

Enhance the Technique of Relevance Feedback for Content-based Multimedia Retrieval by using Mining Algorithm

Aasma S. Mujawar

Department of Computer Engineering,
Smt.Kashibai Navale College of Engineering,
Off Singhad Road,Vadgaon(Bk),
Pune- 411041,India.

Kosbatwar Shyam P.

Department of Computer Engineering,
Smt.Kashibai Navale College of Engineering,
Off Singhad Road,Vadgaon(Bk),
Pune- 411041, India.

ABSTRACT

Today images, multimedia are immensely important in information retrieval system. In existing relevance feedback technique, there is semantic gap between high level concepts and low level features of images as well as videos, another drawback is according to user requirement we cannot retrieve relevant multimedia data (images, videos) from multimedia database and image database. To overcome from this drawback in Content based Multimedia Retrieval (CBMR), using navigation pattern relevance feedback technique to retrieve most relevant videos, images from multimedia data according to user requirement. To provide efficient and effective retrieval of content based multimedia data and images from multimedia database like video data, images by using relevance feedback technique and mining algorithm.

General Terms

Data mining algorithm: K-means, Apriori algorithm.

Keywords

Content-based Multimedia Retrieval, Relevance Feedback, Query Image Reweighting, Query Expansion, Query Point Movement.

1. INTRODUCTION

Today, images and multimedia are immensely important in information retrieval system. Retrieving useful information from large scale database is called as data mining. In text mining, database contain file names, data description, some keywords which create huge, redundant data in database. To overcome from these drawback content-based image and multimedia and image retrieval system has been introduced. Therefore number of image retrieval algorithm proposed for content based image retrieval system to extract features like color, texture, shape. In these retrieval techniques like query reweighting semantic gap between high level concepts (user's concepts) and low level features of images. Another problem is that extracted images are not relevant images as per user requirement. This problem is called as superfluous browsing. To solve such problems relevance feedback technique is introduced to extract relevant images for images as well as video retrieval. In existing RF studies like Query By Example, QBIC there is common problem of redundant browsing. The main problem is that to extract relevant images from given query as image according to user's requirement. Relevance feedback means users pick image from image dataset and retrieve relevant images until user does not satisfied with retrieved images.

In terms of effectiveness, the proposed direction-finding pattern search algorithm consists of three query strategies consist of Query Point Movement (QPM), Query Reweighting (QR) and Query Expansion (QEX) to deal with visual multiplicity problem. Using these techniques relevant images can retrieve by optimizing retrieval iterations.

2. RELATED WORK

A **Relevance-Feedback** based approach to CBIR, in which a computer interacts to refine high-level queries to representations based on low-level features. Relevance feedback is a powerful technique used in traditional text-based information retrieval systems. It is the process of automatically adjusting an existing query using the information feedback by the user. In previously retrieved objects such that the adjusted query is a better approximation to the user's information need. In the relevance-feedback-based approach, the retrieval process is interactive between the computer and the human. Under the assumption that high-level concepts can be captured by low-level features, the relevance feedback technique tries to establish the link between high level concepts and low-level features from the user's feedback. Furthermore, the burden of specifying the weights is removed from the user. The user only needs to mark which images he or she thinks are relevant to the query. It has proven to be very effective for improving retrieval performance. Relevance feedback is used to reduce semantic (meaning) gap in images.

Low level features: 1. Color
2. Texture
3. Shape

High level concepts: User intention or requirement. (Example: **Fresh** Apple) Relevance feedback is used to reduce semantic (meaning) gap in images. That is mapping of low level features and high level concepts.

