Point Pattern Matching Based on Line Graph Spectral Context
and Descriptor Embedding

Jun Tang
Key Lab of IC&SP, Ministry of Education
Anhui University
Hefei, China
tangjunahu@gmail.com

Ling Shao, Simon Jones
Department of Electronic and Electrical Engineering
The University of Sheffield
Sheffield, UK
ling.shao, simon.m.jones@sheffield.ac.uk

Abstract

Spectral methods have been extensively studied for point pattern matching. In this work, we aim to render the spectral matching algorithm more robust for positional jitter and outliers. We concentrate on the issue of spectral representation for point patterns. A local structural descriptor, called the line graph spectral context, is proposed to characterize the attribute of point patterns, making it fundamentally different from the available representation approaches at the global level. For any given point, we first construct a line graph using its neighboring points. Then the eigenvalues of various matrix representations associated with the obtained line graph are used as the point descriptor. Furthermore, the similarities between the descriptors are evaluated by comparing their low dimensional embedding via the technique of multiview spectral embedding. The proposed descriptor is finally integrated with a graph-matching framework for establishing the correspondences. Comparative experiments conducted on both synthetic data and real-world images show the effectiveness of the proposed method, especially in the presence of positional jitter and outliers.

1. Introduction

Point pattern matching (PPM) is a critical issue in computer vision and pattern recognition due to its wide range of applications in 3D reconstruction, image registration, object recognition, etc. Over the past two decades, a number of attempts have been made to use graph spectra for matching feature point-sets. Scott and Longuet-Higgins [10] proposed the pioneer work of spectral matching. They showed how to recover correspondences by performing singular value decomposition (SVD) on the point association matrix between different images. The main drawback of this method is its sensitivity to the degree of rotation and the amount of scaling between two point-sets. In order to overcome these problems, Shapiro and Brady [11] proposed a method which uses the intra-image point proximity matrix. A modal representation for the matched point-sets is constructed by utilizing the eigenvectors of the individual proximity matrices. Then correspondences are found by comparing the rows of the modal matrices. Unfortunately, this method fails abruptly when the point-sets are of different sizes. These two pieces of work built a foundation for subsequent research. Carcassoni and Hancock [3] attempted to render Shapiro and Brady’s method robust to positional jitter by combining it with the EM algorithm. Meanwhile, several alternative kernels instead of the Gaussian are provided for constructing the proximity matrix. Furthermore, they used spectral clusters to tackle the problem of different point-set sizes [2]. Provided that the point-sets have a well-defined cluster structure, relatively better performance can be achieved in contrast to that of [3]. Wang and Hancock [15] explored the intrinsic relationship between graph spectral methods and kernel PCA, and a kernelized variant of [11] was presented. Delphonte et al. [6] defined a miscellaneous proximity matrix by incorporating the similarity of SIFT descriptors. Siletti et al. [13] combined a num-

Figure 1. The flowchart of constructing LGSC.
ber of similarity metrics together to construct the proximity matrix, and the method of [10] was employed to find correspondences. The idea of Leordeanu and Hebert [8] is somewhat different from those mentioned above. They conjured that there is a compact clustering structure underlying all the correct matching pairs. According to this assumption, they constructed a graph with potential correspondences as nodes and the compatibility between potential correspondences as edges, and then correspondences are delivered by utilizing the technique of spectral relaxation. Cour et al. [5] extended the work of [8] by imposing a constraint of affine transformation. In a wider sense, the work of Zass and Shashua [18] can be also seen as an extension of [8]. They formulated PPM as a hypergraph matching problem, in which the hyperedge is defined as the affinity beyond pairs.

