SPEEDING UP RELEVANCE FEEDBACK IN IMAGE RETRIEVAL WITH TRIANGLE-INEQUALITY BASED ALGORITHMS

Ziyou Xiong, Xiang Zhou, William M. Pottenger† and Thomas S. Huang

Dept. of Electrical and Computer Engineeing
Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign
Urbana, IL, 61801
E-mail: {zxiong, xzhou2, huang}@ifp.uiuc.edu

†Ph.D. Lehigh University and the
National Center for Supercomputing Applications
19 Memorial Drive West, Bethlehem, PA 18015
E-mail: billp@ncsa.uiuc.edu

ABSTRACT

A content-based image retrieval(CBIR) system has been constructed to integrate relevance feedback with triangle-inequality based algorithms. The system offers typically 20 to 30 times faster retrieving speed with minimum sacrifice of retrieval performance on Corel database consisting of more than 17,000 images. The theoretic framework is built by using triangle-inequality based algorithms at sub-feature level and using relevance feedback techniques at feature level. Results show retrieval performance is clearly improved over the approach with only triangle-inequality based algorithms. A new high level weight updating method for the hierarchical distance model for relevance feedback is proposed.

1. INTRODUCTION

Relevance feedback is one promising technique trying to bridge the gap between high level concepts and low level features[1] in CBIR. This learning technique in general tries to update the similarity measure in form of a generalized weighted Euclidean distance, then uses this measure to compare the distance between query image and images in the database. When the number of images in the database is very large, it is extremely time-consuming to calculate all the distances. Also sorting the distances can take a long time. For large image databases, relevance feedback can not help to solve this long computation problem.

To tackle this problem in large image database retrieval, a number of triangle-inequality based algorithms have been proposed from the University of Washington[2]. In their scheme, several sets of key images are pre-selected. The triangle-inequality principle is used to eliminate any database image whose lower bound of distance to the query image is larger than a given threshold. The number of key images can be made much smaller than that of the database images, thus a majority of direct distance calculations can be avoided. Hence the retrieval speed is greatly increased. However in [2] the system lacks the performance-boosting schemes like relevance feedback so the fixed distance measure can hardly satisfy different users' ways of defining their own similarity measures.

A scheme integrating the aforementioned two approaches can accompany each other for large image database retrieval applications. This paper is a study on the feasibility of such a scheme. The rest of the paper is organized as follows. Idea on integrating the above two schemes is described in Section 2. Our proposed method is presented in Section 3. A new high level weight updating method for the hierarchical distance model for relevance feedback [3] is discussed in Section 4. Section 5 details our experiments and observations. Section 6 concludes and discusses some future research directions.

2. INTEGRATING RELEVANCE FEEDBACK AND TRIANGLE-INEQUALITY BASED ALGORITHMS

Before integrating the above two approaches, there is one major problem that needs to be solved. Triangle-inequality based algorithms have to know the distance measure with respect to a visual feature beforehand. But relevance feed-

This work was supported partly by National Science Foundation Grant CDA 96-24396 and EIA 99-75019

back keeps updating distance measure function. Fortunately there exists more than one layer of distance measure in the relevance feedback framework. For example, the overall similarity between two images depends on how similar they are in color, in texture and in shape and so on. Also let how similar they are in color depend on how similar they are in red, green and blue. Triangle inequality based algorithm can be used at the red, green and blue level using a fixed-weighted distance measure and relevance feedback can be used to update weights associated with color, texture and shape and so on. In this way the best of two worlds can be put together.

3. PROPOSED APPROACH

A relevance feedback CBIR system with triangle-inequality based algorithms is proposed. The relevance feedback is partial in that only the weights associated with visual features are updated. Remember these weights are at higher level in the weight hierarchical tree structure as in [1]. A normalized Euclidean distance measure with fixed weights is used for sub-feature level distance formation.

