A Proposed Framework for a Distributed CBIR System based on Salient Regions and RF Techniques

Ali. I. ELdesouky
Computer and Systems Department
Faculty of Engineering
Mansoura University
Egypt

Hesham A. Ali
Computer and Systems Department
Faculty of Engineering
Mansoura University
Egypt

Noha A. Sakr
Computer and Systems Department
Faculty of Engineering
Mansoura University
Egypt

ABSTRACT

Digital images databases open the way for content-based searching. Content Based Image Retrieval occupies a well ranked position among the research areas as it provides the practical solution for narrowing the semantic gap between the image retrieval process and the human perception. The main objective of this paper is to propose a framework for region content based image retrieval based on a distributed clustered image dataset. The proposed framework introduces a new perspective to measure the similarity between the image query and the clustered dataset images. Moreover, a development by adopting three relevance feedback techniques is used to refine the results of the retrieval system which are the well known Query Point Movement and Query Expansion, besides to the proposed third technique which is Query Modified Re-Weighting technique.

General Terms

Image Retrieval, Relevance Feedback.

Keywords

CBIR, Saliency Regions, Clustering.

1. INTRODUCTION

Large collections of digital images are being created in commerce, government, academia, hospitals, and many other areas along with the increasing in the popularity of the web blogging and the photo sharing. Usually, the way to search these collections was by keyword indexing, or simply by manually browsing through the database by the humans, and surely this process was very time and effort consuming.

Image Retrieval (IR) is a newly adopted methodology aims to extract the user-query relevant images from an image database. Revolutionary digital image databases growth always brings challenges to the IR systems. IR approaches varied between the Annotation Based Image Retrieval (ABIR) or the meta-data approach which is a traditional method for image retrieval that makes use of the meta data of image such as the textual descriptions, captioning, or the keywords to search and retrieve images [1]. ABIR can be either classification based methods or probabilistic modeling based method [2]. However there is also the Content Based Image Retrieval approach in which database images are indexed on the basis of low level features such as color, texture and shape which can be automatically derived from the visual contents of the images

Although many researchers have investigated this topic from many perspectives to develop its performance, Content Based Image Retrieval (CBIR) systems still suffers from a lot of shortcomings. Extracting the semantically related images from large scale images datasets presents a permanent challenge in the recently discussed researches.

Unfortunately CBIR systems have not achieved the commercial success yet, as large commercial image providers still using human indexers to annotate the images with the appropriate keywords. This paper is an attempt to present a framework for CBIR based on measuring the similarity between the query and the distributed image clusters of the database on the basis of detecting the saliency regions of the images. Moreover it retains the user feedback by investigating his relevant results by adapting two techniques for Relevance Feedback, and proposing the third to refine the retrieved results by the CBIR system.

This paper is organized as follows; an overview about the CBIR is discussed in section II. The Concept detection and the problem definition for the semantic gap are presented in section III. In section IV, a brief introduction to the Region Content Based Image Retrieval is discussed. An introduction to the relevance feedback techniques is presented in section V. The related literature is briefly discussed in section VI. A presentation for the proposed framework is introduced in section VII. Finally the conclusion is overviewed in section VIII.

2. CONTENT BASED IMAGE RETRIEVAL (CBIR)

CBIR as a term was evolved since 1990 which is used to describe the process of retrieving desired images from a large collection of images. CBIR uses the visual contents of an image such as color, shape, texture, and spatial layout to represent and index the image [2]. In typical CBIR system, the visual contents of the images in the database are extracted and described by multi-dimensional feature vectors [3]. The operation of a CBIR system can be seen as a series of independent processing stages as shown in figure 1. In general, the development of a CBIR system consists of at least four stages: data acquisition and processing, feature representation, data indexing, query processing, and feedback processing.

Image query regarding the CBIR applications mainly organized into three paradigms [4]. First paradigm is the Query By Example (QBE) in which the query image is an example shown either from the database itself or some external location [2, 4]. Secondly, there is the image query

