Available online at www.sciencedirect.com
SCIENCE DIRECT®
The Journal of Systems and Software 78 (2005) 9–23
The Journal of Systems and Software
www.elsevier.com/locate/jss

Relevance feedback using adaptive clustering for image similarity retrieval

Deok-Hwan Kim [a,*], Chin-Wan Chung [b], Kobus Barnard [c]

[a] Department of Mobile Internet, Dongyang Technical College, 62-160 Kochuk-dong, Kuro-gu, Seoul 152-714, South Korea
[b] Division of Computer Science, Korea Advanced Institute of Science and Technology, 373-1, Kusong-dong, Yusong-gu, Taejon 305-701, South Korea
[c] Computer Science Department, University of Arizona, Gould-Simpson 730, Tucson, AZ 85721, USA

Received 2 September 2004; received in revised form 8 February 2005; accepted 8 February 2005
Available online 9 March 2005

Abstract

Research has been devoted in recent years to relevance feedback as an effective solution to improve performance of image similarity search. However, few methods using the relevance feedback are currently available to perform relatively complex queries on large image databases. In the case of complex image queries, images with relevant concepts are often scattered across several visual regions in the feature space. This leads to adapting multiple regions to represent a query in the feature space. Therefore, it is necessary to handle disjunctive queries in the feature space.

In this paper, we propose a new adaptive classification and cluster-merging method to find multiple regions and their arbitrary shapes of a complex image query. Our method achieves the same high retrieval quality regardless of the shapes of query regions since the measures used in our method are invariant under linear transformations. Extensive experiments show that the result of our method converges to the user's true information need fast, and the retrieval quality of our method is about 22% in recall and 20% in precision better than that of the query expansion approach, and about 35% in recall and about 31% in precision better than that of the query point movement approach, in MARS.

Keywords: Relevance feedback; Image database; Classification; Cluster-merging; Dimension reduction; Content-based image retrieval

1. Introduction

The relevance feedback based approach to content-based image retrieval (CBIR) has been an active research areas in the past few years. A good survey can be found in (Ishikawa et al., 1998; Rui et al., 1998; Rui et al., 1999). Most existing CBIR systems represent images as feature vectors using visual features, such as color, texture and shape. That is, the closer two vectors are, the more similar the corresponding images are. They search images via a query-by-example (QBE) interface. When the systems present a set of images considered to be similar to a given query, the user can pick up the ones most relevant to the given query, and the system refines the query using them, which allows the relevant images be the ones picked up by the user. Relevance feedback based CBIR techniques do not require a user to provide accurate initial queries, but rather estimate the user's ideal query by using relevant images feedback by the user.

Current approaches to CBIR assume that relevant images are physically near the query image in some feature space regardless of visual features. However, the similarity between images perceived by humans does not necessarily correlate with the distance between them

* Corresponding author. Tel.: +8222 6101914; fax: +8222 6101859.
E-mail addresses: dhkim@dongyang.ac.kr (D.-H. Kim), chungcw@islab.kaist.ac.kr (C.-W. Chung), kobus@cs.arizona.edu (K. Barnard).

0164-1212/$ - see front matter
doi:10.1016/j.jss.2005.02.005

in the feature space. That is, semantically relevant images might be spread out in the feature space and be scattered in several clusters rather than one. In this case, traditional relevant feedback approaches (Flickner et al., 1995; Ishikawa et al., 1998; Rui et al., 1997; Rocchio, 1971) do not work well when shifting the query center by linear combination of the relevant images.

