Available online at www.sciencedirect.com

SCIENCE DIRECT®

Image and Vision Computing 23 (2005) 33–49

image and vision COMPUTING

www.elsevier.com/locate/imavis

Study and evaluation of different Fourier methods for image retrieval

Dengsheng Zhang*, Guojun Lu

Gippsland School of Computing and Information Technology, Monash University, Churchill, Vic. 3842, Australia

Received 11 February 2004

Abstract

Shape is one of the primary low-level image features exploited in the newly emerged content-based image retrieval (CBIR). Many shape methods exist. Among these shape methods, Fourier descriptor (FD) is one of the most widely used shape descriptors due to its simple computation, clarity and coarse to fine description capability. FD has been applied to a variety of applications, including image retrieval application. Generally, FD can be acquired in a number of ways, however, FD acquired from different ways can have different retrieval performance. In this paper, we study shape retrieval using FD. Specifically, we study different ways of acquiring FD, the number of FD features needed for general shape description and the retrieval performance of different FD. A Java client–server retrieval framework has been developed to facilitate the study. The retrieval performance of the different FD is tested using a standard shape database and a commonly used performance measurement.
© 2004 Elsevier B.V. All rights reserved.

Keywords: CBIR; Shape; Fourier descriptor; Retrieval

1. Introduction

Owing to the rapid development of digital and information technologies, more and more images are generated and available in digital form. Before one can use any such information, however, it has to be located first. This requires image be effectively and efficiently described to facilitate automatic searching. A new multimedia application, called content-based image retrieval, or CBIR has emerged to address this urgent issue. In CBIR, images are described by a few low-level features such as color, texture, shape or the combination of them.

Shape is one of the several primary low-level image features. Human beings tend to perceive scenes as being composed of individual objects, which can be best identified by their shapes. Besides, as far as query is concerned, shape is simple for user to describe, either by giving example or by sketching. Once images or scenes are broken down into individual objects, they can be exploited to facilitate CBIR. Applications on shape retrieval can be found in many other areas, such as medicine, object recognition, law enforcement, military, etc.

Various shape techniques exist in the literature, these methods can be classified into two categories: region based and contour based. In region-based techniques, all the pixels within a shape are taken into account to obtain the shape representation. Common region-based methods include moment descriptors [15,18,25,34,35] and grid method [8,20,32,33]. Region-based methods can be applied to more general applications than contour-based methods. However, they usually involves more computation and storage. Compared with region-based shape representation, contour-based shape methods are more popular in the literature. The reasons are in three aspects. First, it is generally recognized in the literature that shape can be described solely by its boundary features and human beings are able to discriminate shapes by their contours or outlines. Second, most of real world objects have clear contours, which are readily available. In fact, contour-based shape methods can easily find applications and have produced satisfactory results in many situations. In this sense, applications of contour-based shape techniques are also quite general. Third, contour-based shape descriptors are usually more easily to derive. Contour-based methods

* Corresponding author. Tel.: +61 3 9902 6772; fax: +61 3 9902 6842.
E-mail addresses: dengsheng.zhang@infotech.monash.edu.au (D. Zhang), guojun.lu@infotech.monash.edu.au (G. Lu).

0262-8856/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2004.09.001

represent shape as a 1D signal which is easier to analyze than to analyze 2D signal. Contour-based shape methods include global shape descriptors [25], shape signatures [10,26], autoregressive model [16,17], structural methods [5,22], geometric invariants [19], spectral descriptors [14, 36,38,39] and curvature scale space (CSS) methods [1,2,24].

Among the many existing shape methods, FD has several desirable characteristics, which make it a popular descriptor. These characteristics are: low computation complexity, clarity and coarse to fine description.

Low computation complexity essentially implies stability. Stability means when a descriptor is used to describe different type of shapes or applied to different applications, its performance is least affected. This requires that the derivation of descriptor should depends on as few parameters as possible. This equivalently requires that the derivation process of the descriptor be not complex. Because the more complex the derivation process, the more parameters are involved in obtaining the descriptor; consequently, the descriptor depends on more uncertain or empirical factors. Consequently, when a descriptor derived under delicately controlled conditions is applied to another application or real world applications, its performance is drastically discounted. The many structural methods and geometric invariants are among the examples of these type of unstable descriptors.

Clarity means that a shape descriptor should have a clear perceptual or physical meaning and have a simple interpretation. This characteristic requires not only the modeling process or the feature extraction process be simple, but also requires the normalization and matching process be simple. The simple feature extraction can be spoiled by a complex normalization process or a complex matching process. For example, the extraction of geometric moments is very simple and the extracted features have clear physical meaning. However, the normalization process spoils the physical meaning of the acquired features. In fact, the physical meaning of the seven geometric moment invariants is not known [15].

Coarse to fine description capability means that a shape descriptor can describe shape with incremental accuracy. This characteristic is desirable for efficiency. It is useful in two aspects. On the one hand, the size of the descriptor can be adaptable to particular application depending on the level of accuracy the application requires. On the other hand, during the retrieval, coarse level description can be used to quickly eliminate those irrelevant shapes from the required list and finer level description can be added to refine the initial result.

Many FD methods have been reported in the literature, these include using FD for shape analysis [26,27,39], character recognition [13,29,30], object recognition [4,21,31], shape coding [9], shape classification [17] and shape retrieval [1,14,20,23,40,41]. However, no evaluation has ever been made on different FD methods and their performance. In this paper, we study different FD methods. Specifically, we study different ways of acquiring FD, what type of FD and what number of FD features are suitable for general image description and retrieval. We obtain desirable FD by retrieval performance study and obtain desirable number of FD features by compactness study. The rest of the paper is organized as follows. In Section 2, we describe three types of FD methods, that is conventional FD, affine FD and short-time FD. In Section 3, convergence speed of different FD and affine FD are studied. We conduct test of retrieval performance in Section 4. The retrieval performance of FD and short-time FD are compared in Section 5. In Section 6, we study compactness of FD. Section 7 concludes the paper and outlook future research.

2. FD methods

In this section, we study three types of FD methods, that is, *conventional FD* method, *affine FD* method and *short-time FD* method. Conventional FD method will be described in details in Section 2.1, affine FD will be described in Section 2.2, and short-time FD is introduced in Section 2.3.

2.1. Conventional FD

Conventional FD is derived by applying Fourier transform on a 1D shape signature. Several shape signatures have been used to derive FD, in this section, we study these signatures in Section 2.1.1. We also propose two new shape signatures, that is, *chordlength signature* and *area signature*, in attempt to improve existing signature methods. The derivation of conventional FD will be given in Section 2.1.2.

2.1.1. Shape signatures

In general, a *shape signature* $u(t)$ is any 1D function representing 2D areas or boundaries, it usually uniquely describes a shape. A shape signature usually captures the perceptual feature of the shape. In the following, we assume the shape boundary coordinates $(x(t), y(t))$, $t=0, 1,\ldots, N-1$, have been extracted in the preprocessing stage, t usually means arclength. In our implementation, the shape boundary points are extracted through a preprocessing module which consists of a denoising procedure and a 8-connectivity contour tracing procedure [28]. For notation convenience, different function names will be employed to denote different shape signatures.

2.1.1.1. Position function. *Position function*, or *complex coordinates*, is simply the complex number generated from the boundary coordinates:

$$z(t) = [x(t) - x_c] + i[y(t) - y_c] \tag{2.1}$$

Download English Version:

https://daneshyari.com/en/article/10360251

Download Persian Version:

https://daneshyari.com/article/10360251

Daneshyari.com