Available online at www.sciencedirect.com
ScienceDirect
J. Vis. Commun. Image R. 17 (2006) 970–985
JOURNAL OF VISUAL Communication & IMAGE Representation
www.elsevier.com/locate/jvci

Similarity retrieval of shape images based on database classification

Mao-Hsiung Hung, Chaur-Heh Hsieh *, Chung-Ming Kuo

Department of Information Engineering, I-Shou University, Dahsu, Kaohshiung 840, Taiwan

Received 5 January 2005; accepted 20 September 2005
Available online 15 November 2005

Abstract

Shape is a key visual feature used to describe image content. This paper develops a novel shape-based similarity retrieval system based on database classification which exploits the contour and interior region of a shape efficiently. In this system, the database of shape images is categorized automatically into 11 classes by a simple contour feature. In query, the contour feature of the input image is used to decide which class the query image belongs to. Then, the possible classes are selected dynamically from the database and to form candidate sets with different priority orders. Then, ART region feature is employed to compare the query with the candidate sets according to the priority order. Instead of using the original contour of a shape image directly, we employ a rough version of the original contour for the classification of shapes. The similarity test results indicate that the proposed method improves retrieval accuracy and speed significantly, as compared to ART.

Keywords: Shape similarity retrieval; Database classification; MPEG-7 ART descriptor

1. Introduction

The major issue in content-based image retrieval is to design a mechanism that can search perceptually similar image with a good accuracy and quick response time. For general color images, color, shape, and texture are major attributes to describe the image content. In some applications, such as trademark image retrieval and object recognition, color and texture information either do not exist in the images, or are removed because of their insignificance. This type of images is referred to as shape image, since it contains shape information only. Many shape descriptors have been developed in the literature, which are generally classified into two types: contour-based and region-based. The shape features of the contour-based type are extracted from the contour only of a shape image, whereas the features of the region-based type are obtained from the entire shape region. Recently, Zhang and Lu [1,2] presented an excellent review on the shape representation and description techniques. They compared various descriptors in terms of six principles including retrieval

* Corresponding author. Fax: +886 7 6578944.
E-mail address: hsieh@isu.edu.tw (C.-H. Hsieh).

1047-3203/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2005.09.002

accuracy, storage compactness, matching computation, invariance robustness, general application, and hierarchical representation. The conclusions they achieved are: (1) Zernike moment descriptor (ZMD) [3] is most suitable for region shape retrieval, (2) Fourier Descriptor (FD) is better than curvature scale space (CSS) descriptor for contour shape retrieval, and (3) ZMD outperforms CSS and FD. In [4], Kim et al. proposed angular radial transform (ART) descriptor, which is derived from a new basis function to preserve the orthogonality and rotation invariance like ZMD. In addition, ART descriptor has capacity to consider both complexities in radial and angular directions. The experimental result in Kim's work indicates that the retrieval performance of ART exceeds that of ZMD. Moreover, ART descriptor is simple and robust. Thus, their conclusion is that ZMD descriptor can be replaced by ART. Recently, MPEG-7 study group suggested two types of shape descriptors: CSS [5,6] and ART [7,8]. The former is contour-based and used to describe an object shape with a single contour; whereas the latter is region-based used for complex objects which consist of multiple contours such as trademark images.

In the past few years, many researchers have developed new descriptors for shape similarity retrieval aiming at improving retrieval accuracy and/or retrieval speed. However, most of them are effective only to retrieve the shapes that undergo geometrical deformations such as rotation, scaling, translation, and perspective transform, etc. In such case, the query image and retrieval images are highly pixel-correlated. Actually, many semantically similar shapes usually have less pixel correlation among them, thus resulting in large descriptive distances (matching distance). Accordingly, the retrieved results may not correctly meet human expectations. This can be illustrated by the example shown in Fig. 1. The two images in the figure are dissimilar from viewpoint of pixel correlation, but they are semantically the same because they both represent "bull head."

Eakins et al. [9] presented a trademark retrieval system called Artisan (automatic retrieval of trademark images by shape analysis) to improve retrieval accuracy. Artisan segments all trademark images into several key shape components, and then derives all features from these component boundaries. To mirror human perception, it groups such components into families and represents these families explicitly as image elements. Jain and Vailaya [10] proposed a two-stage trademark retrieval system to address both retrieval accuracy and computational efficiency. In the first stage, the edge directions and the invariant moments are used to prune the database to a reduced set of plausible candidates. In the second stage, a deformable template model eliminates false matches passing the first stage. Because simple shape features are used in the first stage, the system improves computational efficiency significantly. However, the pruning stage may miss the samples that are similar to the query [11].

Recently, the techniques of relevance feedback with multiple feature descriptors have been developed to address the problems of similarity retrieval [11–14]. The method first let a user label relevant/non-relevant images to the query on each retrieved result. The system then automatically adjusts the weight value of each descriptor according to the user's feedback. Because humans are involved in the adaptation loop, the system improves the retrieval results from the viewpoint of human perception. The major drawbacks of the relevance feedback are high system complexity and computational load.

Generally, two complex objects are considered similar if they look like both in the boundary (contour) and in the region details. Thus, the similarity measure for a complex object should consider contour of the object as well as the interior region of the object. To exploit the two attributes efficiently, a novel shape-based retrieval system based on database classification is developed. In the system, a simple contour feature is used

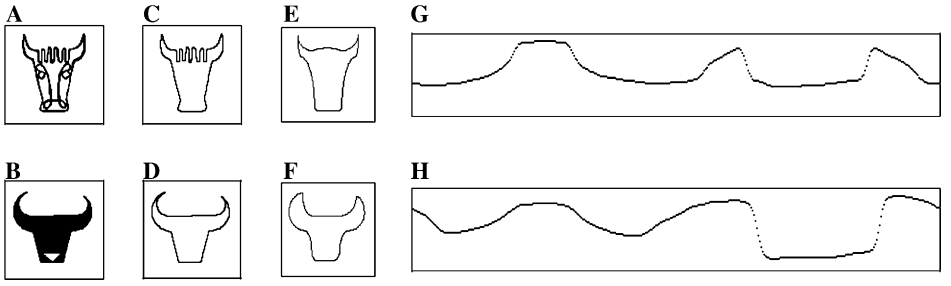

Fig. 1. Contour and signature of shape images, (A) and (B): shape images, (C) and (D): original contours of (A) and (B), (E) and (F): contour envelopes, (G) and (H): signatures of (E) and (F).