Pattern Recognition 57 (2016) 115–135

Contents lists available at ScienceDirect

Pattern Recognition

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

Shape matching by part alignment using extended chordal axis transform

Z. Yasseen [a,*], A. Verroust-Blondet [a], A. Nasri [b]

[a] Inria, Paris 2 rue Simone Iff CS, 42112 75589 Paris Cedex 12, France
[b] College of Computing, Fahad Bin Sultan University, KSA

ARTICLE INFO

Article history:
Received 16 February 2015
Received in revised form
18 February 2016
Accepted 15 March 2016
Available online 2 April 2016

Keywords:
Shape descriptors
Shape matching
Salient features
Ordered part correspondence
Chordal axis transform
Dynamic 2D shape segmentation
Dynamic programming

ABSTRACT

One of the main challenges in shape matching is overcoming intra-class variation where objects that are conceptually similar have significant geometric dissimilarity. The key to a solution around this problem is incorporating the structure of the object in the shape descriptor which can be described by a connectivity graph customarily extracted from its skeleton. In a slightly different perspective, the structure may also be viewed as the arrangement of protruding parts along its boundary. This arrangement does not only convey the protruding part's ordering along the anti clockwise direction, but also these parts on different levels of detail. In this paper, we propose a shape matching method that estimates the distance between two objects by conducting a part-to-part matching analysis between their visual protruding parts. We start by a skeleton-based segmentation of the shape inspired by the Chordal Axis Transform. Then, we extract the segments that represent the protruding parts in its silhouette on varied levels of detail. Each one of these parts is described by a feature vector. A shape is thus described by the feature vectors of its parts in addition to their angular and linear proximities to each other. Using dynamic programming, our algorithm finds a minimal cost correspondence between parts. Our experimental evaluations validate the proposition that part correspondence allows conceptual matching of precisely dissimilar shapes.
© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Shape descriptors and similarity measures have been a central research issue for decades. One of the main motivations is to obtain a retrieval mechanism over a database of images. Proposed approaches fall into numerous categories [46] such as contour-based, region-based, transforms, moments, scale space, or spatial interrelation. Many challenges and requirements withstand shape descriptors. One with vital importance and considerable difficulty is tolerance for intra-class variety. Objects may be labeled the same while having disparate geometries (see Fig. 1). While the tails of the two airplanes in Fig. 1 (or the backs of the two chairs) have dissimilar shapes, they match within their contexts. Their geometric properties such as size and angle of protrusion become relative to the other protruding parts of the shape. This part alignment process allows matching conceptually similar but precisely dissimilar shapes. Although the part correspondence shown in Fig. 1 is trivial for a human, it is a composite process of dynamic segmentation, numeric representation, and search for an optimal solution when it comes to machines. The part that is concerned with dynamic segmentation also belongs to human intuition. In fact, many studies in cognitive science support the theory that humans visually segment objects into simpler parts [10,27]. Propositions concerning the visual segmentation process draw relations with geometric primitives [8], negative minima of curvature along the boundary [7,44], or skeleton-based segmentation [35] of the shape.[1] Moreover, these studies grant theories concerning numeric representations that contribute in the segmentation process. Cohen and Singh [10] conducted a series of

* Corresponding author.
E-mail address: zyasseen@gmail.com (Z. Yasseen).

[1] Hoffman and Richards [20] introduced the 'minima rule' which states that human vision defines part boundaries at negative minima of curvature on silhouettes. Biederman [8] proposed a recognition-by-components theory that assumes the existence of a finite set (a total of 36) of simple geometric components that make up an object-part vocabulary. Cave and Kosslyn [9] questioned whether the segmentation is performed prior to recognition and aggravate the role of spatial relations among components. Neri [27] presented experimental evidence to support the recognition-by-part hypothesis as opposed to recognition-as-a-whole. Delorme et al. [13] studied the influence of the presence of diagnostic animal parts and their relative spatial configuration on visual processing. Higgins and Stringer [19] proposed a role for independent motion of body parts in its segmentation.

http://dx.doi.org/10.1016/j.patcog.2016.03.022
0031-3203/© 2016 Elsevier Ltd. All rights reserved.

