ISSN: 2321-7782 (Online)
Volume 2, Issue 3, March 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Paper / Case Study
Available online at: www.ijarcsms.com

RF for Content-Based Image Retrieval by Mining Navigation Patterns

Anju C R[1]
M-tech Computer Science and Engineering
KMCT College of Engineering Calicut
Calicut University
Calicut City – India

Subi S[2]
Assistant Professor, Department of CSE
KMCT College of Engineering Calicut
Calicut University
Calicut City – India

Abstract: CBIR process extracts the image features first then it compares with images from the database and returns the results to the user. These methods are impractical for real applications due to redundant browsing and exploration convergence. Thus proposes a new user navigation pattern based feedback system to support content based image retrieval. Using NPRF method, high quality of image retrieval on RF can be achieved in limited number of feedbacks and navigation pattern obtained by user query log. Again to refine the search by reweighting of query terms based on the distribution of these terms in the relevant and non relevant documents retrieved is done in response to those queries. The experimental results reveal that NPRF outperforms other existing methods significantly in terms of precision, coverage, and number of feedbacks. In view of very large data sets, this technique includes parallel and distributed computing techniques.

Keywords: Image Retrieval, Content-based image retrieval, Relevance feedback, Feature extraction, Navigation pattern.

I. INTRODUCTION

Image retrieval is an important topic in the field of pattern recognition and artificial intelligence. Image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval utilize some method of adding metadata such as captioning, keywords, or descriptions to the images so that retrieval can be performed over the annotation words. In CBIR, images are indexed by their visual content, such as color, texture, shapes. The semantic gap between low-level features and high-level concepts handled by the user is one of the main problems in image retrieval. On the other hand, the relevance feedback has been used on many CBIR systems such as an effective solution to reduce the semantic gap. The goal in image analysis is to extract useful information for solving application-based problems. The first step to this is to reduce the amount of image data using methods after the features have been extracted, then analysis can be done.

Content-based means that the search analyzes the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. The term "content" in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself. Combining CBIR search techniques available with the wide-range of potential users and their intent can be a difficult task. An aspect of making CBIR successful relies entirely on the ability to understand the user intent. CBIR systems can make use of relevance feedback, where the user progressively refines the search results by marking images in the results as "relevant", "not relevant", or "neutral" to the search query, then repeating the search with the new information.

In this project Navigation-Pattern-based Relevance Feedback (NPRF) is used to achieve the high retrieval quality of CBIR with RF by using the discovered navigation patterns. To solve problems propose an approach called Navigation-Pattern-Based Relevance Feedback (NPRF) which is used to get efficient and effective image retrieval quality with user interaction. CBIR combined with RF is a widely used technique to get high quality of image retrieval. Relevance feedback is a tool where end user involvement is more to improve the performance of the system. CBIR with RF are combined to discover Navigation Patterns that takes user intentions taken in to account.

II. LITERATURE SURVEY

The most relevant work is the work done by M.D. Flickner, H. Sawhney, W. Niblack, [1] and they devolepe QBIC system to explore CBIR l methods. QBIC allows queries on large image and video databases based on example images, sketches & drawings, color &texture patterns. QBIC uses of image and video, content-computable properties of colour, texture, shape and motion of images, videos and their objects-in the queries.

In the work of D.Harman, [2] issue was the relationship between term reweighing (modifying term weights based on term used in relevant and non-relevant documents) and query expansion (the addition of new terms from relevant documents). The work investigates the relationship between term reweighting and query expansion, and most results for query expansion and showed that adding only selected terms from retrieved relevant documents was more effective than adding all the terms. The selection measures were based on picking the top 20 from ail retrieved relevant documents for a given query

Y. Rui, T. Huang, M. Ortega, and S. Mehrotra, [3] systems is built to establish the basis of CBIR the usefulness of the proposed approaches is limited the gap between high level concepts and low level features and subjectivity of human perception of visual content .This paper proposes a relevance feedback based interactive retrieval approach which effectively takes into account the above two characteristics in CBIR. During the retrieval process the users high level query and perception subjectivity are captured by dynamically updated weights based on the users feedback

X. Jin and J.C. French, [4] semantically related images are often scattered across several visual clusters. Although traditional Content-based Image Retrieval (CBIR) technologies may utilize the information contained in multiple queries (gotten in one step or through a feedback process), this is only a reformulation of the original query. In this paper, we present a novel approach to relevance feedback which can return semantically related images in different visual clusters by merging the result sets of multiple queries. Further research topics, such as achieving candidate queries' visual diversity. As a result these strategies only get the images in some neighbourhood of the original query as the retrieval result. This severely restricts the system performance. Relevance feedback techniques are generally used to mitigate this problem. K. Vu, K.A. Hua, and N. Jiang, [5] Query-by-example is the most popular query model for today's image retrieval systems. A typical query image contains not only relevant objects (e.g., Eiffel Tower), but also irrelevant image areas (e.g., the background). Referred to as noise in this paper, has limited the effectiveness of existing image retrieval systems

