Pattern Recognition 55 (2016) 183–197

Contents lists available at ScienceDirect

Pattern Recognition

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

Object matching with hierarchical skeletons

Cong Yang *, Oliver Tiebe, Kimiaki Shirahama, Marcin Grzegorzek *

Pattern Recognition Group, Institute for Vision and Graphics, University of Siegen, Siegen D-57076, Germany

ARTICLE INFO

Article history:
Received 19 August 2015
Received in revised form
20 January 2016
Accepted 23 January 2016
Available online 5 February 2016

Keywords:
Skeletonisation
Hierarchical skeleton
Skeleton evolution
Object representation
Object matching

ABSTRACT

The skeleton of an object provides an intuitive and effective abstraction which facilitates object matching and recognition. However, without any human interaction, traditional skeleton-based descriptors and matching algorithms are not stable for deformable objects. Specifically, some fine-grained topological and geometrical features would be discarded if the skeleton was incomplete or only represented significant visual parts of an object. Moreover, the performance of skeleton-based matching highly depends on the quality and completeness of skeletons. In this paper, we propose a novel object representation and matching algorithm based on hierarchical skeletons which capture the shape topology and geometry through multiple levels of skeletons. For object representation, we reuse the pruned skeleton branches to represent the coarse- and fine-grained shape topological and geometrical features. Moreover, this can improve the stability of skeleton pruning without human interaction. We also propose an object matching method which considers both global shape properties and fine-grained deformations by defining singleton and pairwise potentials for similarity computation between hierarchical skeletons. Our experiments attest our hierarchical skeleton-based method a significantly better performance than most existing shape-based object matching methods on six datasets, achieving a 99.21% bulls-eye score on the MPEG7 shape dataset.

1. Introduction

Shape is an expressive abstraction of the visual pattern of an object. While there are many different approaches [1] using shape for object matching, nearly all of them face the same challenge: object deformation. As shown in Fig. 1, the shapes of the same object are visually different depending on its deformations. To overcome this, on the one hand, various robust shape descriptors [2–11] are designed to capture both local and global geometric properties. On the other hand, some holistic [12,13] and elastic [8,14,2] matching algorithms are proposed to handle the ambiguous correspondences. Among the above-mentioned research efforts, skeleton is an important shape descriptor for deformable object matching since it integrates both geometrical and topological features of an object.

Fig. 2 shows an overview of the skeletonisation process to convert a given shape (a) into a skeleton (d). Specifically, a skeleton is defined as a connected set of medial lines along the limbs of its shape [15]. From a technical point of view, such a skeleton is extracted by continuously collecting centre points of maximal tangent disks touching the object boundary on two or more locations, as shown in Fig. 2 (b) and (c). The centre point of a maximal tangent disk is referred to as a skeleton point. The sequence of connected skeleton points is called a skeleton branch. A skeleton point having only one adjacent point is an endpoint (the skeleton endpoint). A skeleton point having three or more adjacent points is a junction point. The skeletons described above usually lead to a better performance than contour or other shape descriptors in the presence of partial occlusion and articulation of parts [16]. This is because skeletons have a notion of both the interior and exterior of the shape [16], and are useful for finding the intuitive correspondence of deformable shapes.

However, a skeleton is sensitive to the deformation of an object's boundary because little variation or noise of the boundary often generates redundant skeleton branches that may seriously disturb the topology of the skeleton [17–19]. Furthermore, a large number of skeleton branches may cause the overfitting problem and high computation complexity. Though skeleton pruning [12,17] approaches can remove the inaccurate or redundant branches while preserving the essential topology, they normally require manual intervention to produce visually pleasing skeletons. Moreover, the performance of skeleton-based matching highly depends on the quality and completeness of skeletons.

To overcome these problems, we propose a hierarchical skeleton-based object matching method. A hierarchical skeleton is a set of skeletons that represent an object at different levels. More specifically, during the skeleton pruning process, we store all the pruned branches until the skeleton is pruned to the simplest form. These

* Corresponding authors. Tel.: +49 271 740 3972; fax: +49 271 740 1 3972.
E-mail address: cong.yang@uni-siegen.de (C. Yang).

http://dx.doi.org/10.1016/j.patcog.2016.01.022
0031-3203/

Fig. 1. An illustration of shapes which significantly vary depending on deformations.