Parallel to the problem formation in [3] and follow their term symbols, our problem is formed as follows: Find an ideal query Q = $[\vec{q_1}, ..., \vec{q_i}, ... \vec{q_M}]$ from those images specified as relevant by the user to minimize:

$$J = \vec{\pi}^T \cdot \vec{d} \tag{1}$$

where M is the number of visual features, subject to the following constraint:

$$\sum_{i=1}^{M} \sqrt{u_i} = 1 \tag{2}$$

where elements of $\vec{\pi}$ are relevance weights given to relevant images and elements d_n of $\vec{d} = [d_1, ..., d_n, ..., d_N]^T$ represents overall distance between the n-th relevant images and the ideal query Q. N is the number of relevant images. d_n is explicitly expressed as a linear combination of distances g_{ni} with respect to the i-th feature such as color, texture and shape.

$$d_n = \vec{u}^T \cdot \vec{g_n} \tag{3}$$

and $\vec{g_n} = [g_{n1}, ..., g_{ni}, ..., g_{nM}]^T$. In our system color, texture and structure features are used, so there M = 3.

$$g_{ni} = [(\vec{x_{ni}} - \vec{q_i})^T (\vec{x_{ni}} - \vec{q_i})]^{1/2} \tag{4}$$

is an Euclidean distance with fixed weights and i = 1, ..., M.

With Lagrange multiplier method, the above problem is transformed into minimizing the following entity: $L = \vec{\pi}^T \times \vec{d} - \lambda(\sum_{i=1}^{M} \sqrt{u_i} - 1)$.

Solution to q_i is:

$$\vec{q_i}^T = \frac{\vec{\pi}^T X_i}{\sum_{n=1}^{N} \pi_n} \tag{5}$$

where each row of matrix X_i is a feature vector representing the i-th feature for an image. The derivation is very similar to that in [3, 4].

Solution of u_i is such that first derivative of L with respect to u_i is zero, i.e,

$$\frac{dL}{du_i} = \sum_{n=1}^{N} \pi_n g_{ni} + \frac{-\lambda}{2\sqrt{u_i}} = 0 \tag{6}$$

If $\sum_{n=1}^{N} \pi_n g_{ni}$ is denoted as f_i then u_i is solved to be: $u_i = (\frac{\lambda}{2 f_i})^2$. Plug u_i into $\sum_{i=1}^{M} \sqrt{u_i} = 1$, we have $\frac{\lambda}{2} = \frac{1}{\sum_{i=1}^{M} \frac{1}{f_i}}$. Thus:

$$u_i = (\frac{\frac{1}{f_i}}{\sum_{j=1}^{M} \frac{1}{f_j}})^2 \tag{7}$$

With distance measure specified in Equation (4), a set of key images can be selected for each visual feature. The k-means clustering method is used to cluster the database and the centroid image from each of these clusters is chosen as key image. This method is one of the methods used in [2].

Algorithm speeding up relevance feedback with triangle-inequality is summarized as follows:

1. Initialization:
   - For each visual feature $v_1, v_2, ..., v_M$, the weight vector $\vec{u}$ is initialized with $u_i = \frac{1}{M^2}$ for i = 1, ..., M. In our system M = 3.
2. Iterations:
   - Let the user select images as relevant ones together with their degrees of relevance.
   - Form a query Q based on Equation (5).
   - for each image I in the image database do begin
     (a) calculate the lower bound of distance between I and Q for each feature i over all key images. Let us denote it as l_i. Put all of them in vector form and we get $\vec{l} = [l_1, ...l_i, ..., l_M]^T$.
     (b) Calculate the distance $f_i = \sum_{n=1}^{N} \pi_n g_{ni}$ between Q and relevant images for each visual feature i.
     (c) Update $\vec{u}$ using Equation (7).
     (d) Calculate lower bound of the overall distance between I and Q. Denote it as LB(Q, I) = $\vec{u}^T \cdot \vec{l}$.