X.S. Zhou and T.S. Huang, [6]. The retrieval of images based directly on their visual content (content-based image retrieval, CBIR) a search session begins by presenting an example image (or to the search engine as a visual query, then the engine returns images that are visually similar to the query image. More recently, the concept of semantic gap has been extensively used in the CBIR research community to express the discrepancy between the low-level features that can be readily extracted from the images and the descriptions that are meaningful for the users. The RF mechanism implemented in a search engine should attempt to minimize the amount of interaction between the user and the engine required for reaching good results. In the following, we present the main issues related to relevance feedback for image retrieval and we review recent developments in this domain. We then mention a few promising research directions for the near future.

III. EXISTING SYSTEM

A number of powerful retrieval algorithms have been proposed. Content-Based Image Retrieval (CBIR) is the mainstay of current image retrieval systems is to present an image with set of low level features such as colour, texture, and shape. The most common method for comparing two images in content-based image retrieval is using an image distance measure. Distance measure compares similarity of images in various dimensions such as color, texture, shape. For example a distance of 0 signifies an exact match with the query, with respect to the dimensions that were considered, a value greater than 0 indicates various degrees of similarities between the images. Search results then can be sorted based on their distance to the queried image. Many measures of image distance have been developed.[3]Computing distance measures based on color similarity is achieved by computing a color histogram for each image that identifies the proportion of pixels within an image holding specific values. Examining images based on the colours they contain is one of the most widely used techniques because it can be completed without regard to image size or orientation.[3] Texture measures look for visual patterns in images and how they are spatially defined. Textures are represented by Texel's which are then placed into a number of sets, depending on how many textures are detected in the image.[2] These sets not only define the texture, but also where in the image the texture is located. Texture is a difficult concept to represent. The identification of specific textures in an image is achieved primarily by modelling texture as a two-dimensional gray level variation. The relative brightness of pairs of pixels is computed such that degree of contrast, regularity, coarseness and directionality may be estimated.[2] Shape does not refer to the shape of an image but to the shape of a particular region that is being sought out. Shapes will often be determined first applying segmentation or edge detection to an image. [2]Other methods use shape filters to identify given shapes of an image. These conventional approaches for image retrieval are based on the computation of the similarity between the user's query and images via a query by example (QBE) [1] system. Despite the power of the search strategies, it is very difficult to optimize the retrieval quality of CBIR within only one query process. To solve such problems, in the QBE [1] system, the users can pick up some preferred images to refine the image explorations iteratively. The feedback procedure, called Relevance Feedback (RF), repeats until the user is satisfied with the retrieval results. Although a number of RF studies is made on CBIR, and some common problems, namely redundant browsing and exploration convergence.[3] Relevance feedback refers to a set of approaches learning from an assortment of users' browsing behaviours on image retrieval. [4]

IV. PROPOSED SYSTEM

Propose a novel method named Navigation- Pattern-based Relevance Feedback (NPRF) to achieve the high retrieval quality of CBIR with RF by using the discovered navigation patterns. The navigation patterns mined from the user query log according to the discovered patterns and obtain a set of relevant images in an online query refinement process. Thus, the problem of redundant browsing and exploration convergence is successfully solved. In short, navigation pattern in NPRF Search can be regarded as an optimized search path to converge the search space toward the user's intention effectively. Navigation-Pattern-based Relevance Feedback (NPRF) is an efficient method to achieve high efficiency and effectiveness of CBIR in coping with the large-scale image data. In terms of efficiency, the iterations of feedback are reduced by using the navigation patterns discovered from the user query log .NPRF makes use of the discovered navigation patterns to converge the search space toward the user's intention effectively. High quality of image retrieval on RF can be achieved in a small number of feedbacks. For content-based image retrieval, user interaction with the retrieval system is crucial since flexible formation and modification of queries can only be obtained by involving the user in the retrieval procedure.

User interfaces in image retrieval systems typically consist of a query formulation part and a result presentation part.[4] Specifying what kind of images a user wishes to retrieve from the database can be done in many ways. Category browsing is to browse through the database according to the category of the image. For this purpose, images in the database are classified into different categories according to their semantic or visual content . Query by concept is to retrieve images according to the conceptual description associated with each image in the database and provide an example image from which images with similar visual features will be extracted from the database. Queries may be formed by drawing several objects with certain properties like color, texture, shape, sizes and locations. In most cases, a coarse sketch is sufficient, as the query can be refined based on retrieval results. Query by example allows the user to formulate the query a query by providing an example image.[1] The system converts the example image into an internal representation of features. Images stored in the database with similar features are then searched, if the query image is not in the database, and query by internal image example, if otherwise. For query by internal image, all relationships between images can be pre-computed. The main advantage of query by example is that the user is not required to provide an explicit description of the target, which is instead computed by the system. It is suitable for applications where the target is an image of the same object or set of objects under different viewing conditions. Most of the current systems provide this form of querying. Query by group example allows user to select multiple images. The system will then find the images that best match the common characteristics of the group of examples. In this way, a target can be defined more precisely by specifying the relevant feature variations and removing irrelevant variations in the query. In addition, group properties can be refined by adding negative examples.[5] Many recently developed systems provide both query by positive and negative examples. Two types of operation used in proposed approach is i) Online image retrieval and ii) Offline knowledge discovery

