Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 357–366
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Image retrieval based on indexing and relevance feedback

Sanjoy K. Saha [a,*], Amit K. Das [b], Bhabatosh Chanda [c]

[a] *CSE Department, Jadavpur University, Calcutta 700 032, India*
[b] *CST Department, B.E. College (DU), Howrah 711 103, India*
[c] *ECS Unit, Indian Statistical Institute, Calcutta 700 035, India*

Available online 22 May 2006

Abstract

In content based image retrieval (CBIR) system, search engine retrieves the images similar to the query image according to a similarity measure. It should be fast enough and must have a high precision of retrieval. Indexing scheme is used to achieve a fast response and relevance feedback helps in improving the retrieval precision. In this paper, a human perception based similarity measure is presented and based on it a simple yet novel indexing scheme with relevance feedback is discussed. The indexing scheme is designed based on the primary and secondary keys which are selected by analysing the entropy of features. A relevance feedback method is proposed based on Mann–Whitney test. The test is used to identify the discriminating features from the relevant and irrelevant images in a retrieved set. Then emphasis of the discriminating features are updated to improve the retrieval performance. The relevance feedback scheme is implemented for two different similarity measure (Euclidean distance based and human perception based). The experiment justifies the effectiveness of the proposed methodologies. Finally, the indexing scheme and relevance feedback mechanism are combined to build up the search engine.

Keywords: CBIR; Similarity measure; Human perception based similarity measure; Multidimensional indexing; Relevance feedback; Mann–Whitney test

1. Introduction

In a CBIR system, the feature extraction module computes various types of low-level features like shape, texture and color from an image. The search module retrieves the images similar to the query image from the database using a similarity measure based on the features. It should be fast and precision of retrieval should be high.

Apart from the feature set being used by the system, precision of retrieval depends on the similarity measure adopted by the search module. It is evident from the literature that various distance/similarity measures have been adopted by the CBIR systems. Mukherjee et al. (1999) have used template matching for shape based retrieval. A number of systems (Niblack et al., 1993; Srihari et al., 2000) have used Euclidean distance (weighted or unweighted) for matching. Other schemes include Minkowski metric (Fournier et al., 2001), self organising map (Laaksonen et al., 2000), proportional transportation distance (Vleugels and Veltkamp, 2002), etc. For matching multivalued features like color histogram or texture matrix, a variety of distance measures are deployed by different systems. It includes schemes like quadratic form distance (Niblack et al., 1993), Jaccard's co-efficient (Lai, 2000), histogram intersection (Gevers and Smeulders, 2000), etc. The details on combining the distance of various type of features is not available. But, it is clear that Euclidean distance is the most widely used similarity measure.

The simplest approach to search nearest neighbours is the linear search requiring O(n) distance (dissimilarity) computations where n is the number of images in the database. Obviously, it is prohibitive for large value of n. Thus, to obtain a fast response from the retrieval module, an indexing scheme is required. Lots of research work have been carried on in this direction which can be classified

* Corresponding author.
E-mail addresses: sks@becs.ac.in (S.K. Saha), amit@becs.ac.in (A.K. Das), chanda@isical.ac.in (B. Chanda).

0167-8655/$ - see front matter
doi:10.1016/j.patrec.2006.04.005

as: (a) spatial access methods and (b) distance based indexing methods. In spatial access methods, an image is represented by a finite set of features. Distance between two images is the Euclidean distance between their feature vectors. The Euclidean distance space is used to divide the database into spatial clusters. This is the basic principle in KD tree (Bentley and Friedman, 1979), quad tree (Samet, 1984) and R-tree family (Guttman, 1984; Leutenegger et al., 1997). While searching, some of the clusters are pruned to achieve faster response (Faloutsos et al., 1994). The spatial access methods suffer from curse of dimensionality. As the number of features (dimension) increases, computational cost also increases exponentially (Sebastian and Kimia, 2002). Actually it becomes impractical if dimension exceeds twenty (Weber, 1998). In distance based indexing methods, the database is organised on the basis of the distance of the database elements with respect to one or more key/pivot elements. Elements with similar distance from the key elements are grouped into a cluster and when querying the database, triangle inequality is used to discard some clusters (Berman and Shapiro, 1999). This method reduces computational cost but has a poor indexing efficiency. Moreover, proper selection of key is crucial. A k nearest neighbour graph (knn graph) based scheme has been proposed by Sebastian and Kimia (2002) where each node represents a database element and it is connected to its k nearest neighbour. This method is not guaranteed to provide the nearest neighbour of the query image in spite of the extended neighbourhood concept. Moreover, for a large database it is not practical.

