Pattern Recognition 65 (2017) 82–96

Contents lists available at ScienceDirect

Pattern Recognition

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

Local part chamfer matching for shape-based object detection

Qian Yu[a], Hui Wei[a,b,*], Chengzhuan Yang[a]

[a] Laboratory of Cognitive Model and Algorithm, School of Computer Science, Fudan University, No. 825 Zhangheng Road, Shanghai 201203, PR China
[b] Shanghai Key Laboratory of Data Science, School of Computer Science, Fudan University, No. 825 Zhangheng Road, Shanghai 201203, PR China

A R T I C L E I N F O

Keywords:
Chamfer matching
Shape matching
Shape detection
Shape-based object detection

A B S T R A C T

Chamfer matching is one of the elegant and powerful tools for shape-based detection in cluttered images. However, the chamfer matching methods, including oriented chamfer matching (OCM) and directional chamfer matching (DCM), tend to produce bad detections due to deformation of object shapes and cluttering in the scene. To improve detection accuracy of these chamfer matching methods, we propose local part oriented chamfer matching (LPOCM) and local part directional chamfer matching (LPDCM). First, shape templates and discriminative contour fragments are learned, and then a shape representation is built using a Markov random field (MRF). Finally, the template detection in an input image is formulated as an inference in the MRF. Experimental results for benchmark datasets including ETHZ Shape Classes, INRIA Horses and Weizmann Horses clearly demonstrate that the proposed LPOCM and LPDCM significantly improve the detection accuracy of OCM and DCM without sacrificing much time efficiency.

1. Introduction

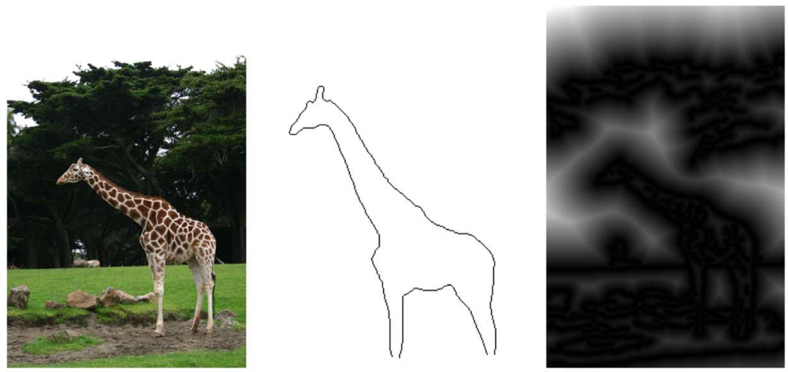
The classical chamfer matching was originally proposed by Barrow et al. [2] as a technique for finding an object similar to template in a cluttered image. A modified version, Hierarchical Chamfer Matching [4], is performed in an image pyramid. The classical chamfer matching is a powerful tool in many computer vision tasks, demonstrating speed, robustness to noise, and invariance to position, scale and rotational changes. However, it disregards the important orientation of edge pixels. In response to this limitation, Shotton et al. [24] proposed oriented chamfer matching (OCM) and Liu et al. [14] proposed directional chamfer matching (DCM), both of which are modifications of classical chamfer matching that consider the edge-pixel orientations to improve detection accuracy.

During the classical chamfer matching, we first extracted the edge of the input image and obtained the distance transformation map as shown in Fig. 1(c) where the lighter the pixel, the longer the distance from the pixel to its nearest edge pixel. Next, we slid the template across the edge map. For each sliding, we computed distance from each pixel in template to the closest edge in the input image by using the distance transformation map and then average the distances for all pixels in template to calculate the chamfer distance. The bounding-box of the template at each location is considered a detection. Note that many detections with high chamfer distances must be discarded because their high chamfer distances reflect low similarity between the template and the object. The intersection-over-union ratio (IoU) between the detection and the ground-truth bounding-box is usually employed to determine positive and negative detections. Due to deformation of object shapes and cluttering in the scene, many detections suffer from unsatisfactory IoUs, bad template localization and high chamfer distances. Fig. 1 illustrates these drawbacks through example. Supposing the template scale fixed, we slid the giraffe template to the red template's location in Fig. 1(d). The template localization appears decent since the template and the giraffe share a large degree of correct overlap. However, the giraffe in the input image presents a large local deformation with respect to the template, especially in head part. As a result, not only is the DCM distance high, but also the IoU between the detection (the dashed green rectangle) and the ground-truth bounding-box (the solid blue rectangle) is unsatisfactory. Fig. 1(e) shows the best DCM detection. Although this detection has the lowest DCM distance, the template localization and detection IoU are poor. Generally, these drawbacks appear in many DCM (and OCM) detections, which would lower detection accuracy.