It is well-known that spectral matching methods are notoriously susceptible to positional jitter and outliers. Although many works have been dedicated to this issue, it remains a major bottleneck. In this work, the properties of graph spectra are used for locating correspondences, which inherits the essential idea of [11]. In [11, 2, 3, 15], the eigenvectors are used to characterize the attribute domain of a point-set at the global level. For such a representation, the influence of positional jitter can be alleviated to some extent by combining it with some optimization strategies. However, the presence of outliers would lead to the real difficulty, because we need to truncate the eigenvectors for comparison and information loss/distortion is inevitable during this procedure. And the algorithm performance may degrade greatly due to this operation. Therefore, the key to rendering spectral methods robust for positional jitter and outliers lies in the effective spectral representation of point patterns. On the other hand, we have observed that the effect of eigenvalues has been paid less attention to and only restricted to scaling eigenvectors [2, 15] in almost all the available spectral matching methods. In fact, eigenvalues are powerful facilities to model data by providing a compact summary of the graph structure, which has been successfully applied to object recognition and image classification. For instance, Shokoufandeh et al. [12] used the eigenvalues of the directed acyclic graph (DAG) to encode the hierarchical structure of images and applied it to the task of 3D object recognition. In [9], Luo et al. used eigenvalues as features to vectorize graphs. Wilson and Zhu [16] showed the potential of using the eigenvalues of different matrix representations (i.e., adjacency matrix, Laplacian matrix, normalized Laplacian matrix, etc.) in image classification. To address the weakness of previous works, we propose a novel method to characterize point patterns, in which we use eigenvalues to construct a local structural descriptor rather than developing a global representation. Intuitively, positional jitter and outliers might have less influence on a local spectral representation in contrast to the global one. Meanwhile, the discriminative power of eigenvalues can provide an effective means for representing the attribute domain.

The main contributions of our work are as follows. First, to the best of our knowledge, we are among the first to use eigenvalues as local features for PPM, although the discriminative power of eigenvalues has been verified in object recognition and image classification. Second, we introduce the concept of line graph into spectral matching and demonstrate its promising results, which has been paid little attention to in the application of graph theory. Third, we present a strategy that can effectively tackle the issue of multipletspectra representation for point patterns.

2. Methodology

We first provide the formulation of PPM. Given two related feature point-sets \( X = \{x_i | i = 1, 2, \ldots, M\} \) and \( Y = \{y_j | j = 1, 2, \ldots, N\} \), the set of correspondences between two matched point-sets is given by \( \phi: X \rightarrow Y \). Without loss of generality, we assume \( M \neq N \).

2.1. Proposed Spectral Descriptor

Fig. 1 depicts the procedure of constructing the proposed spectral descriptor, which we refer to as line graph spectral context (LGSC). For convenience, we use the notation set of point-set \( X \) to describe the construction procedure. For any point \( x_i \in X \) and a predefined series \( \Theta = \{d_1, d_2, \ldots, d_T\} \), we give the following computational method:

**Step 1:** For each \( d_i \in \Theta \), select the \( d_i \) nearest neighbors of \( x_i \) and obtain a sub point-set \( \Omega_{it} \).

**Step 2:** Construct a weighted star graph \( G_{it} = \{V_{it}, E_{it}\} \) on the sub point-set \( \Omega_{it} \), where \( V_{it} \) and \( E_{it} \) are the vector (point) set and the edge set, respectively. The edges only exist between \( x_j \) and \( x_j' \in \Omega_{it} \). And the edge weight is defined as the Euclidean distance between its associated vectors.

**Step 3:** Transform graph \( G_{it} \) into its corresponding **line graph** \( \mathbf{H}_{it} \). The adjacency matrix of line graph \( \mathbf{H}_{it} \) is defined as:

\[
A_{it}(p, q) = \begin{cases} 
\exp\left(-\frac{|e_p - e_q|^2}{2\sigma^2}\right) & p \neq q \\
0 & p = q \\
e_p, e_q & e_p, e_q \in E_{it}
\end{cases}
\]  

(1)

The Laplacian matrix and the Quasi-Laplacian matrix of \( \mathbf{H}_{it} \) are defined as \( \mathbf{L}_{it} = \mathbf{D}_{it} - \mathbf{A}_{it} \) and \( \mathbf{Q}_{it} = \mathbf{D}_{it} + \mathbf{A}_{it} \), respectively, where \( \mathbf{D}_{it} \) is the diagonal degree matrix and its diagonal elements are given by \( \mathbf{D}_{it}(p, p) = \sum_{q=1}^{N} A_{it}(p, q) \).