The precision of the set of retrieved images can be improved through a relevance feedback mechanism. As the importance of the features vary for different queries and applications, to achieve better performance, different emphases have to be given to different features and the concept of relevance feedback (RF) comes into picture. Relevance feedback is a learning mechanism to improve the effectiveness of information retrieval systems. For a given query, the CBIR system retrieves a set of images according to a predefined similarity measure. Then, user provides a feedback by marking the retrieved images as relevant to the query or not. Based on the feedback, the system updates the emphasis of individual feature and retrieves a new set. The classical RF schemes can be classified into two categories: query point movement (query refinement) and re-weighting (similarity measure refinement) (Rocchio, 1971). In the query point movement method, the goal is to improve the estimate of the ideal query point by moving it towards the relevant examples and away from bad ones. Rocchio's formula (Rocchio, 1971) is frequently used to improve the estimation iteratively. In (Huang et al., 1997), a composite query is created based on relevant and irrelevant images. Various systems like Quicklook (Ciocca et al., 2001), Drawsearch (Sciascio et al., 1999) have adopted the query refinement principle. In the re-weighting method, the weight of the feature that helps in retrieving the relevant images is increased and importance of the feature that hinders this process is reduced. Rui et al. (1998) have proposed weight adjustment technique based on the variance of the feature values. Systems like ImageRover (Sclaroff et al., 1997), RETIN (Fournier et al., 2001) use re-weighting technique. A close study of past work indicates that re-weighting technique is widely used.

Thus, the search module of a CBIR system has to deal with the number of interrelated issues like similarity measure, indexing and relevance feedback scheme. This is necessary to satisfy the diverging requirements like high precision and fast response. Keeping all these in mind, in this work, we have proposed a human perception based similarity measure and an indexing scheme and finally combined with a novel relevance feedback scheme.

The paper is organised as follows. Section 2 presents a human perception based similarity measure and an indexing scheme is presented in Section 3 which approximates the proposed similarity measure. Mann–Whitney test based relevance feedback scheme is presented in Section 4. Section 5 discusses the experimental system and result. Finally, it is concluded in Section 6.

2. Similarity measure

The collection of features (often referred to as feature vector) convey, to some extent, visual appearance of the image in quantitative terms. Image retrieval engine compares the feature vector of the query image with that of the database images and presents to the users the images of highest similarity (i.e., least distance) in order as the retrieved images. However, it must be noted that the elements in the feature vector carry different kinds of information: shape, texture and color, which are mutually independent. Hence, they should be handled differently as suited to their nature.

The early work shows that most of the schemes deal with Euclidean distance, which has number of disadvantages. One pertinent question is how to combine the distance of multiple features. Berman and Shapiro (1999) proposed following operations to deal with the problem:

$$\text{Addition : distance} = \sum_i d_i \tag{1}$$

where, d_i is the Euclidean distance of ith features of the images being compared. This operation may declare visually similar images as dissimilar due to the mismatch of only a few features. The effect will be further pronounced if the mismatched features are sensitive enough even for a minor dissimilarity. The situation may be improved by using

$$\text{Weighted sum : distance} = \sum_i w_i d_i \tag{2}$$

where, w_i is the weight for the Euclidean distance of ith feature. The problem with this measure is that selection on