Neurocomputing 122 (2013) 193–197

Contents lists available at ScienceDirect

Neurocomputing

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

Partial correspondence based on subgraph 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 25 March 2013
Received in revised form
30 May 2013
Accepted 12 June 2013
Communicated by S. Choi
Available online 4 July 2013

Keywords:
Feature correspondence
Structural model
Subgraph matching
GNCGCP

ABSTRACT

Exploiting both appearance similarity and geometric consistency is popular in addressing the feature correspondence problem. However, when there exist outliers the performance generally deteriorates greatly. In this paper, we propose a novel partial correspondence method to tackle the problem with outliers. Specifically, a novel pairwise term together with a neighborhood system is proposed, which, together with the other two pairwise terms and a unary term, formulates the correspondence to be solved as a subgraph matching problem. The problem is then approximated by the recently proposed Graduated Non-Convexity and Graduated Concavity Procedure (GNCGCP). The proposed algorithm obtains a state-of-the-art accuracy in the existence of outliers while keeping $\mathcal{O}(N^3)$ computational complexity and $\mathcal{O}(N^2)$ storage complexity. Simulations on both the synthetic and real-world images witness the effectiveness of the proposed method.

1. Introduction

Feature correspondence, aiming to find a reasonable assignment between local feature sets of different images, is one fundamental problem in computer vision and pattern recognition and is extensively applied in many tasks including object detection and recognition, camera self-calibration, 3D reconstruction and tracking. The extracted local features can hardly be used by most current objective object recognition, 3D reconstruction algorithms, unless they are put into correspondence. Though the correspondence problem has been studied for decades, it is still a challenging problem.

Starting from using only the appearance descriptor such as SIFT descriptor [1], bag-of-words model [2] which get good results in some computer vision tasks, recently much more effort in this area has been devoted to the incorporation of structural information into the appearance cues. They thus formulate the correspondence problem as a combination of the unary term and the pairwise term which relate to the appearance similarity and geometric consistency, respectively. Inspite of some controversy over the effectiveness of the structural constraints [3,4], some most recent works [5,6] which pay more attention to the robustness and distinctive ability of the structural model witness obvious performance improvements on some benchmark datasets and reconfirm the usefulness of structural cues.

However, the correspondence is still a challenging task when there exist outliers, which are inevitable in many practical applications, e.g. when matching an object in complex background or part of the object is occluded. Such a correspondence problem with unequal feature points in two images in the presence of outliers is denoted as partial correspondence. Many current partial correspondence methods deteriorate greatly as the outlier number increases [7] and some other methods [8], especially the adjacency matrix based methods [6,9–11] cannot even deal with the partial problem.

In this paper, we propose a novel partial feature correspondence method, with two main contributions listed below:

1. A pairwise term together with a neighborhood system which describes the coherence of key points is proposed. The coherence prior means that the adjacent points more likely locate in the same image region, either in object region or background region.
2. An effective and efficient combinatorial optimization framework named Graduated-NonConvexity-and-Graduated-Concavity-Procedure (GNCGCP) [12] is introduced to solve the partial correspondence problem.

In addition, two pairwise terms based on our previously proposed directed distance and direction descriptors [6] with remarkable distinctive ability are adapted to the partial situation. Together with the appearance term, coherence term and GNCGCP, the whole scheme achieves state-of-the-art performance, and at the same time enjoys $O(N^3)$ computational complexity and $O(N^2)$ storage complexity, where N is the key point number.

* 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).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.06.031

There has been other literature on the partial correspondence issue [5,7,13]. The proposed method differentiates mainly from two aspects: (1) the directed structural model based objective function; and (2) the GNCGCP based optimization algorithm.

2. Proposed method

2.1. Objective function

Given two image feature sets $G \in \mathbb{R}^{m\times s}, H \in \mathbb{R}^{n\times s}, m \le n$, where m, n are the key point numbers and s is the appearance descriptor dimension. The partial correspondence problem is formulated as finding a good assignment, or equivalently a partial permutation matrix $P \in \mathbb{R}^{m\times n}$ between the key points in G and H to get a correspondence criterion $F(P, G, H)$ minimized. In this paper, we utilize

$$F(P,G,H) = w^{app}F^{app}(P) + w^{dis}F^{dis}(P) + w^{dir}F^{dir}(P) + w^{coh}F^{coh}(P),$$
$$\text{s.t. } P \in \mathcal{P}, \mathcal{P} = \left\{ P | P_{ij} = \{0,1\}, \sum_{j=1}^{n} P_{ij} = 1, \sum_{i=1}^{m} P_{ij} \le 1, \forall i,j \right\} \tag{1}$$

where w^{app}, w^{dis}, w^{dir}, w^{coh} are weights of the four terms.

Next we will first introduce a neighborhood system, and then give the explanation for each term in (1).

When using the complete neighborhood system [6,9] where all the key points are mutually connected, the connection between two points far away from each other often brings in noise rather than makes the structural model robust, especially in the non-rigid correspondence. Instead we utilize a local neighborhood system where a key point i is connected to its K nearest neighboring points $\mathcal{N}_K(i)$, and is considered as infinitely far away from all the other points. A neighborhood matrix $N \in \mathbb{R}^{n\times n}$ is given as

$$N_{ij} = \begin{cases} 1/K & \text{if } j \in \mathcal{N}_K(i), \\ 0 & \text{if } j \notin \mathcal{N}_K(i). \end{cases} \tag{2}$$

where i, j are two key points. The local neighborhood system improves the correspondence performance as will be illustrated in Section 3. Based on the system, several pairwise terms are introduced or modified below.

