Image and Vision Computing 33 (2015) 57–67

Contents lists available at ScienceDirect

Image and Vision Computing

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

Feature correspondence based on directed structural model matching☆

Xu Yang, Hong Qiao, Zhi-Yong Liu *

State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 4 September 2013
Received in revised form 27 April 2014
Accepted 2 November 2014
Available online 11 November 2014

Keywords:
Feature correspondence
Directed structural model
Graph matching

ABSTRACT

Feature correspondence lays the foundation for many tasks in computer vision and pattern recognition. In this paper the directed structural model is utilized to represent the feature set, and the correspondence problem is then formulated as the structural model matching. Compared with the undirected structural model, the proposed directed model provides more discriminating ability and invariance against rotation and scale transformations. Finally, the recently proposed convex–concave relaxation procedure (CCRP) is generalized to approximately solve the problem. Extensive experiments on synthetic and real data witness the effectiveness of the proposed method.

1. Introduction

As a fundamental problem in computer vision and pattern recognition, feature correspondence plays an important role in many tasks such as image retrieval, object recognition, 3D reconstruction, and bio-informatics. Although the correspondence using only appearance descriptors such as SIFT descriptor [24], shape context feature [1] or bag-of-words model [20] got good results in some tasks such as object detection and image classification [32,10,35,36], recently much effort has been devoted to the incorporation of spatial information into the appearance cue [6,33,25,2,15,26]. Consequently, the feature correspondence can be in general formulated as a combination of a unary term related to the appearance similarity and a pairwise term describing the spatial consistency.

Intuitively, the pairwise constraints could improve the correspondence performance. For example, if we add the distance constraint, then a pair of feature points close to each other in one set is less likely to be assigned to the points far away from each other in another set. However, researchers [6] showed that the appearance only based methods performed comparably or even better than the structural ones. In fact, the performance of pairwise constraints depends largely on the structural model constructed, and an unstable descriptor may result in bad performance in practice. For instance, the same object in different images may appear in different scales, orientations, and suffers from other types of distortions. It is thus far from easy to construct a stable structural model invariant to these geometric transformations. The distance descriptor is robust against rotation, but often vulnerable to scale and other types of transformations. The orientation descriptor is robust against scale, but not rotation.

In this paper, we propose a robust directed structural model to tackle the equal-sized correspondence problem. By representing the feature set by the directed structural model or equivalently the directed graph, the correspondence is formulated as a directed graph matching and then approximately solved by generalizing the recently proposed convex–concave relaxation procedure (CCRP).

1.1. Related works

Feature correspondence mainly consists of two parts, model construction and optimization algorithm. We give a brief review of the related works in literature from the two aspects.

The proposed model is related to [33,19,27,28,9,3,37] which models utilize both the distance and orientation descriptors, though in different ways. Considering the distance descriptor, some models [33,19,27,9,37] are sensitive to the scale transformation by directly taking the distance as the descriptor. In [28,3] they normalize the distance by the longest distance between features, which makes it invariant to scale transformation, but the performance may deteriorate sharply when false features exist, because they tend to affect the global measurement. By contrast, the proposed distance descriptor takes a local measure as the normalizer, which is more robust than the global normalization. Considering the orientation descriptor, many existing models utilize or equivalently utilize the angle between an edge and the horizontal axis as the orientation descriptor. Such descriptors are sensitive to the rotation transformation. In [27] they utilize the angle between dominant orientation of one feature and related edge as the descriptor, which however is limited to the feature types with local orientation like SIFT [24]. By contrast, the proposed orientation descriptor based on the object orientation is invariant to rotation.

☆ This paper has been recommended for acceptance by Bodo Rosenhahn.

* Corresponding author. Tel.: +86 10 6263 6367.
E-mail addresses: xu.yang@ia.ac.cn (X. Yang), hong.qiao@ia.ac.cn (H. Qiao), zhiyong.liu@ia.ac.cn (Z.-Y. Liu).

http://dx.doi.org/10.1016/j.imavis.2014.11.001
0262-8856/

There exist many graph matching algorithms for the correspondence problem combining both appearance and structural cues, about which the readers are advised to refer to [8] for a comprehensive survey. Generally speaking, the graph matching problem is a typical NP-hard problem, which makes the approximation necessary in practical applications. Though some approximate matching algorithms [14,31] could deal with the directed graphs, they often suffer from high computational and storage complexity. Some other state-of-the-art algorithms recently proposed, including the spectral method [19,9,34,7,29], semi-definite based method [30] and path following methods [38,39], are applicable only on undirected graphs. As an extension of the path following algorithm, the CCRP [23,22] is generalized to deal with the directed problem.

The directed structural model and the objective function are proposed in detail in Section 2, and the optimization algorithm is given in Section 3. Following the extensive experiments on both synthetic and real data in Section 4, Section 5 finally concludes the paper.

2. Directed structural model

In this section, we will first present the directed structural model and then give the objective function for the correspondence problem.

