Pattern Recognition Letters 000 (2016) 1–7

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Co-spectral for robust shape clustering☆

Song Bai, Zhiyong Liu, Xiang Bai*

School of Electronic Information and Communications, Huazhong University of Science and Technology, 430074 Wuhan, China

A R T I C L E I N F O

Article history:
Available online xxx

Keywords:
Shape clustering
Co-training
Shape context
Skeleton path

A B S T R A C T

Shape clustering is a difficult visual task due to large intra-class variations and small inter-class variations induced by shape articulation, rotation, occlusion, *etc*. To tackle this problem, we attempt to leverage the complementary nature among features of different statistics (*e.g.*, skeleton-based descriptors and contour-based descriptors) for robust clustering. In this paper, a similarity fusion framework based on spectral analysis is proposed. The proposed method, which we call co-spectral, is a spectral clustering algorithm. It has two inborn merits for shape clustering: (1) it can automatically make use of the complementarity of various shape similarities based on a co-training framework; (2) it does not require shape representations to be vectors. Co-spectral is evaluated on several popular shape benchmarks. The experimental results demonstrate that co-spectral outperforms the state-of-the-art algorithms by a large margin.

1. Introduction

Shape clustering [16] is a fundamental problem in pattern recognition with applications to shape matching [7,18], recognition [13], retrieval [5,6] and classification [4]. Given a collection of shapes $\mathcal{S} = \{s_1, s_2, \ldots, s_N\}$ where N denotes the number of shapes, the aim of shape clustering is to divide all the shape instances into K clusters $\mathcal{C} = \{c_1, c_2, \ldots, c_K\}$ according to a pre-defined similarity measure.

The key issues in shape clustering lie in two aspects. First, shape data is usually not represented by vectorial features. Instead, tree [8], matrix [13,27] and string [17] are more widely-used in shape analysis. Hence, some clustering algorithms that require vectorial representations as inputs, *e.g.*, K-means [28], are not applicable directly. Second, there are large intra-class variations and small inter-class variations, like articulation, rotation, occlusion, *etc*. Nevertheless, it is difficult to design generic and discriminative shape features to handle all the common deformations. In most cases, a certain descriptor only focuses on a specific geometric structure of shapes. For example, Shape Context (SC) [13], as a representative contour-based descriptor, works well with rigid shapes. In contrast, Inner Distance Shape Context (IDSC) [27], which replaces the Euclidean distance used in SC with geodesic distance, is better at dealing with articulated shapes.

To address the above issues, we introduce spectral clustering as an elegant mathematical tool for the shape clustering task. Spectral clustering operates on a weighted affinity graph, where the nodes in the graph represent the data points and the edge weights measure the similarities between two adjacent data points. Therefore, it can deal with arbitrary types of input data, as long as the pairwise similarities are avaliable. This property is crucial for shape clustering, since many shape similarity measures have no vectorial representations. Moreover, by exploiting the properties of Laplacian of the affinity graph, spectral clustering can capture the main patterns across categories and diminish the negative influences of noisy attributes. As a result, spectral clustering is more robust to shape outliers.

Considering the limitation of using only one type of similarity measure, it can be expected that an effective method which integrates multiple complementary similarities can boost the performance of shape clustering remarkably. Nevertheless, it is very difficult to fuse multiple descriptors in shape clustering, since no prior or extra information can be used to judge the discriminative power of different features in such an unsupervised task. To our best knowledge now, no methods have properly addressed the feature fusion issue in the shape clustering task.

In this paper, based on spectral clustering, a co-trained spectral clustering algorithm is presented. The proposed method inherits the nice properties of spectral clustering as introduced above. Similarity fusion is automatically done based on the co-training framework [41] that exploits the complementary nature among different shape descriptors. Moreover, a density-based seed is exerted to co-trained spectral clustering in order to avoid local minima and provide stable performances consistently. At last, since co-training is not guaranteed to converge as suggested in [23], we propose a simple yet effective consensus-based voting scheme to aggregate

☆ This paper has been recommended for acceptance by Gabriella Sanniti di Baja.