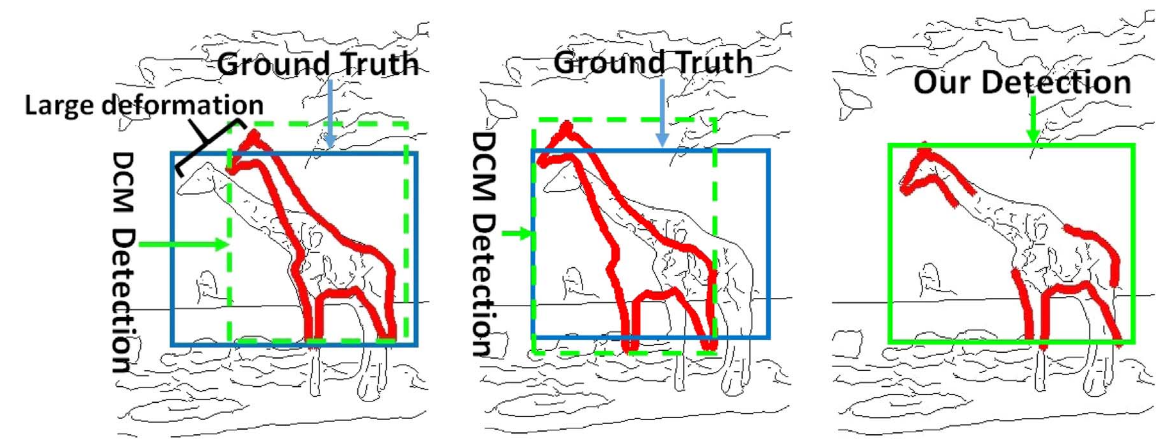
To at least partially overcome the above drawbacks, local part DCM (LPDCM) and local part OCM (LPOCM) are proposed for shape-based object detection in cluttered images. Our general idea for LPDCM and LPOCM is to divide the model shape into contour fragments (or shape parts), and then find candidates in the input image. Finally, the candidate combination with the correct geometric constraint and low chamfer distance result in final object detection. As shown in Fig. 1(e), four contour fragments were correctly located, with low DCM distances and detection close to the ground truth. This detection is superior to

* Corresponding author at: Laboratory of Cognitive Model and Algorithm, School of Computer Science, Fudan University, No. 825 Zhangheng Road, Shanghai 201203, PR China.
E-mail addresses: yuqian12@fudan.edu.cn (Q. Yu), weihui@fudan.edu.cn (H. Wei), chengzhuanyang13@fudan.edu.cn (C. Yang).

http://dx.doi.org/10.1016/j.patcog.2016.11.020
Received 7 May 2016; Received in revised form 9 October 2016; Accepted 23 November 2016
Available online 18 December 2016
0031-3203/



(a) Input image (b) Template (c) Distance Tramsformation Map



(d) DCM Detection(High DCM Distance) (e) Best DCM Detection(Bad Template Localization) (f) Our Detection

Fig. 1. Illustration of the drawbacks of chamfer matching methods. Fig. 1(c) shows a distance transformation map, where the lighter the pixel, the longer the distance from the pixel to its nearest edge. Fig. 1(d) shows a DCM detection with the high DCM distance especially in the head part, despite decent template localization. Fig. 1(e) shows the best DCM detection with the lowest DCM distance but bad template localization. Both DCM detections in Fig. 1(d–e) show unsatisfactory IoUs. Fig. 1(f) shows our detection with four contour fragments correctly located and the detection rectangle close to the ground truth.

DCM detections for two reasons. First, object representation using separate fragments is much more flexible than using single rigid template. The geometric constraint between any pair of the separate fragments is elastic, making dealing with deformation between the template and the object shape in an input image easier. Second, separate fragments can be better located than a single global and rigid template. Note that, although the chamfer matching methods can provide a relatively rough pixel-to-pixel correspondence, their main goal is to detect objects with bounding-boxes in cluttered images. Therefore, final detections are not seriously affected by omitting some template contours.

Fig. 2 illustrates general processing employed by LPOCM and LPDCM. The giraffe template was divided into four contour fragments to form a graph where each node represents a fragment and each edge represents the geometric constraint between the corresponding two fragments. The contour fragments are respectively located in the input image given in Fig. 1(a) using OCM or DCM, and their candidates are shown in Fig. 2(c). Only candidates that collectively lie within the correct geometric constraints associated with the overall template and have low chamfer distances can be assembled into the final giraffe shape. Hence, the process of the proposed chamfer matching becomes a combinatorial problem in the graph representation. If the graph in Fig. 2(b) is considered as a probabilistic graphical model, such as Markov random field (MRF), the combinatorial problem encountered in our work is easily solved by the inference in the MRF. Let the contour fragments correspond to MRF nodes, where detected candidates are states of the MRF nodes. Then, the solution to the combinatorial problem can be got by inference in the MRF. As shown in Fig. 2(d), the detection result is the best combination of candidates having good geometric constraint and fragment localization.

The remainder of this paper is organized as follows. Related works are reviewed in Section 2. Basic chamfer matching definitions are then reviewed in Section 3. In Section 4, local part oriented chamfer matching (LPOCM) and local part directional chamfer matching (LPDCM) are proposed, and then shape templates and discriminative contour fragments are learned. In Section 5, the proposed methods are tested, and their performance compared with those of OCM and DCM on the popular benchmarks of ETHZ shape classes, INRIA horses and Weizmann horses. Finally, this paper is concluded in Section 6.

2. Related work

Chamfer matching was originally proposed by Barrow et al. [2]. To reduce the computational load, Hierarchical Chamfer Matching was proposed [4]. Opelt et al. [20] proposed a chamfer matching method that first learned codebooks of contour fragments, and then used chamfer distance to match the learnt fragments to edge images. However, all of these classic methods ignore the orientation of the