(e) Compare LB(Q, I) with a given threshold T. If LB(Q, I) $\geq$ T, then eliminate I.

- Sort those images that are not eliminated according to their lower bounds in ascending order. Return a specific number of images with shortest distance to Q. In our system, this number is 20.

3. Continue the above iteration until convergence.

4. DISCUSSION

The constraint on u_i's in Equation (2) in [3] is a bit different from the one proposed in this paper. There the constraint takes the following form:

$$\sum_{i=1}^{M} \frac{1}{u_i} = 1 \tag{8}$$

and the update formula for u_i is:

$$u_i = \frac{\sum_{j=1}^{M} \sqrt{f_j}}{\sqrt{f_i}} \tag{9}$$

The relationship between u_i and f_i in both Equation (9) and Equation (7) is worth studying. In both of these equations, when f_i changes while all f_j's are kept constant for j = 1, ... i-1, i+1, ..., observations are that the bigger f_i is, the smaller u_i is. This is in agreement with intuition about relevance feedback. The bigger the distance between two images with respect to one feature is, the less important this feature is as far as overall distance is concerned. This means both of them can be used to derive u_i's.

From Equation (9), the following relationship can be derived:

$$\frac{u_i}{u_j} = \sqrt{\frac{f_j}{f_i}} \tag{10}$$

while from Equation (7), the relationship becomes:

$$\frac{u_i}{u_j} = (\frac{f_j}{f_i})^2 \tag{11}$$

Results in the above two equations show that relation in Equation (11) is a better choice to execute relevance feedback as it represents an squared inverse relationship between u_i and f_i instead of squared root inverse. In our proposed relevance feedback scheme, this relationship is adopted. This is also the reason the constraint in Equation (2) is used rather than the one in Equation (8) as in [3].

5. EXPERIMENTAL RESULTS

5.1. Testbed, Features, Key Selection and System

The Corel image database is used as the testbed. For each image group labeled by experts, the first image is used as the initial query. Only those retrieved images that are in the same group as the initial query image are regarded as relevant and are used to form the new query. In all our experiments, equal weights are assigned to relevant images. Although our system supports multiple level weight-assigning functionality, equal weights are used in order to obtain an objective evaluation.

Fig. 1. Interface of the Content-based Image Retrieval System. The lower-left button is for randomly displaying and browsing images. Sliders are used to specify query image.

An image retrieval system is developed based on the MARS system at University of Illinois at Urbana-Champaign[5]. The numbers of relevant images retrieved in the top 20 returns are recorded. Interface of the system is in Figure 1. The same set of visual features(color, texture and structure) are used as those in the system in [3]. 9 key images for color feature, 10 key images for texture feature and 18 for structure feature are selected using the k-means clustering algorithm as in [2].

5.2. Execution time, Retrieval Performance

On average, it takes 1.5 seconds for MARS to calculate the distance measure, calculate distances between query and all images and sort them. In camparison, it takes less than 0.05 seconds for the proposed system on a Pentium 3 600Hz processor. There are more than 17,000 distance calculations(vector multiplication) and sorting for MARS but only 37 for the proposed system, 17,000 distance comparison(only addition and subtraction operations) and sorting less than 500 distances. The significant speed gain is no surprise because of the triangle-inequality based algorithms.

One display of the retrieval results is shown in Figure 2. For each initial query image, four iterations are done with