* Corresponding author. Fax: +86 87543236.
E-mail address: xbai@hust.edu.cn (X. Bai).

http://dx.doi.org/10.1016/j.patrec.2016.07.014
0167-8655/

the clustering results of different iterations without impairing the performance too much.

The rest of the paper is organized as follows. We give a brief review of shape clustering algorithms in Section 2. Our proposed method is introduced in Section 3. The experimental evaluations and comparisons are conducted in Section 4. Conclusions are drawn in Section 5.

2. Related work

In recent years, many algorithms are proposed to address the shape clustering task. They can be coarsely divided into two categories: contour-based methods and skeleton-based methods.

In [24], a new similarity measure between a single shape and a shape group is defined, and it serves as the basis for a soft K-means like framework to enable robust clustering. Clustering in [35] is achieved by building on a differential geometric representation of shapes and geodesic lengths as shape metrics. Yankov et al. [38] take isomap clustering using a rotationally invariant metric, which can detect the intrinsic nonlinear embedding in which the shape examples reside. In [29], the elastic properties of shape boundaries are investigated and clustering is done using dynamic programming based on the elastic geodesic distance.

Demirci et al. [19] construct a medial axis graph for shape silhouettes. For every two graphs, a many-to-many correspondence between graph nodes [20] is established. These correspondences are used later to recover the abstracted medial axis graph. An information theoretic framework is presented in [37]. It attempts to learn a mixture of tree unions that best accounts for the observed samples using a minimum encoding criterion. In [21], a game theoretic clustering approach is developed, which can simultaneously learn categories from examples and the similarity measures related to them. Shen et al. also propose a skeleton-based clustering algorithm in [33] on the common structure skeleton graph (CSSG), which can discover intrinsic structural information of shapes belonging to the same cluster.

Most aforementioned shape clustering methods are either contour-based or skeleton-based. There are also some methods that are not descriptor-based, such as Laplacian spectrum [30] or minimum spanning trees [40]. The method proposed in this paper is descriptor-based. It combines complementary shape features in a unified framework, thus providing much better performances.

3. Proposed method

3.1. Similarity measure

Contour-based descriptors and skeleton-based descriptors are two main streams in shape analysis. Contour-based descriptors deliver the distribution of shape boundary points. They are more stable to affine transformations. By contrast, skeleton-based descriptors convey the structure of object skeletons, thus are more adequate in non-ridge analysis. The complementary nature between them has been extensively testified in shape recognition [9,32].

Two similarity measures are implemented in this paper, *i.e.*, Shape Context (SC) [13] and Skeleton Path [8], with the former one as a representative contour descriptor and the latter one as a skeleton descriptor.

Shape contetxt. Given a certain shape $s_q \in \mathcal{S}$, we extract its outer contour represented by n discrete points $\nu = \{\nu_1, \nu_2, \dots, \nu_n\}$ in the plane. Around each point $\nu_i \in \nu$, we construct a log-polar coordinate space with 12 bins for dividing angle space and 5 bins for dividing radius space. As a consequence, the k-th element in the shape context histogram of ν_i is computed via

$$p_i(k) = |\{\nu|\nu \in bin(k), \nu \neq \nu_i, \nu \in \nu\}|, \tag{1}$$

where |.| measures the cardinality of the input set.

Let $q = \{q_1, q_2, \dots, q_n\}$ and $p = \{p_1, p_2, \dots, p_n\}$ denote two sets of shape context histograms of s_q and s_p respectively. Their point-wise matching cost is measured using χ^2 distance as

$$C(p_i, q_j) = \frac{1}{2}\sum_k \frac{[p_i(k) - q_j(k)]^2}{p_i(k) + q_j(k)}. \tag{2}$$

After obtaining the matching cost $C(p_i, q_j)$ for all pairs of elements $p_i \in p$ and $q_j \in q$, Hungarian algorithm is applied to find the optimal correspondence as

$$H(\pi) = \arg\min_\pi \sum_i^n C\big(q_i, p_{\pi(i)}\big), \tag{3}$$

where π is a permutation indicating that the matching is one-to-one.