- **The gap between high level concepts and low level features.** The assumption that the computer-centric approach makes is that the high-level concepts to low-level features mapping is easy for the user to do. While, in some cases, the assumption is true, e.g., mapping a high-level concept (**Fresh** Apple) to low-level features (color, texture shape).
- **The subjectivity of human diversity.** Different persons, or the same person under different situation, may perceive the same visual content differently. This is called human intension diversity. The subjectivity exists at various levels. For example, one person may be more interested in an image's color feature, while another may be more interested

in the texture feature. Even if both people are interested in texture, the way they perceive the similarity of texture may be quite different. To solve such multiplicity problem one novel method proposed names user satisfied Relevance Feedback to achieve improved retrieval quality of CBMR Content-based Multimedia Retrieval using navigation pattern has been proposed .

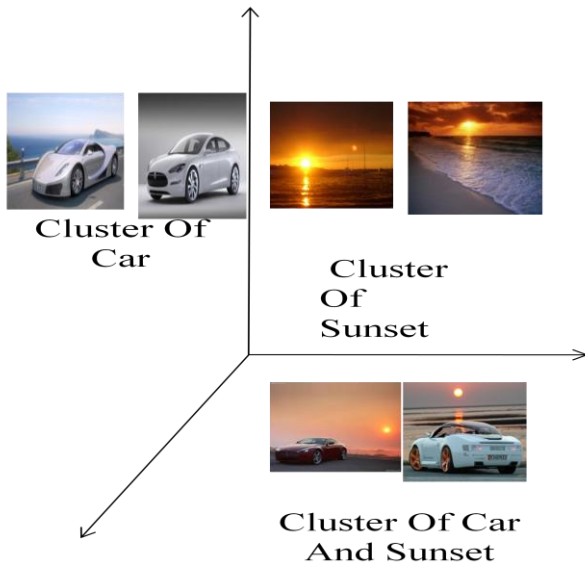


Figure 1: Example Of Visual Multiplicity

A set of approaches as per user's requirement retrieval proposed like Photobook does not give exact description of semantic features. Such problems are called as Visual Multiplicity problem. These problems are widely present in real time image application. Therefore maintain log tables and improve query retrieval strategies which help to improve retrieval quality of images as well as videos in area of RF.

2.1 Query- Image Reweighting:

In previous work contain what image features are important for those images (relevant images) select up by the users at each feedback (iteration).The notion behind QR is that, if the i th feature f_i exists in positive (relevant) examples frequently, the system assigns the higher degree to f_i . QR like approach was first proposed by Rui , which convert image feature vectors to weighted-term vectors in early version of Multimedia Analysis and Retrieval System (MARS). In query reweighting, feature weights are keep on changing to connect with high level user concepts and low level image features.

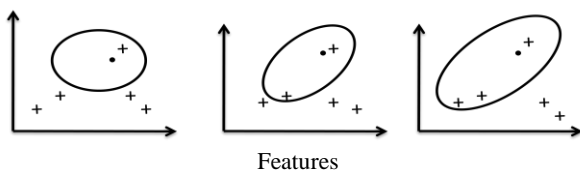


Figure 2: Relevance Feedback with Query Reweighting

2.2 Query -Image Point Movement

To improve the accuracy of image retrieval is moving the query point towards the point of the user's intension in feature space. It gives positive (relevant) examples as a new query point at each iterations. After changing or updating new query point at each feedback, the query point should be move to

region of the user's requirement. The well-known space-vector formula proposed by Rocchio is as follows:

$$Q_i = Q_{i-1} + \alpha \sum_{j=1}^{nr} R_j / nr - \beta \sum_{j=1}^{nir} IR_j / nir$$

where. Q_i is the vector of the i^{th} query.
 IR_j is the vector of the j th irrelevant image.
 nr is cardinality of relevant images.
 nir is cardinality of irrelevant images.

