point-sets under nonrigid deformation.

decomposition has to be done for each point in our method, it is only performed on the graph Laplacian denoting small neighboring sub-point-sets. Therefore, our method achieves moderately better performance with respect to run time.

We then turn to the results on the CMU house images. The values of the control parameter in the compatibility matrix for the method of Leordeanu and Hebert and the method of Cour *et al.* are tuned as 5 and 10, respectively. The average accuracies are summarized in Table IV, where the accuracy is computed as the ratio between the average of correct pairs and the average of found pairs. Meanwhile, some matching results are demonstrated in Fig. 11, where green lines indicate correct correspondences while red lines denote mismatches. Our method achieves the best performance with respect to accuracy and stability. When the difference of viewing angles becomes large, the accuracies of the other four methods show an obvious tendency of degradation. The worst results are produced by Pang's method. Although Pang's method is exactly a fast matching algorithm, its performance with respect to accuracy is not satisfactory.

Next, we study the results on the image pairs. We show the experimental comparison in Table V, where the results are reported as the number of correct pairs/the number of found pairs and the accuracies. Some matching results are plotted in Fig. 12. In the evaluation on these different kinds of images, the proposed method achieves the best relative performance, followed by the method of Cour *et al*. In summary, the conclusion here is basically consistent with that drawn from the experiments on the CMU house images, which verify that the descriptor provides an effective means for characterizing point patterns.

Finally, we discuss the weakness of the proposed method. We investigate the performance of the proposed algorithm in the presence of nonrigid deformation. Fig. 13 plots the matching result on two Chinese character point-sets under nonrigid deformation [44]. It can be observed that our method cannot handle this situation well. We consider that this disadvantage arises from two aspects.

1) Euclidean distances are used to construct the spectral descriptor and they may not be preserved well under nonrigid deformation.
2) It is difficult to achieve a good matching result for point-sets under nonrigid deformation using only the proposed matching method so that the common framework of

iterative correspondence and transformation estimation is needed.

Although we concentrate on handling outliers and positional jitter in this paper, in our ongoing research, we will attempt to combine the thin plate spline model [16] with our matching method and consider alternative ways to define edge weights that are able to resist nonrigid deformation to some extent.

VI. CONCLUSION

In this paper, we have studied local spectral representations of point patterns in order to improve the robustness of spectral matching methods against positional jitter and outliers. The proposed local descriptor uses the distribution of normalized Laplacian spectra to characterize feature points. We have combined the descriptor with the approximate distance order for recovering correspondences. Extensive experiments verify that our method is robust to positional jitters and outliers.

In the future work, we will pay more attention to utilize some other spectral properties. Applying hypergraph to our model is another promising direction.

REFERENCES

[1] H. Qiao, X. Xi, Y. Li, W. Wu, and F. Li, "Biologically inspired visual model with preliminary cognition and active attention adjustment," *IEEE Trans. Cybern.*, Doi: 10.1109/TCYB.2014.2377196.

[2] Y. Wan *et al.*, "A quasi-dense matching approach and its calibration application with Internet photos," *IEEE Trans. Cybern.*, vol. 45, no. 3, pp. 370–383, Mar. 2015.

[3] C.-C. Tsai, H.-Y. Lin, J. Taur, and C.-W. Tao, "Iris recognition using possibilistic fuzzy matching on local features," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 1, pp. 150–162, Feb. 2012.

[4] B. Li, Q. Meng, and H. Holstein, "Articulated pose identification with sparse point features," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 34, no. 3, pp. 1412–1422, Jun. 2004.

[5] H. Liu and S. Yan, "Common visual pattern discovery via spatially coherent correspondences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, San Francisco, CA, USA, 2010, pp. 1609–1616.

[6] J. Han, K. N. Ngan, M. Li, and H.-J. Zhang, "A memory learning framework for effective image retrieval," *IEEE Trans. Image Process.*, vol. 14, no. 4, pp. 511–524, Apr. 2005.

[7] D. Tao, X. Tang, X. Li, and X. Wu, "Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 7, pp. 1088–1099, Jul. 2006.

[8] U. Gaur, Y. Zhu, B. Song, and A. Roy-Chowdhury, "A 'string of feature graphs' model for recognition of complex activities in natural videos," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Barcelona, Spain, 2011, pp. 2595–2602.