**Step 4:** Perform SVD on \( \mathbf{A}_{it}, \mathbf{L}_{it} \) and \( \mathbf{Q}_{it} \). For \( \mathbf{A}_{it} \), we have:

\[
\mathbf{A}_{it} = \mathbf{U}_{it} \Sigma_{it} \mathbf{U}_{it}^T
\]  

(2)

where \( \mathbf{U}_{it} \) is composed of the eigenvectors of \( \mathbf{A}_{it} \) and the diagonal elements of \( \Sigma_{it} \) are the absolute values of the eigenvalues.
values of $A_{it}$ sorted in a descending order. Similar expressions can be obtained for $L_{it}$ and $Q_{it}$. Consequently, we have three vectors composed of eigenvalues to sketch the structure information of the sub point-set:

$$
S^A_{it} = \{ |\lambda^1_{it}|, |\lambda^2_{it}|, \ldots, |\lambda^{|Q_{it}|}_{it}| \}
$$

$$
S^L_{it} = \{ |\lambda^1_{it}|, |\lambda^2_{it}|, \ldots, |\lambda^{|Q_{it}|}_{it}| \}
$$

$$
S^Q_{it} = \{ |\lambda^1_{it}|, |\lambda^2_{it}|, \ldots, |\lambda^{|Q_{it}|}_{it}| \}
$$

(3)

Through the aforementioned computation, we obtain a feature descriptor $desc(x_i)$ for $x_i \in X$:

$$
desc(x_i) = \{ S^A_{i1}, S^L_{i1}, S^Q_{i1}, \ldots, S^A_{iT}, S^L_{iT}, S^Q_{iT} \}
$$

(4)

To make it clearer, we give some further explanation for LGSC.

In step 3, the concept of line graph is used. The line graph of an undirected graph represents the adjacencies between edges of the original graph. Formally, given graph $G_{it}$, its line graph $H_{it}$ is a graph such that each vertex of $H_{it}$ represents an edge of $G_{it}$ and two vertices of $H_{it}$ are adjacent if and only if their corresponding edges share a common vertex in $G_{it}$. We briefly describe the reason for using the line graph. Note that the weighted graph obtained in step 2 is a star graph. It is easy to know that the rank of its adjacency matrix is 2. That is, the number of its non-zero eigenvalues is 2. Obviously the descriptors may not be discriminative if the eigenvalues of the star graph are used as features. Consequently, we convert the star graph into its corresponding line graph to obtain more non-zero eigenvalues. Meanwhile, as the obtained line graph is still a connected graph, the number of the non-zero eigenvalues of its Laplacian matrix or Quasi-Laplacian matrix does not change.

In step 4, we use three types of spectra to represent the sub point-set. The motivation is two-fold: (1) eigenvalues are the compact reflection of a graph structure, so we believe that using multiple spectra can make the representation more distinctive; (2) according to [16], using more than one kind of spectrum may reduce the ambiguities caused by cospectrality.

In addition, LGSC holds the property of some geometric transformation invariance.

i) LGSC is invariant to translation and rotation.

As LGSC consists of relative distances, obviously it is invariant to translation. As for rotation invariance, we only need to prove its permutation invariance due to the rotation invariance of the selected nearest neighbors. Let $A' = \Psi A \Psi^T$ be an adjacency matrix derived from rearranging the sub point-sets with the permutation matrix $\Psi$. Substitute $A = U \Delta U^T$ into $A' = \Psi A \Psi^T$, we have $A' = \Psi U \Delta U^T \Psi^T = (\Psi U) \Delta (\Psi U)^T$. Since $A'$ is a real symmetric matrix, its result of SVD is unique and the diagonal values of $\Delta$ are the singular values of $A'$. Similar conclusions can be made for the Laplacian matrix $L$ and the Quasi-Laplacian matrix $Q$, and hence LGSC is rotation invariance.

ii) LGSC is invariant to scaling.

The selected sub point-set is scale invariant. We can easily make LGSC invariant to scaling only by tuning $\sigma$ in Eq. (1). For instance, we can compute the inner average distances $l_X$ and $l_Y$ of point-sets $X$ and $Y$, respectively, then we set $\sigma_Y = \sigma_X l_Y/l_X$.