2.1. Distance descriptor

A feature g_i in a feature set $G = \{g_i\}_{i=1}^{N}$ is a key point with its spatial coordinates $l_i \in \mathbb{R}^2$ and the appearance descriptor $f_i \in \mathbb{R}^{d_f}$ around it. For instance, by utilizing a 128-dimensional SIFT histogram abstract from the patches around each key point as the appearance descriptor, there is $d_f = 128$.

Distance is probably the most commonly used structural descriptor between key points, due to its invariance to rotation. However, the distance itself is vulnerable to the scale transformation, which consequently makes a normalization procedure necessary. Commonly used normalization should be the ones normalized by the maximal or average distance between feature points. However, the performance of such normalization may deteriorate greatly when encountering outliers. To overcome this drawback, the distance descriptor is given by,

$$a_{ij}^{dis} = \exp\left(-\frac{\|l_i - l_j\|^2}{\max\limits_{j=1\cdots N}\|l_i - l_j\|^2}\right). \tag{1}$$

Generally $a_{ij}^{dis} \neq a_{ji}^{dis}$, which means the distance descriptor is directed. And the descriptor is less affected by the outliers than the global normalization, as shown in Fig. 1.

2.2. Orientation descriptor

The orientation descriptor is robust to the scale transformation but in general sensitive to rotation. To overcome this problem, a fixed orientation $\overline{d}$ with respect to object rotation is proposed as follows,

$$\overline{d} = \sum_{\substack{i=1\cdots N\\ l_i \neq \overline{l}}} \frac{l_i - \overline{l}}{\|l_i - \overline{l}\|}, \tag{2}$$

where

$$\overline{l} = \frac{1}{N}\sum_{i=1\cdots N} l_i$$

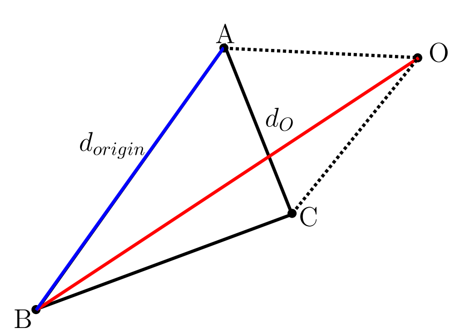
is treated as the object center. A simple illustration of the object orientation $\overline{d}$ is given in Fig. 2(a). It is observed that the two 'Houses' retain the same object orientation under scale and rotation transformations.

Based on $\overline{d}$, the orientation descriptor is given by

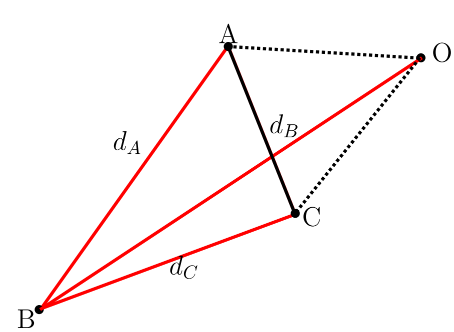
$$a_{ij}^{ori} = \begin{cases} \frac{1}{\pi} arccos\left(\frac{(l_i - l_j)}{\|l_i - l_j\|}\frac{\overline{d}}{\|\overline{d}\|}\right) & \text{if } i \neq j \wedge \|\overline{d}\| \neq 0 \\ 0 & \text{Otherwise.}\end{cases}, \tag{3}$$

It is the angle between $\overline{d}$ and the link between features g_i and g_j, normalized by π. Generally $a_{ij}^{ori} \neq a_{ji}^{ori}$, which means that the orientation descriptor is directed. Fig. 2(b) gives an illustration of the orientation descriptor.

Fig. 3 gives a simple comparison between the orientation descriptors in directed and undirected models, where G' is gotten by a reflection transformation over G. The object orientation of G (G') is denoted by $\overline{d}$ $\left(\overline{d'}\right)$. In the undirected model, the acute angle between edge AD and $\overline{d}$, i.e. α, is utilized as the orientation descriptor a_{AD}. Then $a_{AD} = a_{DA} = a_{A'D'} = a_{D'A'} = \alpha$, and it can be noticed that this descriptor makes no distinction between points A and D. If AB and CD are parallel, there may be multiple correspondence solutions, i.e., $[ABCD]$ may be assigned to $[A'B'C'D']$ or equivalently to $[D'C'B'A']$. And the detailed explanation is given in A. In the directed model, there are $a_{AD} = a_{A'D'} = \alpha$ and $a_{DA} = a_{D'A'} = \beta$, which imply that there is a unique correspondence solution even in the case that AB and CD are parallel. Thus, the orientation descriptor in the directed model provides a more discriminating ability.



(a) Normalization by global distance maximum.



(b) Normalization by local distance maximum.

Fig. 1. Effect of the proposed distance descriptor. When using global normalization on the left, an outlier O replaces the original maximum distance d_{origin} by d_O and thus affects all the normalized distance descriptors. While the proposed distance descriptor is normalized by the local distance maximum. Thus d_A and d_C are the same as in the situation without O, and only d_B is different from the original d_{BA}.