Implementing the relevance feedback concerns the computation of a new query point (or points) in a feature space and the change of a distance function. As shown in Fig. 1(a), early studies (Ishikawa et al., 1998; Rui et al., 1997) represent a new query as a single point and change the weights of feature components to find an optimal query point and an optimal distance function. In this case, a single point is computed using the weighted average of all relevant images in the feature space. The contours represent equi-similarity lines. Meanwhile, a recent study (Porkaew and Chakrabarti, 1999) represents a new query as multiple points to determine the shape of the contour as shown in Fig. 1(b). This approach uses a clustering method (Charikar et al., 1997) to compute new query points using query results (relevant images) based on the user's relevance judgement. It is assumed that the relevant images are mapped to points close together according to the similarity measure. A single large contour is constructed to cover all query points and the system finds images similar to them. However, if the feature space and the distance function for the user's perception are quite different from those for the system, the relevant images are mapped to disjoint regions of arbitrary shapes in the feature space. That is, the relevant images may be ranked below other retrieved images for the given query. In order to converge rapidly to the user's information need at the higher semantic level, the system should find the images similar to any of the query points as in Fig. 1(c). A query that retrieves the images similar to any of the query points is called a disjunctive query. Especially, a complex image query is represented as disjoint multiple regions since semantically related images can be scattered in several visual regions rather than one.

In this paper, we propose a new adaptive classification and cluster-merging method (Qcluster: query clustering) to find multiple regions and their arbitrary shapes of contours for a given complex image query. Also we propose an approach to the relevance feedback using multiple query points to support disjunctive queries.

Fig. 2 shows the proposed relevance feedback mechanism. At the first stage, an example image submitted by the user is parsed to generate an initial query $Q = (q, d, k)$, where q is a query point in the feature space, k is the number of images in the query result returned by the system, and d is the distance function. The query point q is compared with images in the database using the distance function d. According to d, the result set consisting of k images close to q, $\text{Result}(Q) = \{p_1, \ldots, p_k\}$, is returned to the user.

At the next stage, the user evaluates the relevance of images in Result(Q) by assigning a relevance score to each of them. Based on those scores, the relevant set, $\text{Relevant}(Q) = \{p'_1, \ldots, p'_m\}$, is obtained. In this paper, we present a new adaptive clustering method consisting of two processes: the classifying process and the cluster-merging process. The proposed classifying process places each element of the relevant set, Relevant(Q), in one of the current clusters or a new cluster. Then, the proposed cluster-merging process finds the appropriate number of clusters by merging certain clusters to reduce the number of query points in the next iteration. Finally, representatives of clusters generated from relevant images in the classified set make up the set of new query points. A new query, $Q' = (q', d', k)$ with a set of new query points q' and a new distance function d', is computed and then used as an input for the second round.

After some iterations, the loop ends up with the final result set close to Result(Q_{opt}), where $Q_{\text{opt}} = (q_{\text{opt}}, d_{\text{opt}}, k)$ is the optimal query.

Our approach to the relevance feedback allows multiple objects to be a query. We refer to them as a multipoint query. When the user marks several points as relevant, we cluster sets of relevant points and choose the centroids of the clusters as their representatives. Then, we construct the multipoint query using a small number of good representative points. At the classifying process, Bayesian classification function (Fisher, 1938) is used. Statistics such as mean and covariance of each cluster, which were computed from the previous itera-

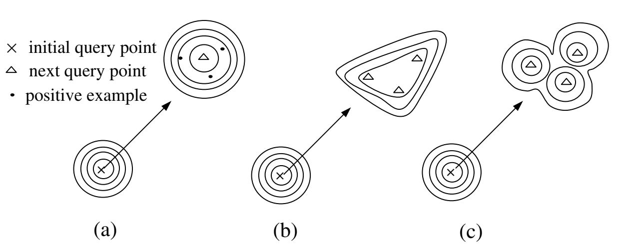

Fig. 1. Query shape. (a) Query point movement (b) Convex shape (multipoint) and (c) Concave shape (multipoint).

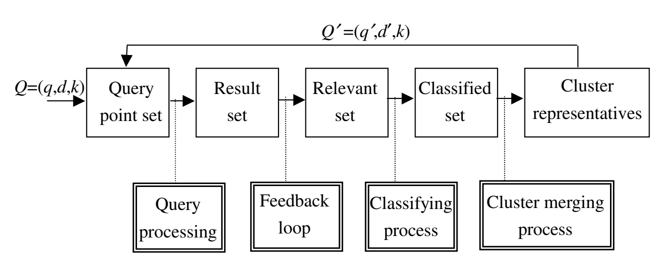

Fig. 2. Overall structure of the proposed method.