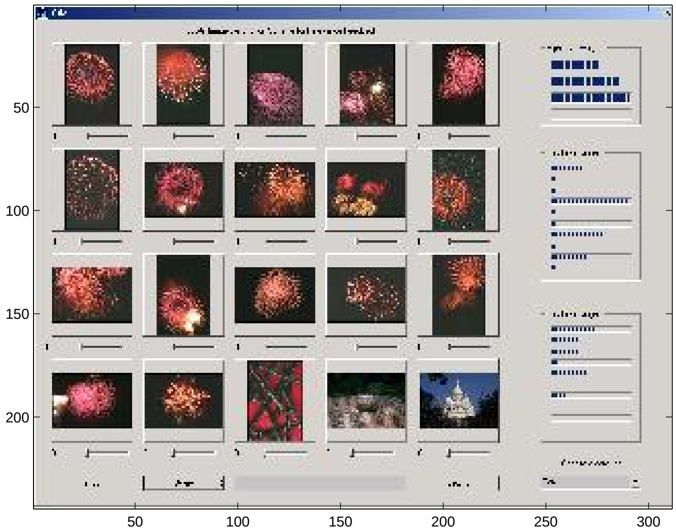

Fig. 2. Retrieval results using the upper-left image as query.

both systems. The average number of relevant returns over all initial queries is put into the following table.

Table 1: Number of Relevant Images in Top 20 Returns

	$0_{th} Iteration$	1_{st}	2_{nd}	3_{rd}
MARS	2.6784	3.2982	3.6433	3.7018
Proposed	2.4442	3.0854	3.2965	3.4725

The following observations can be drawn:

1. Relevance feedback improves the retrieval performance over the fix-weighted system such as FIDS[2]. The results on the first column in the above table correspond to those achieved by systems using triangle-inequality based algorithms alone.

2. The average retrieval performance of our approach is competitive against but not so good as MARS which uses multi-level weight updates for relevant feedback. This is in agreement with our expectation. Weight updates at sub-feature level boost retrieval performance for MARS, but they are omitted in the proposed system in exchange for computation gain. Our effort in developing weight updates as in Equation (7) has some positive gain in performance. In fact our record shows that of all 171 classes that have been tested, our approach out-performs MARS in 30 classes, under-performs in 71 classes and performs equally in 70 classes.

6. CONCLUSIONS AND FUTURE WORK

An image retrieval system has been constructed to integrate relevance feedback with triangle-inequality based algorithms. The system offers typically 20 to 30 times faster retrieving speed with minimum sacrifice of retrieval performance on Corel database consisting of more than 17,000 images. Taking into account the significant gain in speed and the loss in performance, we believe the research efforts are worthwhile.

Several research directions are ahead. How to improve retrieval performance of the current system may be the first choice. Better key images selection can contribute from one direction. Better relevance feedback schemes such as the ones using both positive and negative samples can contribute from another direction. We are also looking for better image features, exploring compressed domain techniques.

7. ACKNOWLEDGEMENT

We'd like to thank Dr. Irwin E. Alber, Dr. Morton S. Farber from Boeing Corporation and Ms. Nancy Yeager from National Center for Super-computing Applications for their kind support and encouragement. The third author, William M. Pottenger, gratefully acknowledges the aid of his Lord and Savior Jesus Christ in his work.

8. REFERENCES

[1] Yong Rui, Thomas S. Huang, Michael Ortega, and Sharad Mehrotra, "Relevance feedback: A power tool for interactive content-based image retrieval," *IEEE Tran on Circuits and Systems for Video Technology, Special Issue on Segmentation, Description, and Retrieval of Video Content*, vol. 8, pp. 644–655, September 1998.

[2] A. Berman and L. G. Shapiro, "A flexible image database system for content-based retrieval," *Computer Vision and Image Understanding*, vol. 75, pp. 175–195, July/August 1999.

[3] Yong Rui and Thomas Huang, "Optimizing learning in image retrieval," *Proc. of IEEE Int Conf on Computer Vision and Pattern Recognition(CVPR)*, June 2000.

[4] Y. Ishikawa, R. Subramanya, and C. Faloutsos, "Mindreader: Query databases through multiple examples," *Image Processing, ICIP 99. Proceedings. 1999 International Conference on*, vol. 2, pp. 570 –574, 1999.

[5] Y. Rui, T. S. Huang, and S. Mehrotra, "Contented-based image retrieval with relevance feedback in mars," *Image Processing, ICIP 97. Proceedings. 1997 International Conference on*, October 1997.