[9] D. Tao, X. Li, X. Wu, and S. J. Maybank, "General tensor discriminant analysis and Gabor features for gait recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 10, pp. 1700–1715, Oct. 2007.

[10] G. L. Scott and H. C. Longuet-Higgins, "An algorithm for associating the features of two images," *Proc. Roy. Soc. London B, Biol. Sci.*, vol. 244, no. 1309, pp. 21–26, 1991.

[11] L. S. Shapiro and J. M. Brady, "Feature-based correspondence: An eigenvector approach," *Image Vis. Comput.*, vol. 10, no. 5, pp. 283–288, 1992.

[12] M. Carcassoni and E. R. Hancock, "Spectral correspondence for point pattern matching," *Pattern Recognit.*, vol. 36, no. 1, pp. 193–204, 2003.

[13] H. F. Wang and E. R. Hancock, "Correspondence matching using kernel principal components analysis and label consistency constraints," *Pattern Recognit.*, vol. 39, no. 6, pp. 1012–1025, 2006.

[14] M. Leordeanu and M. Hebert, "A spectral technique for correspondence problems using pairwise constraints," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 2. Beijing, China, 2005, pp. 1482–1489.

[15] T. Cour, P. Srinivasan, and J. Shi, "Balanced graph matching," in *Advances in Neural Information Processing Systems*, vol. 19. London, U.K.: MIT Press, 2007, pp. 313–350.

[16] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.

[17] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[18] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis.*, Graz, Austria, 2006, pp. 404–417.

[19] B. Luo, R. C. Wilson, and E. R. Hancock, "Spectral embedding of graphs," *Pattern Recognit.*, vol. 36, no. 10, pp. 2213–2230, 2003.

[20] A. Shokoufandeh, D. Macrini, S. Dickinson, K. Siddiqi, and S. W. Zucker, "Indexing hierarchical structures using graph spectra," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 7, pp. 1125–1140, Jul. 2005.

[21] S. Umeyama, "An eigendecomposition approach to weighted graph matching problems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 5, pp. 695–703, May 1988.

[22] M. Pilu, "A direct method for stereo correspondence based on singular value decomposition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, San Juan, Puerto Rico, 1997, pp. 261–266.

[23] M. Carcassoni and E. R. Hancock, "Correspondences matching with modal clusters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1609–1615, Dec. 2003.

[24] E. Delponte, F. Isgrò, F. Odone, and A. Verri, "SVD-matching using SIFT features," *Graph. Models*, vol. 68, no. 5, pp. 415–431, 2006.

[25] A. Silletti, A. Abate, J. D. Axelrod, and C. J. Tomlin, "Versatile spectral methods for point set matching," *Pattern Recognit. Lett.*, vol. 32, no. 5, pp. 731–739, 2011.

[26] H. Qiu and E. R. Hancock, "Graph matching and clustering using spectral partitions," *Pattern Recognit.*, vol. 39, no. 1, pp. 22–34, 2006.

[27] Y. Pang, M. Shang, Y. Yuan, and J. Pan, "Scale invariant image matching using triplewise constraint and weighted voting," *Neurocomputing*, vol. 83, pp. 64–71, Apr. 2012.

[28] R. C. Wilson, E. R. Hancock, and B. Luo, "Pattern vectors from algebraic graph theory," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 7, pp. 1112–1124, Jul. 2005.

[29] R. C. Wilson and P. Zhu, "A study of graph spectra for comparing graphs and trees," *Pattern Recognit.*, vol. 41, no. 9, pp. 2833–2841, 2008.

[30] B. Xiao, E. R. Hancock, and R. C. Wilson, "Graph characteristics from the heat kernel trace," *Pattern Recognit.*, vol. 42, no. 11, pp. 2589–2606, 2009.

[31] B. Xiao, E. R. Hancock, and R. C. Wilson, "Geometric characterization and clustering of graphs using heat kernel embeddings," *Image Vis. Comput.*, vol. 28, no. 6, pp. 1003–1021, 2010.

[32] J. Tang, L. Shao, and X. Zhen, "Robust point pattern matching based on spectral context," *Pattern Recognit.*, vol. 47, no. 3, pp. 1469–1484, 2014.