$F^{app}(P)$ is a unary term which measures the appearance similarity between feature sets. We define it as

$$F^{app}(P) = \mathrm{tr}(CP^T) \tag{3}$$

where $\mathrm{tr}(\cdot)$ is the matrix trace, $C \in \mathbb{R}^{m\times n}$ is the cost matrix where c_{ij} denotes the dissimilarity between the two appearance descriptors g_i and h_j normalized by the largest value in C. SIFT descriptor, shape context descriptor, bag-of-words model and other appearance descriptors could be adopted in this term depending on their typical uses, and the dissimilarity measures vary accordingly, where the chi-square distance for histogram based appearance descriptor such as SIFT and the Euclidean distance are two common choices.

$F^{dis}(P)$ is a pairwise term which measures the geometric consistency from the distance aspect. We define it as

$$F^{dis}(P) = \|A_G^{dis} - PA_H^{dis}P^T\|_F^2 \tag{4}$$

where $\|\cdot\|$ is the Frobenius matrix norm defined as $\|A\|_F = \sqrt{\sum_i \sum_j A_{ij}^2} = \sqrt{\mathrm{tr}(A^TA)}$. $A_G^{dis} \in \mathbb{R}^{m\times m}, A_H^{dis} \in \mathbb{R}^{n\times n}$ are the adjacency matrices describing the distance attributes of feature sets defined as

$$a_{\overrightarrow{ij}}^{dis} = \begin{cases} \exp\left(-\frac{\|l_i - l_j\|^2}{\max_{j\in\mathcal{N}_K(i)}\|l_i - l_j\|^2}\right) & \text{if } j \in \mathcal{N}_K(i) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where l_i, l_j are the locations of key points i and j respectively. The distance descriptor is directed since generally $a_{\overrightarrow{ij}}^{dis} \ne a_{\overrightarrow{ji}}^{dis}$, which makes it more distinctive between i and j [6], and the descriptor normalized by the local distance maximum makes itself less affected by the outliers compared with the traditional normalization [6].

$F^{dir}(P)$ is a similar pairwise term as $F^{dis}(P)$ but from the orientation aspect. We define it as

$$F^{dir}(P) = \|A_G^{dir} - PA_H^{dir}P^T\|_F^2 \tag{6}$$

where $A_G^{dir} \in \mathbb{R}^{m\times m}, A_H^{dir} \in \mathbb{R}^{n\times n}$ are the adjacency matrices for the direction attribute defined as

$$a_{\overrightarrow{ij}}^{dir} = \begin{cases} \frac{1}{\pi}\arccos\left(\frac{(l_i - l_j)}{\|l_i - l_j\|}\frac{\overline{d}}{\|\overline{d}\|}\right) & \text{if } j \in \mathcal{N}_K(i) \wedge \|\overline{d}\| \ne 0 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

where $\overline{d}$ is denoted as the object orientation which is a relatively fixed direction with respect to the object rotation. When matching two objects with clear background, an effective definition is

$$\overline{d} = \sum_{\substack{i=1\cdots n \\ l_i \ne \overline{l}}} \frac{l_i - \overline{l}}{\|l_i - \overline{l}\|} \tag{8}$$

where $\overline{l} = (1/n)\sum_{i=1\cdots n} l_i$ can be regarded as the center of the object [6]. While with complex background or given the prior that there is rare rotation between two objects, e.g. Chinese character matching, it is better to utilize horizon direction instead. Similarly to $a_{\overrightarrow{ij}}^{dis}$, $a_{\overrightarrow{ij}}^{dir}$ is also a directed descriptor which makes it more robust and distinctive [6].

$F^{coh}(P)$ is a pairwise penalty term which penalizes the correspondence status difference between neighboring key points. That means the neighboring key points should locate in the coherent region—both in the object region or both in the background [5]. We define it as

$$F^{coh}(P) = -\|PN_HP^T\|_F^2 \tag{9}$$

where N_H is the neighborhood matrix of H given by (2). Minimizing this term means that when a key point in H is selected then more of its neighboring points are preferred to be selected.

2.2. Feature correspondence algorithm

The problem (1) is a combinatorial optimization problem which could be solved by the graph matching algorithms [14]. The feature points could be viewed as the vertices of the graph model and the pairwise relations describe the edge attributes. Then the unary term and pairwise term measure the similarity of vertices and consistency of edges, respectively. However, the above problem is an NP-hard problem with factorial complexity [10]. To make the problem computationally tractable, some approximations are necessary, for which a comprehensive review is referred to, e.g., [15].

Here we will adopt the recently proposed Graduated Non-Convexity and Graduated Concavity Procedure (GNCGCP) to minimize (1). The GNCGCP was proposed as a general framework for the discrete optimization problem over the set of partial permutation matrices $\mathcal{P}$. It has been proved to realize exactly the Convex–Concave Relaxation Procedure [9,10,16] but in a much simpler manner; it does not involve explicitly the convex or concave relaxation functions, which are typically difficult to construct [7,9–11] and greatly hinder the real application of CCRP. The GNCGCP exhibited competitive or even better accuracy than traditional CCRP, and meanwhile enjoys a low computational and storage complexity as CCRP. Since GNCGCP does not need convex