Signal Processing: *Image Communication* 24 (2009) 572–586

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Farthest point distance: A new shape signature for Fourier descriptors

Akrem El-ghazal [a,*], Otman Basir [a], Saeid Belkasim [b]

[a] *Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1*
[b] *Department of Computer Science, Georgia State University, Atlanta, GA 30303, USA*

ARTICLE INFO

Article history:
Received 7 August 2008
Received in revised form
15 February 2009
Accepted 7 April 2009

Keywords:
Fourier descriptors
Image retrieval
Shape signatures

ABSTRACT

Shape description is an important task in content-based image retrieval (CBIR). A variety of techniques have been reported in the literature that aims to represent objects based on their shapes. Each of these techniques has its pros and cons. Fourier descriptor (FD) is one of these techniques a simple, yet powerful technique that offers attractive properties such as rotational, scale, and translational invariance. Shape signatures, which constitute an essential component of Fourier descriptors, reduce 2-D shapes to 1-D functions and hence facilitate the process of deriving invariant shape features using the Fourier transform. A good number of shape signatures have been reported in the literature. These shape signatures lack important shape information, such as corners, in their representations. This information plays a major role in distinguishing between different shapes.

In this paper, we present the farthest point distance (FPD), a novel shape signature that includes corner information to enhance the performance of shape retrieval using Fourier descriptors. The signature is calculated at each point on a shape contour. This signature yields distances calculated between the different shape corners, and captures points within the shape at which the human focuses visual attention in order to classify shapes.

To reach a comprehensive conclusion about the merit of the proposed signature, the signature is compared against eight popular signatures using the well-known MPEG-7 database. Furthermore, the proposed signature is evaluated against standard boundary- and region-based techniques: the curvature scale space (CSS) and the Zernike moments (ZM). The FPD signature has demonstrated superior overall performance compared with the other eight signatures and the two standard techniques.

1. Introduction

The ease and convenience of capturing and transmitting digital images between digital cameras and image databases is a contributing factor in the immense growth of image databases. These databases cover a wide range of applications including military, environmental, astronomy, transportation, aviation, medical, and multimedia. The storage format of such image data is relatively standardized; however, the effective retrieval of images from such databases remains a significant challenge.

Typically, images in a database are retrieved based on either textual information or content information. Early retrieval techniques were based on textual annotation of images. Images were first annotated with text, then searched based on their textual tags. However, text-based techniques have many limitations due to their reliance on

* Corresponding author.
E-mail addresses: aelghaza@engmail.uwaterloo.ca (A. El-ghazal), obasir@uwaterloo.ca (O. Basir), sbelkasim@cs.gsu.edu (S. Belkasim).

0923-5965/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.image.2009.04.001

manual annotation—a process that is tedious and error-prone, especially for large data sets. Furthermore, the rich content typically found in images and the subjectivity of the human perception make the task of describing images using words a difficult, if not impossible, task.

To overcome these difficulties, content-based image retrieval (CBIR) was proposed [1]. This approach to image retrieval relies on the visual cues of images, rather than textual annotations, to search for images, and therefore has the potential to respond to more specific user queries. CBIR techniques use such visual contents as color, texture, and shape to represent and index images. The increasing interest in using shape features of objects for CBIR is not surprising, since shape is a more intrinsic property of objects than color and texture, and given the considerable evidence that natural objects are recognized primarily based on their shape [2,3]. A survey of users on the cognition aspects of image retrieval indicates that users are more interested in retrieval based on shape than on color and texture [4]. However, retrieval based on shape content remains a more difficult task than that based on other visual features [2].