[33] B. Mohar and Y. Alavi, "The Laplacian spectrum of graphs," *Graph Theory Comb. Appl.*, vol. 2, pp. 871–898, 1991.

[34] H. Qiao, Y. Li, T. Tang, and P. Wang, "Introducing memory and association mechanism into a biologically inspired visual model," *IEEE Trans. Cybern.*, vol. 44, no. 9, pp. 1485–1496, Sep. 2014.

[35] Z. Zhang and D. Tao, "Slow feature analysis for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 3, pp. 436–450, Mar. 2012.

[36] J. Han *et al.*, "Representing and retrieving video shots in human-centric brain imaging space," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2723–2736, Jul. 2013.

[37] Y. Yang, F. Nie, S. Xiang, Y. Zhuang, and W. Wang, "Local and global regressive mapping for manifold learning with out-of-sample extrapolation," in *Proc. AAAI Conf. Artif. Intell.*, Atlanta, GA, USA, 2010, pp. 649–654.

[38] W. Liu and D. Tao, "Multiview Hessian regularization for image annotation," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2676–2687, Jul. 2013.

[39] F. R. Chung, *Spectral Graph Theory*. vol. 92. Providence, RI, USA: AMS Bookstore, 1997.

[40] X. Yang, H. Qiao, and Z.-Y. Liu, "Feature correspondence based on directed structural model matching," *Image Vis. Comput.*, vol. 33, pp. 57–67, Jan. 2015.

[41] Y. F. Zheng and D. Doermann, "Robust point matching for nonrigid shapes by preserving local neighborhood structures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 4, pp. 643–649, Apr. 2006.

[42] M. Pelillo, "The dynamics of nonlinear relaxation labeling processes," *J. Math. Imag. Vis.*, vol. 7, no. 4, pp. 309–323, 1997.

[43] R. Sinkhorn, "A relationship between arbitrary positive matrices and doubly stochastic matrices," *Ann. Math. Statist.*, vol. 35, no. 2, pp. 876–879, 1964.

[44] H. Chui and A. Rangarajan, "A new point matching algorithm for nonrigid registration," *Comput. Vis. Image Understand.*, vol. 89, no. 2, pp. 114–141, 2003.

[45] J.-M. Geusebroek, G. J. Burghouts, and A. W. Smeulders, "The Amsterdam library of object images," *Int. J. Comput. Vis.*, vol. 61, no. 1, pp. 103–112, 2005.

[46] G. Griffin, A. Holub, and P. Perona, "Caltech-256 object category dataset," California Inst. Technol., Pasadena, CA, USA, Tech. Rep. CNS-TR-2007-001, 2007.

Jun Tang is an Associate Professor with the School of Electronics and Information Engineering, Anhui University, Hefei, China. His current research interests include computer vision, machine learning, and pattern recognition.
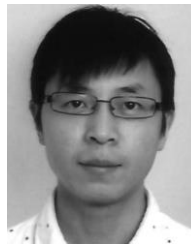
Ling Shao (M'09–SM'10) is currently a Professor with the Department of Computer Science and Digital Technologies, Northumbria University, Newcastle upon Tyne, U.K. He was a Senior Lecturer with the Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield, U.K., from 2009 to 2014, and a Senior Scientist with Philips Research, Eindhoven, The Netherlands, from 2005 to 2009. His research interests include computer vision, image/video processing, and machine learning. He is an Associate Editor of the IEEE Transactions on Image Processing, the IEEE Transactions on Cybernetics, and several other journals. He is a fellow of the British Computer Society and the Institution of Engineering and Technology.

Xuelong Li (M'02–SM'07–F'12) is a Full Professor with the Center for OPTical IMagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China.

Ke Lu was born in Ningxia, China, in 1971. He received the master's degree from the Department of Mathematics and the Ph.D. degree from the Department of Computer Science both from Northwest University, Kirkland, WA, USA, in 1998 and 2003, respectively.

He was a Post-Doctoral Fellow at the Institute of Automation, Chinese Academy of Sciences, Beijing, China, from 2003 to 2005. He is currently a Professor with the University of the Chinese Academy of Sciences. His current research interests include computer vision, 3-D image reconstruction, and computer graphics.