based on a given type identified by the derived features. Thirdly, there is the query image provided by the abstract attributes such as some named events or emotional significance [4].
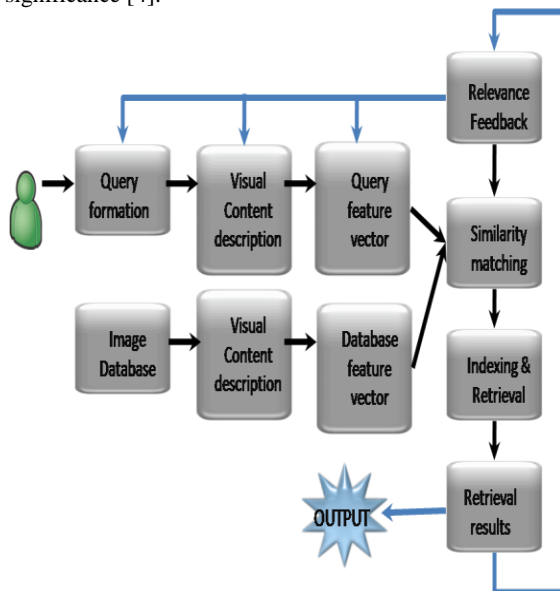

Fig 1: A typical CBIR system

3. CONCEPT DETECTION AND SEMANTIC GAP

CBIR technologies have shown a lot of limitations regarding lack of the support of high level semantic knowledge and the fact of being far away from the human query perception. Although the user seeks the semantic similarity, the database can only provide the mathematical similarity by means of data processing. While image retrieval has been active over the years, an emerging new and possibly more challenging field is arising which is automatic concept recognition from the visual features of image. Yet, there is what is called the semantic gap which may be defined as the lack of coincidence between the information that can be extracted from the visual data and the interpretation that the same data have for a user in a given situation [4]. Shortly, it can be defined as the gap between the human vision and the results of the CBIR systems. Many solutions were proposed to reduce the semantic gap such as [5]:

1. Incorporating the query concepts with the low level features by using the machine query learning tools.
2. Using objects ontology to define high level concepts.
3. Generating semantic templates to support high level Information Retrieval.
4. Introducing Relevance Feedback (RF) into retrieval process for continuous learning of user intention.
5. Making use of visual contents and textual information.

Concisely, to narrow down this semantic gap, two main proposed techniques have been widely used which are: investigating the region-based features to represent the focus of the user's perceptions of the image content. Also, adapting learning techniques, such as Relevance Feedback techniques to learn the user's intentions [6]. Though, many limitations still exist in the current approaches, which make these approaches to suffer from lack of accuracy.

4. REGION CONTENT BASED IMAGE RETRIEVAL

User always seeks his own interested objects or regions in the query image. This matter may lead onto considering the local features as more profitable descriptors other than the global ones to capture the human conception. Region Based Image Retrieval was recently proposed as an extension to CBIR. RBIR is strongly believed to retain the semantic meaning of each region when it segments the image into reasonable regions. Initially the image got to be segmented into specific regions by detecting the regions boundaries. The detected regions are being represented by its local feature and the weights of its importance. Finally and based on the regions representation, the similarity between the query regions and the other images are calculated to determine the relevant and irrelevant targets [2]. Though, RBIR still lacks the quality for many reasons such as inaccurate region segmentation, high dimensionality of the extracted local features and determining which similarity measure to apply.

5. RELEVANCE FEEDBACK

One of the interactive learning techniques is relevance feedback (RF), which was initially introduced into CBIR during mid 1990's [7]. The main idea behind RF concept is to allow the user to participate in refining the CBIR results by judging its given results. During retrieval process, the user interacts with the system and rates the relevance of the retrieval images, according to his subjective judgment. This judgment may be just to rank it as relevant or irrelevant or he may use some kind of granularities.

Using this information, the system dynamically learns the user's intention, and gradually presents better results [1].
RF techniques should consider a lot of design aspects to enhance its online functionality such as [8]:

1. Few numbers of iterations not to cause tire for the user.
2. Reasonable amount of feedbacks after each iteration when the system gives his response for this iteration.
3. Feature extraction process may not consume much time to avoid user frustration.

6. RELATED WORK

Many contributions to the different areas of CBIR were investigated in the literature. In this context a brief overview about the most well known proposals is presented.