2.3 Query-Image Expansion (QEX)

Query Reweighting and Query Point Movement cannot increase the quality of Relevance Feedback, Query Expansion has been another technique in the solution to increase quality of image retrieval recently. The search strategies, such as QR and QPM, cannot completely satisfy the user's interest to extract relevant images .Using Query Expansion, user picked image feature added into image result set at each level of iteration. For this reason, the modified version of MARS [9] groups the similar relevant points into several clusters, and selects good points from these cluster to the multipoint query

3. PROPOSED SYSTEM

3.1 Relevance Feedback Workflow

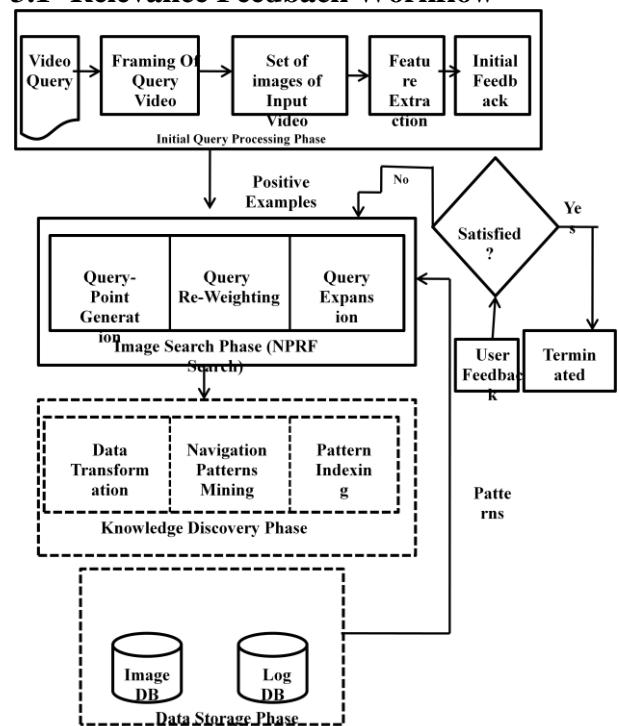


Figure 3: Workflow of Relevance Feedback

Initial Query Processing Phase:

In this phase the feature weight, retrieve the visual features from the original query image to find the similar images. From relevant images dataset user select or picked up re further analyzed at the first iteration.

Image Search Phase:

In the search phase, user intention is to extend the one search point to multiple search points by adding relevant pattern. The proposed user intention of relevance feedback for image and multimedia retrieval search algorithm can help to

overcome from visual multiplicity image problem. In this phase, a new query point at each feedback is generated by the adding relevant images examples. Then, using the k-nearest images, new query point can be found by expanding the weighted query. The search procedure does not stop unless the user is satisfied with the retrieval results

Knowledge Discovery Phase:

By adding user behaviour in image retrieval can be viewed as one type of knowledge discovery. As result, this phase concerns with the construction of the user satisfied relevance feedback model by taking relevant patterns from user browsing behaviour.

Data Storage Phase:

Datawarehouse which store integrated, time-variant and nonvolatile collection of useful data including images, navigation patterns, log files, image features. The data storage is very helpful to improve the quality of image retrieval. In this data phase calculating association mining rule from the image databases can be calculated and updated dynamically to maintain dataset for getting positive examples.

The first query process is called **initial feedback**. From dataset of relevant images, user pick relevant images valuable information to the image search phase, including new feature weights, new query point and the user’s intention. Then, by using three search strategies, with respect to QPM, QR and QEX are integrated together to find the desired images.

