Available online at www.sciencedirect.com
SCIENCE DIRECT®

Pattern Recognition Letters 27 (2006) 123–132

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Multiscale Fourier descriptors for defect image retrieval

Iivari Kunttu [a,*], Leena Lepistö [a], Juhani Rauhamaa [b], Ari Visa [a]

[a] *Tampere University of Technology, Institute of Signal Processing, P.O. Box 553, FIN-33101 Tampere, Finland*
[b] *ABB Oy, Process industry, P.O. Box 94, FIN-00381 Helsinki, Finland*

Received 12 November 2004; received in revised form 4 August 2005
Available online 10 October 2005

Communicated by R. Manmatha

Abstract

Shape is an essential visual feature of an image and it is widely used to describe image content in image classification and retrieval. In this paper, two new Fourier-based approaches for contour-based shape description are presented. These approaches present Fourier descriptors in multiple scales, which improves the shape classification and retrieval accuracy. The proposed methods outperform ordinary Fourier descriptors in the retrieval of complicated shapes without increasing computational cost.

Keywords: Shape retrieval; Shape classification; Fourier descriptor; Multiresolution shape description

1. Introduction

The description of the shapes in image is an essential task in the field of pattern recognition. During recent decades, a number of approaches and solutions have been presented to characterize different shapes. The primary purpose of the early shape descriptors was shape classification, whereas during recent years the use of shape description in image retrieval has received increasing interest (Del Bimbo, 2001). For example, in multimedia applications, content-based image retrieval (CBIR) plays a significant role.

In CBIR systems, one of the problems is to answer the question: "Which of the database images contain the most similar shapes to the query image"? This type of image retrieval is called shape similarity-based retrieval (Mehtre et al., 1997). In this kind of retrieval, the aim is to find similar shapes from the database as accurately as possible. On the other hand, the classification accuracy (effectiveness) of a certain descriptor is not an adequate measure for its usefulness in the retrieval. Due to the increasing number of online retrieval solutions, computational lightness (efficiency) is nowadays considered equally important as effectiveness (Zhang and Lu, 2004). A recently introduced multimedia standard, MPEG-7 (Manjunath et al., 2002), has set several principles for measuring a shape descriptor. The principles are good retrieval accuracy, compact features, general application, low computational complexity, robust retrieval performance, and hierarchical coarse to fine representation (Kim and Kim, 2000). These principles were used as criteria in the study of Zhang and Lu (2004), in which common shape description techniques were reviewed. Another review of the state of the art in shape description techniques is provided by Loncaric (1998).

Shape description techniques can be divided into two types: region- and boundary-based techniques (Costa and Cesar, 2001). Region-based methods consider the whole area of an object. Different moments (Hu, 1962), including for example Zernike moments (Teague, 1980) are popular descriptors. Boundary-based shape descriptors use only the object boundary in the description of the object shape. The most common boundary-based shape descriptors are chain codes (Freeman and Davis, 1977),

* Corresponding author. Tel.: +358 3 3115 4393; fax: +358 3 3115 4989.
E-mail address: Iivari.Kunttu@tut.fi (I. Kunttu).

0167-8655/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2005.08.022

Fourier descriptors (Persoon and Fu, 1977) and simple descriptors such as circularity (Costa and Cesar, 2001), eccentricity, convexity, principle axis ratio, circular variance and elliptic variance (Iivarinen and Visa, 1998). Recently, growing research interest has been focused on curvature scale space (CSS) shape representation (Mokhtarian and Mackworth, 1986) that has been selected to be used as the boundary-based shape descriptor of MPEG-7 standard (Bober, 2001). However, despite the fact that the Fourier descriptor method is over 30 years old (Granlund, 1972; Persoon and Fu, 1977), it is still found to be a valid shape description tool. In fact, Fourier descriptor has proved to outperform most other boundary-based methods in terms of retrieval accuracy and efficiency. This has been verified in several comparisons. Kauppinen et al. (1995) made a comparison between autoregressive models (Dubois and Glanz, 1986) and Fourier-based descriptors in shape classification. In most cases, Fourier descriptors proved to perform better than autoregressive models. In a comparison made by Mehtre et al. (1997), the retrieval ability of chain codes, Fourier descriptors, and different moments were compared in shape similarity-based retrieval. In this case, the best results were obtained by using moments and Fourier descriptors. In the recent studies of Zhang and Lu (2003a,b), Fourier descriptors and Zernike moments outperformed CSS-representation in terms of retrieval accuracy and efficiency. Similar results were also obtained by Kunttu et al. (2004).