Skeleton path. We implement skeleton path proposed in [8] as the second similarity measure, which is based on skeleton analysis. In this subsection, we give a brief review of skeleton path. One can refer to the study in [8] for more details if necessary.

Assuming that the skeleton curve is one pixel wide, three kinds of points are defined: end point, junction point and connection point. The end point is defined as the skeleton point owning one adjacent point. The skeleton point with more than two adjacent points is named a junction point, and the rest are connection points. To build the skeleton graph, both end points and junction points are chosen as the nodes of the graph. The edges between two adjacent nodes are the skeleton branches between them.

Given a pair of nodes u, ν in a skeleton graph, the skeleton path $P(u, \nu)$ is defined as the shortest path along the skeleton graph between u and ν. Let $P(u_q, \nu_q)$ and $P(u_p, \nu_p)$ represent two skeleton paths from shape s_q and s_p respectively. Their path distance is defined as

$$pd(P(u_q, \nu_q), P(u_p, \nu_p)) = \sum_{i=1}^M \frac{(r_{qi} - r_{pi})^2}{r_{qi} + r_{pi}} + \alpha\frac{(l_q - l_p)^2}{l_q + l_p}, \tag{4}$$

where r_{qi} and $r_{pi} (0 \leq i \leq M)$ represent the radii of the maximal disks centered at the M sample points of skeleton paths $P(u_q, \nu_q)$ and $P(u_p, \nu_p)$ respectively. l_q and l_p are their lengths, and α is the weighting factor.

Assume that the skeleton graph of s_q, denoted as $E_q = \{e_{q_1}, \dots, e_{q_{nq}}\}$, has nq end points, and the skeleton graph of s_p, denoted as $E_p = \{e_{p_1}, \dots, e_{p_{np}}\}$, has np end points. The pairwise distance between each pair of end points e_{q_i} and e_{p_j}, referred as $ed(e_{q_i}, e_{p_j})$, is computed via Optimal Subsequence Bijection (OSB) as in [8]. Then we can get a $nq \times np$ distance matrix $\Phi(E_q, E_p)$:

$$\begin{pmatrix} ed(e_{q_1}, e_{p_1}) & ed(e_{q_1}, e_{p_2}) & \dots & ed(e_{q_1}, e_{p_{np}}) \\ ed(e_{q_2}, e_{p_1}) & ed(e_{q_2}, e_{p_2}) & \dots & ed(e_{q_2}, e_{p_{np}}) \\ \dots & \dots & \dots & \dots \\ ed(e_{q_{nq}}, e_{p_1}) & ed(e_{q_{nq}}, e_{p_2}) & \dots & ed(e_{q_{nq}}, e_{p_{np}}) \end{pmatrix} \tag{5}$$

After applying the Hungarian algorithm on $\Phi(E_q, E_p)$, we can get the optimal correspondence $\varphi : \{e_{q_1}, \dots, e_{q_{nq}}\} \to \{e_{p_1}, \dots, e_{p_{np}}\}$ between end points in E_q and those in E_p. Thus the matching cost, also the dissimilarity between two shapes, is obtained.

3.2. Co-trained spectral clustering

As a representative of graph-based clustering algorithms, spectral clustering exploits the properties of Laplacian of the affinity graph. Spectral clustering algorithms are usually divided into two

Please cite this article as: S. Bai et al., Co-spectral for robust shape clustering, Pattern Recognition Letters (2016), http://dx.doi.org/10.1016/j.patrec.2016.07.014