The next critical issue is how to effectively evaluate the similarities between the proposed descriptors. We may simply concatenate all these features together as a new vector and then compare the resulting vectors. However, this concatenation often does not make sense because it ignores the specific statistical property of each feature. On the other hand, our multi-spectra descriptor can be treated as an approach of using multiple features from different views to characterize an object. Therefore, we leverage multi-view spectral embedding (MSE) [17] to encode the eigenvalues obtained from various matrix representations so as to achieve a semantically meaningful embedding. In consideration of space, we omit the introduction of this algorithm and leave the details to the original paper [17]. For each $x_i \in X$, we have $\{S^A_{it}, S^L_{it}, S^Q_{it}\}$ for $x_i \in X$ and $\{S^A_{jt}, S^L_{jt}, S^Q_{jt}\}$ for $y_j \in Y$. By fusing such multiple feature vectors via MSE, we obtain low-rank representations $\hat{S}_{it}$ for $x_i \in X$ and $\hat{S}_{jt}$ for $y_j \in Y$, and hence the similarity between $x_i \in X$ and $y_j \in Y$ can be given by:

$$
w_{ij} = sim(x_i, y_j) = \exp(-\beta \prod_{t=1}^{T} ||\hat{S}_{it} - \hat{S}_{jt}||)
$$

(5)

where $\beta$ is a smoothing coefficient.

2.2. Applying LGSC to PPM

Due to the unavoidable ambiguities of the local structural descriptor caused by the potential existence of similar subsets, directly applying it is unlikely to achieve a satisfied matching result. As a common approach, the combination with spatial constraints is frequently used to refine the matching results delivered by local descriptors [7, 4]. Here we resort to embedding the proposed spectral descriptor with a typical graph-matching framework [19]. In a nutshell, both the similarities between descriptors and the constraints of the geometric neighborhood are taken into account when finding correspondences. Next we elaborate the details of the proposed matching algorithm. The matching objective function is defined as:

$$
F(X, Y, \phi) = \alpha F_a(X, Y, \phi) + (1 - \alpha) F_g(X, Y, \phi)
$$

(6)

where the constant $\alpha$ weighs the feature similarities $F_a(X, Y, \phi)$ and $F_g(X, Y, \phi)$ denote the constraints of the
geometric neighborhood. We set \( \alpha = 0.8 \) throughout this paper. \( F_\alpha(X, Y, \phi) \) is given by:

\[
F_\alpha(X, Y, \phi) = \sum_{i=1}^{M} w_{i, \phi(i)} \tag{7}
\]

As for \( F_\gamma(X, Y, \phi) \), we borrow the idea from [19] to define it as:

\[
F_\gamma(X, Y, \phi) = \sum_{i=1}^{M} \sum_{a \in N_i} \delta(\phi(i), \phi(a)) + \sum_{j=1}^{N} \sum_{b \in N_j} \delta(\phi^{-1}(j), \phi^{-1}(b)) \tag{8}
\]

where \( N_i \) and \( N_j \) denote the neighbors of \( x_i \in X \) and \( y_j \in Y \), respectively. And the metric of neighbors is given by:

\[
\delta(i, a) = \begin{cases} 
1 & a \in N_i \\
0 & a \notin N_i 
\end{cases} \tag{9}
\]

Then we represent the matching function \( f \) in a multivariate form, which can be arranged as a matrix \( P \) with dimension \((M + 1) \times (N + 1)\):

\[
P = \begin{bmatrix}
p_{11} & \cdots & p_{1N} & p_{1, nil} \\
\vdots & \ddots & \vdots & \vdots \\
p_{M1} & \cdots & p_{MN} & p_{M, nil} \\
p_{nil, 1} & \cdots & p_{nil, N} & 0
\end{bmatrix} \tag{10}
\]

Matrix \( P \) comprises two parts: the inner \( M \times N \) part of it represents correspondences, and the remainder is defined to tackle outliers. If a point \( x_i \in X \) matches a point \( y_j \in Y \), then \( p_{ij} = 1 \); otherwise, \( p_{ij} = 0 \). In order to impose the constraint of one-to-one correspondences, the binary matching matrix \( P \) is subject to:

\[
\begin{cases}
\sum_{j=1}^{N+1} p_{ij} = 1, i = 1, 2, \cdots, M \\
\sum_{i=1}^{M+1} p_{ij} = 1, j = 1, 2, \cdots, N
\end{cases} \tag{11}
\]