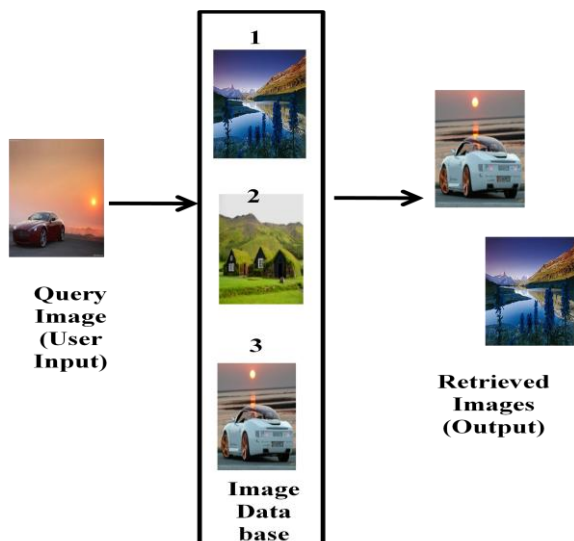


Figure 4: Example Of Visual Multiplicity

3.2 Relevance Feedback (User Intention)

Algorithm

- Step1. Start
- Step 2. Initially feature extraction of all the images of the database is performed.
- Step 3. User inputs a query image.
- Step 4. Feature extraction of query image is performed.
- Step 5. Relevance feedback technique is implemented using query reweighting, query expansion, query point movement done.
- Step 6. Result is displayed.
- Step 7. If user satisfied then searching get finished.
- Step 8. Else repeat process of query reweighting, query expansion, query point movement technique.
- Step 9. User puts his feedback after examining the searched result.

Step 10. User repeats this step until and unless he is not satisfied.

Step 11. End.

In this RF, basically, the frequent relevant images retrieved from the user logs and find useful browsing paths to reduced the search direction on RF. In this approach, the users’ common interests can be represented by the discovered frequent patterns (also called frequent itemsets). Through these patterns, the user’s intention can be correctly captured in a reduced query process. In this phase, the Apriori-like algorithm is performed to discover navigation patterns using the transformed data. This pattern finding process is divided into two step

Step 1:

Construction of the cluster transaction table. From table 1, consider, select five query sessions as an example shown in Table 1. A query session can be considered a transaction., the transaction is composed of a query image item and several iteration image items To extract valuable relevant patterns, all query sessions in the transformed log table are called as **transaction table**. From Table 1, we can discover some interesting phenomena. First, though passing different iterations, the paths starting with the same query item lead to the same destination, e.g., Session 1 and Session 2. Second, the paths starting with different query items lead to the same destination, e.g., Session 2 and Session 4. Third, the paths starting with the same query item lead to different destinations, e.g., Session 3 and Session 4. Even the special path, Session 5, may be the other important trail for image finding. This step helps to maintain session of images.

Step 2:

Generation of relevant patterns: This operation emphasize on mining valuable patterns to facilitate image retrieval. As shown the frequent itemsets X whose supports support(X) exceed the presetting minimum support minsup are mined by Apriori-like algorithm. In fact, the generated frequent itemsets can be regarded as sequential relevant patterns For example, as shown in Table 2, the sequential navigation pattern fC11 C32 C42 derived from frequent itemset C11;C32;C42 under the minimum support is 2.

Table 1: Example of Relevant Cluster Patterns

Query Session ID	Item
1	C ₁₁ ,C ₂₁ ,C ₃₂ ,C ₄₂
2	C ₁₁ ,C ₂₃ ,C ₃₂ ,C ₄₂
3	C ₁₂ ,C ₂₁ ,C ₃₁ ,C ₄₂
4	C ₁₂ ,C ₂₁ ,C ₃₁ ,C ₄₂
5	C ₁₃ ,C ₂₂ ,C ₃₂ ,C ₄₃

3.3 Mathematical Model:

1. Assume that by K-means algorithm set of images are found when selection of specific image as input by user .Such set of images consider as query image point (QPImg)_{old} at

$$O = \sum_{x=1}^k \frac{\sqrt{\sum_{j=1}^d (f_j^x - f_j^{qpiold})^2}}{d}$$

2. Visual features of positive example G picked by user then first

Then first averaged into new query point (QPImg)_{new}

3. Consider a set of positive examples
 $G = \{q_1, q_2, q_3, \dots, q_n\}$ and
 Dimensions of the i th feature $F_i = \{f_1, f_2, f_3, \dots, f_n\}$
 4. From the positive example. There upon, the new query point
 $(QPI_{img})_{new}$ implied by G can be define
 $(QPI_{img})_{new} = \{F_1, F_2, \dots, F_b\}$
 Meanwhile, $(QPI_{img})_{new}$ and the positive examples are stored in the log database to enhance the knowledge database. Next, the negative examples are appended to the accumulated into negative set .5. Feature reweighting formula. The new weight of the i th feature F_i is defined as initial feedback. Similarly, for video retrieval first video is divided into number of images using framing technique.