In addition to good retrieval and classification accuracy, there are also other reasons which make Fourier descriptors probably the most popular of the boundary-based shape representations. The main advantages of the Fourier-based shape representations are that they are compact and computationally light. Furthermore, they are easy to normalize and their matching is a very simple process. Also their sensitivity to noise is low when only low frequency Fourier coefficients are used as descriptors.

It has been found that complicated shapes can be effectively characterized by using a description with multiple resolutions (Costa and Cesar, 2001; Mokhtarian, 1995). CSS-representation uses multiple resolutions that are achieved by smoothing the boundary. However, the drawbacks of CSS are relatively low shape classification accuracy and efficiency compared to Fourier descriptors (Zhang and Lu, 2003a,b). Also the matching procedure of the CSS-features is not as straightforward as that of Fourier descriptors. Wavelet transform (Chui, 1992) has been widely used in multiscale image analysis. However, it has only a few applications in the shape description (Yang et al., 1998; Tieng and Boles, 1997). The obtained descriptors are not rotation invariant. Furthermore, the matching scheme of these wavelet representations is more complicated and time consuming than that of Fourier descriptors. This reduces their usability in on-line retrieval solutions.

In this paper, two multiresolution approaches to shape description are presented. The first one, called here *Multiscale Fourier*, utilizes a combination of wavelet and Fourier transforms. *Multiscale Fourier* descriptor is obtained by applying the Fourier transform to the coefficients of the multiscale wavelet transform. Consequently, the Fourier descriptor is formed from multiresolution representation of the shape. The *Multiscale Fourier* approach has given promising results in the classification of general shapes as well as shapes of industrial defects in (Kunttu et al., 2003). In (Kunttu et al., 2004), this method also outperformed CSS-representation and *Contour Fourier* descriptor (Kauppinen et al., 1995) in the shape-based retrieval of different kinds of defect images. The second multiresolution approach presented in this paper is called *Boundary Scale Fourier* descriptor. This descriptor is obtained by using different scales of the boundary line. The scales are achieved by smoothing the boundary line iteratively and applying the Fourier transform to the boundary of different degrees of smoothness. In both of these descriptors, the matching is as simple as in the case of Fourier descriptors, which is a benefit in the on-line retrieval.

The outline of this paper is the following. Section 2 presents the methodology of shape description using Fourier-based methods. In addition to ordinary single-scale Fourier descriptors, both of the proposed multiresolution methods are presented in that section. Section 3 is the experimental part of this paper. In the experiments, the proposed methods are evaluated and compared to ordinary Fourier descriptors in retrieval. For retrieval experiments, a database of real industrial defect shapes is used. In Section 4, the results and performance of the methods are discussed. Section 5 concludes this study.

2. Shape description

In this paper, the shape description methods are based on the object boundary. In shape description, a boundary is usually presented using some shape signature i.e. a function derived from the boundary coordinates. Complex coordinate function (Kauppinen et al., 1995) is a simple and probably the best-known signature used in the Fourier-based shape description. Let (x_k, y_k), $k = 0, 1, 2, \ldots, N-1$ represent the object boundary coordinates, in which N is the length of the boundary. The complex coordinate function $z(k)$ expresses the boundary points in an object centered coordinate system in which (x_c, y_c) represents the centroid of the object:

$$z(k) = (x_k - x_c) + j(y_k - y_c) \tag{1}$$

Hence, using this function, the boundary is represented independent of the location of the object in the image. In this way the translation invariance can be achieved.

2.1. Fourier descriptors

Fourier descriptors characterize the object shape in a frequency domain. The descriptors can be formed for the complex-valued boundary function using the discrete Fou-