Using Geodesic Information in Shape Matching for Object Recognition

Vittal Premachandran* and Ramakrishna Kakarala
School of Computer Engineering, Nanyang Technological University, Singapore
*corresponding author, E-mail: vittalp@pmail.ntu.edu.sg

Abstract

Object recognition in Computer Vision is done by the extraction of features from the object being recognized, and matching them to similar features, which are extracted from known objects. The features could be the object's colour, texture, or shape, or a combination of one or more of them. Humans can usually recognize the object's class from just the shape information. The figure below shows multiple instances of a coffee mug, each having different colour and texture, but all having a characteristic shape. In this paper, we look at how we can exploit this characteristic property by extracting meaningful information from object's contour, which can aid shape matching, and hence help object recognition.

We first extract the object's shape information i.e., its contour, using an edge detector and sample it into a set of uniformly spaced points. We then go on to improve the foundation method used in the recent paper [1], by making changes that capture more of the shape information. We propose the use of geodesic information for defining a relation between two sample points, as this takes into consideration the object's shape characteristics much better than the Euclidean distance. The geodesic distance between any two points is the shortest distance between the same, along the contour of the object. Apart from the distance metric, we also compute, and store, the angular information between two sample points, which is the angle between the tangent at that point and the displacement vector between the two points.

Armed with the distance and the angular information, we now create a shape context at each sampled point, similar to the one used in the famous paper by Belongie et al. in 2002 [2]. A description of the objects in the form of shape context is invariant to similarity transformations. In order to match two shapes, we compare the shape contexts of their respective sample points, followed by dynamic programming to find the best alignment of the points. Such a comparison helps in obtaining a "cost" metric, which is a direct indication of the cost of matching the two shapes. Smaller the cost, better the match, larger the cost, poorer the match, between the two shapes.

We test our shape descriptor on the well-known MPEG-7 dataset, which contains 1400 images, belonging to 20 different classes. The MPEG-7 dataset is a challenging dataset as the objects in each class not only differ by translation, scale, and rotation, but are also deformed and occluded in many cases. From the experiments performed on this database, we show that our method performs better than some of the commonly cited algorithms, which are believed to have state-of-the-art results.

Keywords: object recognition; shape matching; shape descriptor; computer vision

Fig: Examples of coffee mugs each having different colours & texture, but all having a characteristic shape.

References

[1] R.Gopalan, P.Turaga, R.Chellappa, Articulation-invariant representation of non-planar shapes, *ECCV*, 2010

[2] S.Belongie, J.Malik, J.Puzicha. Shape Matching and Object Recognition Using Shape Context, *PAMI*, 2002.