$$w_i = \frac{\frac{\sum_{y=1}^b \sigma_y}{\sigma_i}}{\sum_{z=1}^b \frac{\sigma_i}{\sigma_z}}$$

4. CONCLUSION

To remove long iteration problem of CBMR with RF, it has presented a approach named User Satisfied RF by integrating pattern mining search approach named Pattern Search. In summary, the main feature of RF is to efficiently reduce the retrieval quality of interactive CBMR. On other hand to, relevant patterns derived from the users' long browsing behaviors for minimizing the number of user feedback to match the user's intention by merging three query refinement strategies. As a result, traditional problems such as visual multiplicity and exploration convergence are solved.

5 REFERENCES

[1] Ja-Hwung Su, Philip S. Yu, Wei-Jyun Huang ,Efficient Relevance Feedback for CBIR by Mining User Navigation Patterns IEEETrans On Knowledge And Data Engineering VOL. 23,No March 2011.
 [2] Peng-Yeng Yin ,Bir Bhaneu Integrating Relevance Feedback Techniques for Image Retrieval Using Reinforcement Learning IEEETrans on Pattern analysis and Machine Intelligence, vol. 27.No 10, October 2005.
 [3] K. Porkaew, K. Chakrabarti, and S. Mehrotra, Query Refinement for Multimedia Similarity Retrieval in

MARS, Proc. ACM Int'l Multimedia Conf. (ACMMM), pp. 235-238, 1999.
 [4] Mianchu Chen,Ping Fu,Yuan sun,Hui zhang Image Retrieval Based on Multi-feature similarity score fusion using Genetic Algorithm_The 2nd International Conference on Computer and Automation Engineering (ICCAE), 2010.
 [5] P.Y. Yin, B. Bhanu, K.C. Chang, and A. Dong, Integrating Relevance Feedback Techniques for Image Retrieval Using Reinforcement Learning, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no 10, pp. 1536-1551, Oct. 2005.
 [6] X.S. Zhou and T.S. Huang, Relevance Feedback for Image Retrieval: Comprehensive Review, Multimedia Systems, vol. 8,no. 6, pp. 536-544, Apr. 2003.
 [7] R. Fagin, Combining Fuzzy Information from Multiple Systems, Proc. Symp. Principles of Database Systems (PODS), pp. 216- 226, June1996..
 [8] R. Fagin, Fuzzy Queries in Multimedia Database Systems, Proc.Symp. Principles of Database Systems (PODS), pp. 1-10, June 1998.
 [9] Mianchu Chen,Ping Fu,Yuan sun,Hui zhang Image Retrieval Based on Multi-feature similarity score fusion using Genetic Algorithm_The 2nd International Conference on Computer and Automation Engineering (ICCAE), 2010.
 [10] K. Vu, K.A. Hua, and N. Jiang, Improving Image Retrieval Effectiveness in Query-by-Example Environment, Proc. 2003 ACM Symp. Applied Computing, pp. 774-781, 2003.
 [11] A. Pentland, R.W. Picard, and S. Sclaroff, Photobook: Content-Based Manipulation of Image Databases Int'l J. Computer Vision(IJCV), vol. 18, no. 3, pp. 233-254, June 1996.
 [12] J.J. Rocchio, Relevance Feedback in Information Retrieval. The SMART Retrieval System—Experiments in Automatic Document Processing, pp. 313-323, Prentice Hall, 1971.