Content-Based Image Retrieval: Object Representation by the Density of Feature Points

De Cao Tran, Kinji Ono

Abstract:

A lot of research on multimedia database aims to develop techniques for searching and retrieving multimedia documents based on their contents. This is known as content-based searching and retrieval.

For image retrieval, many works proposed different techniques to extract the visual contents such as color, textual, shape etc. Among of these, shape is important because the semantic meaning of an image is strongly related to the objects in the image. Shape identification and representation offer a way to move from visual objects to real-world objects.

This paper introduces a novel method for shape representation. It is invariant to translation, scale and rotation. Object shape is represented by a density histogram of features points. Searching and retrieving images is carried out by measuring Euclid distance between the query image and the images in the database. Our experimentation on 100 images shows that this method is better than the Angloram representation that is based on the Delaunay triangulation of object feature points.

Key words: Shape representation, image content-based searching, image retrieval.

I. INTRODUCTION

Searching and retrieving image is not bit-by-bit searching. It is not a matching process on the raw data. The searching process is normally based on the "image content" that is the "semantic meaning" of the image in some level of abstraction. The simple approach dealing with this problem is text-based retrieval in which the images are annotated manually by *keywords*. Then, these annotations are organized in a database to support searching and retrieving images [1]. This approach faces two great difficulties. One is time-consume process due to the large amount of effort paid for annotating. The other, more essential, is that the text annotation depends on the subjective human perception. There are inconsistencies in the interpretation of the image content by different annotators. Sometimes, it is very difficult to express precisely the image content by text, for example, describing the shape of an object.

Content-based retrieval was proposed to overcome these difficulties. Instead of being manually annotated by text, the visual features of images are extracted automatically, objectively from raw data. There has been a lot of work [1,2,3,4,5] proposing various techniques for processing images and extracting visual features such as color histogram, textual, shape, etc. Among these visual features, shape is very attractive because the image meaning is strong related to the objects in the image. Shape identification and representation offer a way to move from visual objects to real-world objects. Shape may be sufficient to identify and classify an object completely and accurately [6].

There are two questions.

- How to extract object shape from an image?
- How to represent the object shape for measuring the similarity between them?

For the second issue, three are two important criteria that should be meet [7]:

- The representation of a shape should be invariant to scale, translation and rotation.
- The similarity measure between shapes should conform to human perception.

This paper introduces a novel method for shape representation. It is invariant to translation, scale and rotation. Shape of an object is represented by a density histogram of features points. For measuring the similarity of shape, the Euclid distance between two vectors of two shape representations is used. Our experimentation on 100 images shows that this method is better than the Angloram representation that is based on the Delaunay triangulation of object feature points.

The following section is about the related works. Section 3 devotes to represent our method. Section 4 presents the result of our experimentation. The conclusion and some discussion are in the last section.

II. RELATED WORK

Shape representations can be classified in two categories: boundary-based and region-based. The boundary-based representation uses the outer boundary of the shape while the region-based representation uses entire shape region. The most successful representatives for these two categories are Fourier Description and Moment Invariant [4].

The Fourier descriptor is based on the Fourier transform [ZAH72]. The boundary transformed using Fourier Transform Algorithm is used as image feature. The moments are based on probability density functions as expectations, variance, covariance and skewness [STR95]. The disadvantages of this approach are its high computation cost because the features are computed on entire region and low discrimination power [RUI98]. The Fourier Transform on boundary of region is quicker. It is also a good representation of shape but it is sensitive to the starting point of the shape boundary. Many

De Cao Tran, Kinji Ono, National Institute of Informatics, Japan
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430, Japan.
(email: tcde69@yahoo.ca)