During the last decade, significant progress has been made in both the theoretical and the practical research aspects of shape-based image retrieval [5,6]. There are mainly two approaches to shape representation, namely, the region-based approach and the boundary-based approach (also known as contour-based approach). Region-based techniques often use moment descriptors to describe shapes. These descriptors include geometrical moments [7,8], Zernike moments (ZM) [9,10], pseudo-Zernike moments [11], Legendre moments [9], and Tchebichef moments [12]. Other notable region-based techniques include generic Fourier descriptor (FD) [13], compound image descriptor [14], shape matrix [15], and the grid technique [16]. Although region-based approaches are global in nature and can be applied to generic shapes, they often involve intensive computation and fail to distinguish between objects that are similar [17].

In many applications the internal content of the shape is not as important as its boundary. Boundary-based techniques tend to be more efficient for handling shapes that are describable by their object contours [17]. Many boundary-based techniques have been proposed in the literature, including Fourier descriptors [18–20], curvature scale space (CSS) [21–23], wavelet descriptors [24,25], contour displacement [26], chain codes [27], autoregressive models [28], Delaunay triangulation technique [29], and multi-resolution polygonal shape descriptors [30]. Recently, dynamic programming (DP) has been adopted in order to achieve a high accuracy rate using shape boundaries [31–35]. Even though the DP-based techniques generally offer better performance than other techniques that do not use DP, the DP-based techniques suffer from being computationally expensive, making them impractical for large databases.

Fourier descriptors have proven to be better than other boundary-based techniques in many applications [18–20,36,37]. The traditional FDs are based on applying the Fourier transform to a shape signature. Many shape signatures have been used in Fourier descriptor techniques. However, the complex coordinate is the most frequently used signatures in the literature. Recent work shows that in shape-based image retrieval, the radial distance (RD) signature outperforms the complex coordinates (CC) and other signatures [38]. In order to increase the capability of the Fourier-based technique to capture local features, Eichmann et al. [39] have used the short Fourier transform (SFT). The SFT is not suitable for image retrieval because the matching process using SFT is computationally more expensive than the traditional FDs. Invariance to affine transforms allows considerable robustness in the case of rotating shapes in all three dimensions. Arbter et al. [40,41] have used a complex mathematical analysis and proposed a set of normalized descriptors that are invariant under any affine transformation. Also, Oirrak et al. [42] have used one-dimensional Fourier series coefficients to derive affine invariant descriptors. Zhang and Lu [38] have shown that even though the affine Fourier descriptors [40] was proposed to target affined shape distortion, it does not perform well on the standard affine invariance retrieval set of the MPEG-7 database. This is because the affine Fourier descriptors are designed to work on a polygonal shape under affine transformation and are not designed for a non-rigid shape [38]. Most of the Fourier-based techniques utilize the magnitude of the Fourier transform and ignore the phase information in order to achieve rotation invariance as well as make the descriptors independent from the starting point. However, Bartolini et al. [33] have described a technique in which the phase information is exploited. Kunttu et al. [20] have introduced a multiscale Fourier descriptor for shape-based image retrieval. These descriptors are presented in multiple scales by adopting the wavelet and Fourier transform, which improves the shape retrieval accuracy of the traditional Fourier descriptors. Recently, El-ghazal et al. have described curvature-based Fourier descriptors (CBFD) for shape retrieval. The invariant descriptors of the CBFD technique are derived from the 2-D Fourier transform of the curvature-scale image obtained from the image contour [43].

In general, Fourier descriptors is a promising boundary-based approach for shape-based image retrieval, as the FDs are based on the well-known Fourier theory, making them easy to compute and simple to normalize and interpret. In addition, the computational efficiency and compactness of FDs allow them to be well suited for online image retrieval.

To derive the FDs of an image, the 2-D image is converted to 1-D signature. Many signatures have been proposed in the literature [18,38,44]. The complex coordinates, the radial distance, and the triangular centroid area (TCA) are some notable signatures available to derive FDs. Fourier descriptors derived from different signatures can have significantly different effects on the result of retrieval [38]. In this paper, we propose a novel signature, namely farthest point distance (FPD), and compare it with other frequently used shape signatures.

The paper is organized as follows: Section 2 gives a brief description of commonly used shape signatures.