i) Online Image Retrieval

1) Initial Query Processing Phase: Without considering the feature weight, this phase extracts the visual features from the original query image to find the similar images. Afterward, the good examples picked up by the user are further analyzed at the first feedback.

2) Image Search Phase: In this phase, a new query point at each feedback is generated by the preceding positive examples. Then, the k-nearest images to the new query point can be found by expanding the weighted query[3]. The search procedure does not stop unless the user is satisfied with the retrieval results.

ii) Offline Knowledge Discovery

1) Knowledge Discovery Phase: Learning from users behaviours in image retrieval can b viewed as one type of knowledge discovery. Consequently, this phase primarily concerns the construction of the navigation model by discovering the implicit navigation patterns from users' browsing behaviours. This navigation model can provide image search with a good support to predict optimal image browsing paths.

2) Data Storage Phase: The databases in this phase can be regarded as the knowledge marts of a knowledge warehouse, which store integrated, time-variant, and non volatile collection of useful data including images, navigation patterns, log files, and image features. The knowledge warehouse is very helpful to improve the quality of image retrieval the procedure of constructing rule base from the image databases can be conducted periodically to maintain the validity of the proposed approach. A unified log-based relevance feedback provides framework for integrating log data of user feedback with regular relevance feedback for image retrieval. NPRF works by integrating the navigation pattern mining and a navigation-pattern-based search approach. It is found that none of the existing approaches meets completely the requirements of an accurate CBIR system with relevance feedback because none of the techniques have completely solved the problem of semantic gap.[6] So it is still undecided what the future truly holds for Improving and implementing Relevance Feedback in real world applications as shown in figure1below:

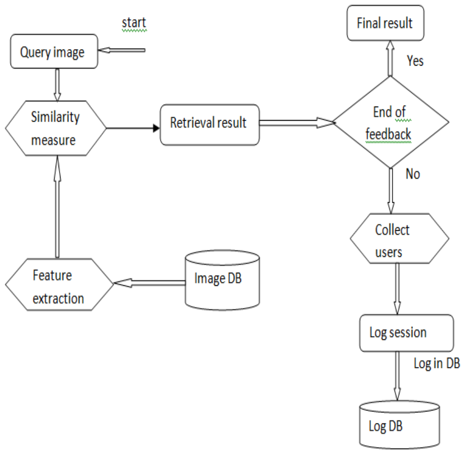

Figure 1: Architecture of NPRF

V. NPRF SEARCH ALGORITHM

NPRF Search algorithm proposed iterative solution to RF, which merges. NPRF Search is proposed to reach the high precision of image retrieval in a shorter query process by using the valuable navigation patterns.

The iterative procedure can be shown accordingly

1. Generate a new query point by averaging the visual-features of positive examples.

2. Find the matching navigation patterns by calculating nearest query seeds (root).

3. Calculate the nearest leaf node from the matching navigation patterns tree.

4. Find the top s relevant visual query points from the set of the nearest leaf node nodes.

5. Finally the top k images are returned to the user.

NPRF Search algorithm is triggered by receiving:

1) A set of positive examples P and negative examples N determined by the user at the preceding feedback,

2) Generate a new query point *Newqp* by G and compute the new feature weights

3) An accuracy threshold thrd.