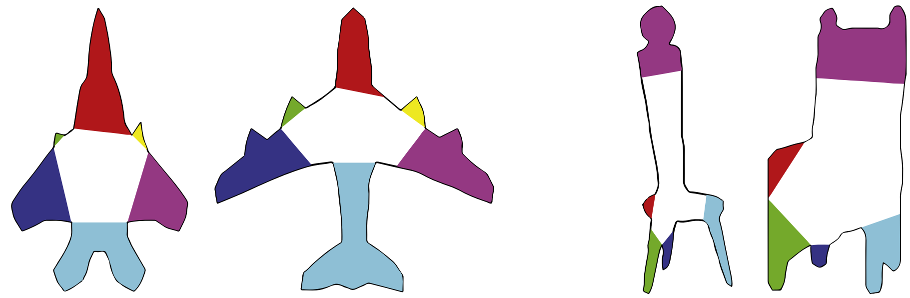

Fig. 1. Matching shapes by aligning their visual parts allows corresponding parts that are conceptually similar but precisely dissimilar.

psychophysical experiments to quantify the influences of segment length, turning angles at part boundaries, and part protrusion on segment identification. De Winter and Wagemans [44] analyzed the role of curvature singularities, symmetry, elongation, proximity, good continuation, and cognitive influences on rules of part formation. Bertamini and Wagemans [7] presented a review on evidence from behavioral, neurophysiological, imaging, and developmental studies to conclude that "convexity and concavity information along a contour is something that the visual system computes and represents as part of how shape is analyzed".

Human perception is known to easily overcome imprecision and occlusion [30]. Hence, we presume a role for these theories while attempting to match the silhouettes of objects taken from different perspectives, with part occlusion, or intra-class diversity. This role helps analyze shapes from the same perspective in which they are perceived by a human inspector: part-based representation independent of level of detail. In this paper, we propose a shape matching scheme inspired by the aforementioned theories and validate its correctness on benchmark datasets. We compute both a skeletal description and a dynamic 2D shape decomposition.

The input data is a binary image representing the silhouette of a single object. We extract the contour, locate corner points, and sample the in-between contour fragments uniformly. The advantage of locating corner points is the inclusion of the sharp features in the sample set. The region is then triangulated using constrained Delaunay triangulation (CDT). The rectified chordal axis transform [29] (CAT) which is a descendant of the medial axis transform, and a set of pruning and merging operations provide a skeleton with an association between skeletal segments and subregions. Using this skeleton, we segment the shape and embed the segmentation in a hierarchy where the leaves are the protruding parts corresponding to the terminal segments defined by the CAT.

The subtrees in this hierarchy group terminal segments with their adjacent internal ones. After pruning away *larger* subtrees using some heuristics, the remaining ones represent the features of the shape on different levels of detail. We refer to all such subtrees and leaves as the *visual parts* that may contribute to the matching process.

Visual parts are always arranged in their anti-clockwise order of appearance along the boundary of the object. Their spatial and angular distances comprise an inter-distance matrix relating every pair of them. Each *visual part* is described by a set of geometric attributes including relative size, eccentricity, circularity, and convexity in addition to a shape signature. The geometric attributes are multiplied by weights and assembled into a feature vector. The distance between two visual parts is then defined by the Euclidian distance between feature vectors plus the squared distance between signatures. The matching process is performed by topologically rotating the position of the 'first part' in the shape. Consequently, more acceptable segmentations of an object will result from matching it to those in the same class. For every configuration, we employ dynamic time warping (DTW) to find a minimal cost match.

Basically, the components of our work are a skeleton-based segmentation, geometric salience measures to describe visual parts and their spatial relations, and a matching scheme built on a part-to-part correspondence between shapes. The shape descriptor is invariant under affine transformations such as scaling, translation, and rotation. It has also demonstrated tolerance for occlusion, articulation, and minor deformations. The contributions of our work are as follows:

- We bring the CAT into shape description and strengthen some of its weaknesses that are revealed by dissimilarities between skeletons of objects of the same class. It is worth mentioning that Prasad [29] does not provide a matching technique adapted to the CAT representation method.
- We alter the customary practice followed when using DTW in closed shapes by rotating visual parts instead of sample points to find the best match.

The matching process based on the protruding parts is simple and visually intuitive compared with the ones based on skeleton matching. Moreover, this matching process can compare two shapes at different levels of detail or with different topologies in a unified way.

The rest of the paper proceeds as follows. We discuss related work that describe shapes by parts or use their skeletons in the process in Section 2. We present our contribution to the CAT segmentation in Section 3, whereas the topological structure that embeds different levels of details is discussed in Section 4. The geometric properties, saliency, and distance measure of parts are detailed in Section 5. In Section 6, we describe the DTW method, and present the retrieval results obtained on Kimia-99 and Kimia-216 datasets [33]. Section 7 presents a brief conclusion of the work presented in this paper.

2. Related work

Hoffman and Richards [20] introduced the 'minima rule' which states that human vision defines part boundaries at negative minima of curvature on silhouettes. Following the 'minima rule' [20], most of the fragment-based shape descriptors [12,23,25,36,43] extract

Download English Version:

https://daneshyari.com/en/article/530039

Download Persian Version:

https://daneshyari.com/article/530039

Daneshyari.com