By using matrix \( P \), the cost function can be formulated as:

\[
C(X, Y, P) = \sum_{i=1}^{M} \sum_{j=1}^{N} w_{ij} p_{ij} + 2(1 - \alpha) \sum_{i=1}^{M} \sum_{a \in N_i} \sum_{b \in N_j} p_{ij} p_{ab} \tag{12}
\]

So the PPM problem is cast into an NP-complete problem of integer quadratic programming. In this paper, we relax \( p_{ij} \) to \([0, 1]\) and use the well-known probabilistic relaxation to find the solution. With respect to the matching objective function, the gradient \( g_{ij} \) is computed by:

\[
g_{ij} = \alpha w_{ij} + 4(1 - \alpha) \sum_{a \in N_i} \sum_{b \in N_j} p_{ab} \tag{13}
\]

According to the method of probabilistic relaxation, \( p_{ij} \) is iteratively updated by:

\[
p_{ij} := p_{ij} g_{ij} / \sum_{k=1}^{N} p_{ik} g_{ik} \tag{14}
\]

Note that only one-way normalization is imposed in Eq. (14). In order to satisfy the constraints in Eq. (11), we convert the probability matrix into a doubly stochastic matrix by alternated row and column normalization, which was suggested in [14] and successfully used in [19]. Since probabilistic relaxation is a local optimal technique, a good initialization is pivotal to obtain an acceptable solution. In this work, we use the similarities evaluated by the LGSC descriptors, as shown in Eq. (5), to initialize the probabilistic matrix and then transform it into the form of doubly stochastic matrix. We set \( p_{i, nil} \) and \( p_{nil, j} \), the probability for a point matching to a dummy point, to 0.2. According to our experiments, 200 rounds of iterative update are enough to obtain a convergent solution. After the update procedure, we determine correspondences by \( p_{ij} \geq 0.6 \) in order to obtain more matching pairs.

### 3. Experiments

We conduct our experiments on both synthetic data and real-world images. As LGSC is a local structural descriptor, we compare it with the typical shape context to evaluate its performance. The comparison method is to embed the shape context into the graph-matching framework in Section 3, where the similarities of the shape context to compute \( F_\alpha(X, Y, \phi) \) and all the experimental settings remain the same. We refer to it as SC+NH for convenience. In [1], the shape context is not rotation invariant. To achieve a fair comparison, we adopt the improved measurement in [19] to compute the shape context: the mass center of a point-set is used as a reference point and the direction from a point to the mass center is employed as the positive x-axis for the local coordinate system. Our method is also systematically compared with two state-of-the-art matching approaches: the method of Zass et al. [18] and the method of Cour et al. [5]. As for the series \( \Theta \), according to our experiments, it is applicable that the size of \( \Theta \) ranges from 3 to 5 and the value of \( d_\epsilon \in \Theta \) approximately ranges from 1/4 to 1/2 of the point-set size. We set \( \Theta = \{10, 13, 16\} \) in the following experiments.

#### 3.1. Synthetic Data

In this part, we aim to investigate the influence of positional jitter and outliers quantitatively. Apart from the aforementioned comparative experiments, here we also show the discriminative power of LGSC without the support of any extra optimization. As the objective function in Eq. (6) includes two parts, we achieve this by setting \( \alpha = 1 \) in Eq.
Figure 2. Effect of varying positional jitter on matching accuracy with a given percentage of outliers. (a)-(c) show the comparison with different methods. (d)-(f) evaluate the performance of various spectral representations, where Laplace,QLaplace and Adjacency denote only the spectrum of the Laplacian matrix, the Quasi-Laplacian matrix or the adjacency matrix is used, respectively. Concatenation represents stacking these three kinds of spectra into a vector. And MSE denotes fusing the miscellaneous spectra via the technique of MSE. (a) Without outliers. (b) The percentage of outliers = 10%. (c) The percentage of outliers = 20%. (d) Without outliers. (e) The percentage of outliers = 10%. (f) The percentage of outliers = 20%.