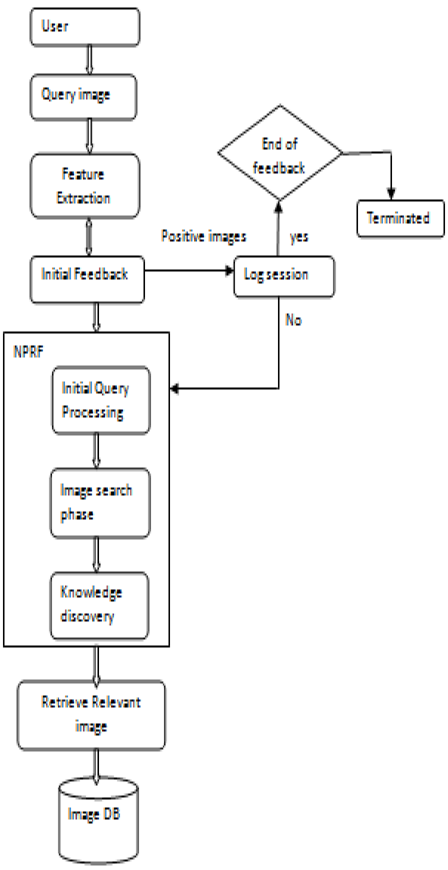

Figure 2: Block Diagram of NPRF search

NPRF Search Algorithm

1. Generate a new query point Newqp by G and compute the new feature weights;

2. Let NM be the accumulated set of negative examples and Nm= NM U N;

3. Store Newqp and G into the log database

4. Initialize each trh,rth,chk=0 and canPnt=ø;

5. for each gi do

6. Determine the special query seed rth with the shortest distance to g

7. rth.chk=1;

8. end for

9. if |G| < thrd then 10. for each n

11. Determine the special seed rth with the shortest distance to nn,

12. count (rth) ++

13. End for

14. Find the seed rth with max (count (rth));

15. rth.chk=0;

16. End if

17. for each trh do

18. if trh.rth.chk=1 then

19. Find the set of the visual query QPT within the leaf nodes of pattern trh;

21. End if

22. End for

23. Find the top s visual query point

24. for i=0 to s do

25. Find the positive image set PM in the transformed log table, which is referred to as sqpt;

26. CanImg=CanImg U RM

27. End for

28. CanImg=| CanImg\ NM|;

29. Rank the image in CanImg;

30. Return the set of top K similar images R;

Advantages:

1.NPRF outperforms other existing methods significantly in terms of precision, coverage, and number of feedbacks.

2. This method provides optimal solution to resolve the problems existing in current RF, such as redundant browsing and exploration convergence.

3. The navigation patterns can assist the users in obtaining the Global optimal results.

Issues:

1. User's profile needed to be integrated into NPRF to further increase the retrieval quality which is not yet integrated.

VI. CONCLUSION

Deal with the long iteration problem of CBIR with RF, we have presented an approach named NPRF by integrating navigation pattern and a navigation- pattern-based search approach named NP Search. The main feature of NPRF is to efficiently optimize the retrieval quality of interactive CBIR the navigation patterns derived from the users' long-term browsing behaviors are used as a support for minimizing the number of user feedbacks. The proposed algorithm NPRF Search performs the navigation- pattern-based search to match the user's intention by merging three query refinement strategies to solve the problems such as visual diversity and exploration convergence. The methods for special data partition and pattern pruning also speed up the image exploration. The experimental results reveal that the proposed approach NPRF is very effective in terms of precision and coverage. Within a very short term of relevance feedback, the navigation patterns can assist the users in obtaining the global optimal results. Moreover, the new search algorithm NPRF Search can bring out more accurate results. In the future, there are some remaining issues to investigate is to integrate user's profile into NPRF to further increase the retrieval quality and then apply the NPRF approach to more kinds of applications on multimedia retrieval or multimedia recommendation.

Acknowledgement

I express my sincere gratitude to my guide Ms Subi S, Assistant Professor of Computer Science and Engineering Department, and I express my sincere gratitude to Mr.Pratap.G.Nair, Dean of Computer Science & Information Technology Departments, KMCT College of Engineering for their valuable guidance. Above all, I thank the almighty for enabling me to be what I am.

References

1. M.D. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Steele, and P. Yanker,"Query by Image and Video Content: The QBIC System," Computer, vol. 28, no. 9, pp. 23-32, Sept. 1995.
2. D. Harman, "Relevance Feedback Revisited," Proc. 15th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval, pp. 1-10, 1992.
3. Y. Rui, T. Huang, M. Ortega, and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," IEEE Trans. Circuits and Systems for Video Technology, vol. 8, no. 5, pp. 644-655, Sept. 1998.
4. X. Jin and J.C. French, "Improving Image Retrieval Effectiveness via Multiple Queries," Multimedia Tools and Applications, vol. 26, pp. 221-245, June 2005.
5. 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.
6. X.S. Zhou and T.S. Huang, "Relevance Feedback for Image Retrieval: A Comprehensive Review," Multimedia Systems, vol. 8,no. 6, pp. 536-544, Apr. 2003.

AUTHOR(S) PROFILE

Anju C R received her B tech degree in Information Technology from SNS College of Technology Coimbatore, Anna University, Chennai, India and is presently doing her final year Master of Technology from KMCT College of Engineering Calicut, Calicut University, Kerala, India. Her research interest includes Data Mining and Knowledge Engineering.

Subi S received her ME degree in Computer Science and Engineering from CIET Coimbatore , India and working as Assistant Professor in Computer Science and Engineering Department, KMCT College of Engineering, Calicut, India