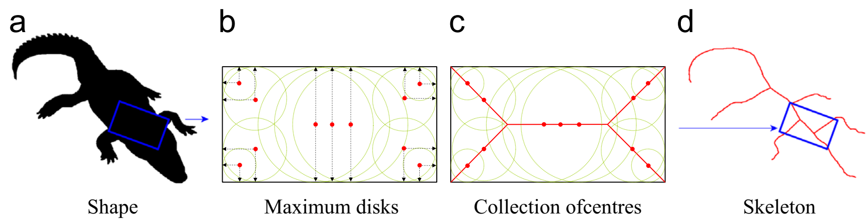

Fig. 2. An overview of the skeletonisation process to convert a given shape (a) into a skeleton (d). (b) and (c) visually illustrate the skeleton extraction process, where the skeleton (red line) of a shape (rectangle) is generated by collecting the centres (red dots) of all discs (green dotted circles) that touch the boundary of the shape on two or more different locations (dotted arrows). (a) Shape. (b) Maximum disks. (c) Collection of centres. (d) Skeleton. (For interpretation of the references to colour in this figure caption, the reader is referred to the web version of this paper.)

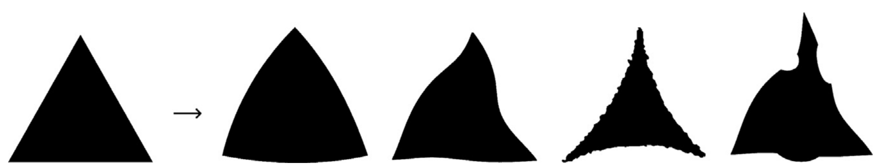

Fig. 3. Examples of shapes that are perceptually similar to the original one, irrespective of fine-grained noises and deformations.

branches are reused to construct the hierarchical skeleton which is favourable for the following reasons: First, it does not need any manual intervention since we consider a set of skeletons rather than a single one. Second, a hierarchical skeleton captures geometric and topological features at different levels along with skeleton pruning. Fine levels feature the small object deformation while skeletons at coarse levels capture global shape deformations. This enables us to develop an object matching algorithm that allows more deformations on finer levels while preserving important global geometrical and topological properties. This design is based on the fact that objects (e.g. the four objects on the right side of the arrow in Fig. 3) reconstructed with the same skeleton topology are still perceptually similar to the original (the triangle on the left side of the arrow in Fig. 3) even though there are some fine-grained noises and deformations.

The third advantage of a hierarchical skeleton is that it can also provide additional information for improving the object matching accuracy. In particular, by looking into the skeleton pruning process, transitions of pruned skeletons from the same category are more similar than those from different ones. This is because skeletons from the same category have more similar branches and these branches on each level have similar effects on the possible skeleton reconstruction. We call this phenomenon skeleton evolution. In Section 5.3, we show that adopting skeleton evolutions improves the performance of object matching.

As the fourth advantage, the hierarchical skeleton is obtained along with the skeleton pruning process, requiring no extra computational cost. Lastly, by limiting levels of hierachical skeletons, we can filter out skeleton branches which represent shape properties irrelevant to matching. This alleviates the overfitting problem.

2. Related work

Several skeletonisation methods have been developed to generate proper skeletons [20–23]. One typical approach is to continuously collect the centre points of maximal tangent disks that touch the object boundary in two or more locations. However, all of the obtained skeletons are sensitive to small changes and noises in the object boundary [24,25]. The intrinsic reason is that a small protrusion on the boundary may result in a large skeleton branch. To solve this problem, Choi et al. and Telea et al. [26–29] proposed algorithms to detect the skeleton in a distance map of the boundary points. Fig. 4(i) shows a skeleton obtained by the method in [28]. Although these methods can preserve some visual parts of a shape, some significant parts are missing. Therefore, they cannot guarantee the completeness of a skeleton. To overcome this, Bai and Latecki present significance measures for skeleton pruning associated with Discrete Curve Evolution (DCE) [12] or Bending Potential Ratio (BPR) [17]. Both methods decide whether or not a skeletal branch should be pruned by evaluating the contribution of its corresponding boundary segment to the overall shape. However, these methods require manual intervention to stop the evaluation and produce visually pleasing skeletons. For example, in Fig. 4, DCE [12] requires a proper stop parameter k to calibrate the pruning power. However, different stop parameters for the same object (the first row in Fig. 4) or the same parameter for different objects (the second row in Fig. 4) lead to visually different skeletons in which some important parts are missing (legs in Fig. 4(a), (b), (e), (f)). Furthermore, even if we find the best stop parameter, skeletons of the same object sometimes differ if the scale is changed (Fig. 4(g) and (h)). This is because the vanishing of shape parts is unavoidable when the resolution decreases [30]. Therefore, fixing k for skeleton pruning is not a proper solution for all objects. In contrast, our hierarchical skeleton is a collection of skeletons obtained by all the stop parameters. This not only eliminates the necessity of manually tuning a stop parameter, but also preserves both the coarse-grained global and fine-grained local properties of a shape.

For skeleton matching, most methods [31–38] only consider one skeleton for a shape. However, the matching performance relies on the quality of a skeleton since it is essential to find the correct corresponding elements. Some methods [18,25,39] promote the matching performance by fusing additional shape descriptors. Though the global matching accuracy could be improved, it requires