In order to provide a baseline, the case of $\alpha = 1$ is also tested for SC+NH. Furthermore, we investigate the benefit of using multiple spectra and MSE. The comparison method is similar to that of the shape context. The synthetic data are generated as follows. We sample randomly and uniformly 40 points on the unit 2D plane. The matched point-set without positional jitter and outliers is generated by applying a random rigid transform in which the random transform parameters are uniformly distributed. And the value ranges are set to $-1 \leq t_x, t_y \leq 1$ and $-\pi \leq \theta \leq \pi$, where $\theta$ is the rotation angle and $t_x, t_y$ denote the offset in X and Y axes, respectively. We create positional jitter by adding Gaussian noise to the position of the matched point-set, in which the mean is zero and the standard deviation is defined as the fraction of the average closest distance of the matched point-set. The outliers are generated by adding a number of random points locating in the range of the matched point-set. Figs. 2(a)-(f) show the accuracy as a function of varying the noise level with a given number of outliers. The accuracy on each noise level is averaged over 100 independent experiments.

We first pay attention to the results shown in Figs. 2(a)-(c). The proposed approach performs the best in the overall experiments. Our approach yields a remarkable improvement over the other three alternatives, especially in the presence of outliers and greater noise. From the results we can conclude that LGSC provides a good estimation on correspondences, which implicitly shows the discriminative power of LGSC. The cases of $\alpha = 1$ reveal two facts: (1) LGSC outperforms the shape context by a good margin and it can achieve meaningful results even under the circumstance of outliers and greater noise, which demonstrates that LGSC is a relatively robust spectral representation for point patterns; (2) the combination with other optimal strategies is necessary to enable satisfied matching results. Then we turn to the results shown in Figs. 2(d)-(f). We can observe that there is no obvious difference between those single-spectrum representations. The straightforward concatenation does not make sense in our experiments, the results from which are even worse than those from a single spectrum. It indicates that the naive concatenation cannot effectively take advantage of the complementary attribute of multiple spectra. On the other hand, the fused multiple-spectra representation encoded by MSE receives a consistent improvement over all the other compared methods, and hence we can draw a conclusion that finding the intrinsic embedding via MSE can facilitate the utilization of multiple-spectra representations.

3.2. Real-World Images

In the experiments on real-world images, we perform feature point matching on the CMU/VASC house sequence which has been widely used as a canonical test set in PP-M, and compare the results with those computed from the alternatives. In order to compute the accuracy easily, we label 40 corresponding feature points manually across all the frames. The hand-labeled feature points are illustrated in
Fig. 4(a). The 0th frame, which is adopted as the template image, is tested against the 10th frame \((i = 1, 2, \ldots, 11)\). Obviously the difficulty of matching increases when the sequence gap becomes large.

To analyze the influence of outliers more precisely, we categorize our comparative experiments into three groups: (1) The template image and the matched image contain 40 feature points, respectively. That is, there are no outliers in the image pairs. (2) Randomly delete 5 points from the template image. That is, there are 5 outliers in the matched frame. (3) 5 points are randomly deleted from the template image and the matched image, respectively, to make each image contain 5 outliers \(i.e.,\) the number of correspondences is 30. For the latter 2 groups of experiments, as the outliers are randomly generated, the precisions are averaged over 100 independent trials.

Figs. 3(a)-(c) show the accuracy curves with respect to sequence gap in different experimental settings. Figs. 4(b)-(d) demonstrate some matching results obtained from our method, where green lines indicate correct matches and red lines indicate incorrect ones. In general, the relative performance of each algorithm is roughly identical to that in the experiments on synthetic data. The method of Cour et al. works the worst, although it achieves almost perfect results when the disparity is trivial. Our method still outperforms the other three methods. It should be noted that the margin of improvement is not as much as that in the former experiment. The main reason is that the influence of disparity is far less than that of random positional jitter, which is easily deduced from the result illustrated in Fig. 3(a).

4. Conclusion

In this paper, we investigated the problem of spectral representation for point patterns. We have proposed a novel local spectral descriptor, namely LGSC, which is essentially different from the previous global representation approaches. Multiview spectral embedding has been used to fuse different descriptors to yield the final representation. Extensive experiments as well as comparisons with the state-of-the-art spectral methods have demonstrated the effectiveness of the proposed method.

Acknowledgements

Jun Tang gratefully acknowledges the support of National Natural Science Foundation of China(Grant No. 61127127) and Natural Science Foundation of Anhui Provincial Education Department (Grant No. 2011KJA008).

References