There is *Marques et al. who* segmented the images out to regions that match salient Regions of Interest (ROI) defined by studying human perception [9]. And *Wang et al.* who constructed their ROI on the basis of the wavelet decomposition of an image and then, they used the lowest frequency sub-band as an approximation of the original image [10]. Also there is *Xiong et al.* who proposed a framework which applied multiple features descriptors to represent local features in their Multi-PRE [11]. And *Fan et al.* who proposed a framework to achieve a middle level understanding of contents of images. They developed their method to find salient objects based on some pre-defined basic vocabularies [12].
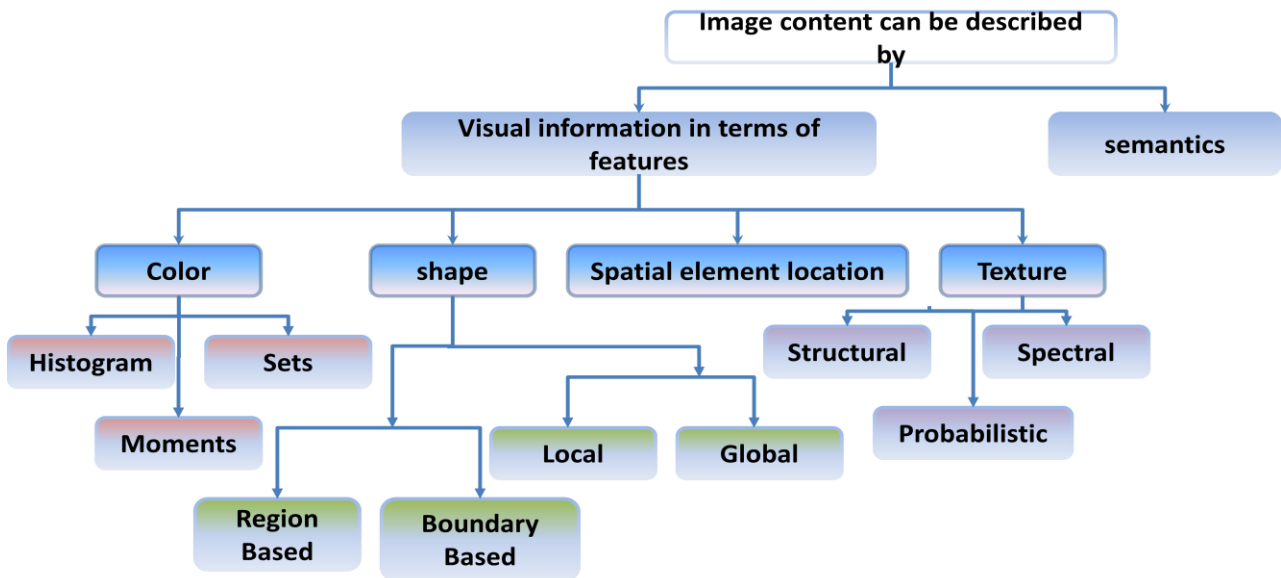
Fig 2: Image content description features varieties

And Chen and Wang who proposed to segment the image into several regions and represent these regions by multi-dimensional fuzzy sets as their local feature representation [13].*Carson et al.* who proposed the Blobworld system which utilized four kinds of local features; color, texture, location and shape to represent blobs (regions) selected by user from the query [14].

7. THE PROPOSED FRAEWORK STRUCTURE

Developing a solution for the high dimensional image content indexing is a crucial challenge when it comes to building a real world large-scale CBIR system. This paper proposes an integrated framework for a CBIR system based on considering three issues to enhance the performance of the system. Initially, it captures the semantic of the images based on its regions content, thus by following the RBIR as a valuable extension to the traditional CBIR. Furthermore, it distributes the image dataset into clusters. Each cluster has the closest image contents according to a similarity measure. This similarity measure is computed based on the saliency regions of the images contents. Moreover, introducing a proposed similarity measure which computes the distance between the given query and the dataset clusters. Finally, it applies three well-known RF techniques to get the best use of the user judgment in guiding the proceeding of the system.

As shown in figure 3, the proposed framework mainly consists of three layers which are: Dataset processing layer, Query processing layer and the Relevance Feedback layer. Each layer contains different modules based on its prior functionality. Here in the upcoming section; a brief overview about each of these modules is being introduced.

7.1 Dataset Processing Layer:

1) Data Processing and Acquisition Module

In this module, each dataset image is being segmented into regions, and then the duplicated regions are removed. Then a log for set of objects is refined by the corresponding object sets detected from the image.

2) Salient Region Detecting Module

In this module, the salient regions for each dataset image are detected. Afterwards, the salient regions are separated from the background. Then, the detected saliency regions are described by Scale Invariant Feature Descriptor (SIFT) descriptor because of its highly invariance to scale, illumination and orientation.

3) Dataset Clustering Module

In this module, the dataset is being clustered by applying the K-means partitioning clustering algorithm. Then, the distributor component does the distributing task by allocating each cluster or set of clusters on separate machines based on the decision taken by the decision maker component to determine the threshold value between these clusters.

7.2 Query Processing Layer:

1) Query Formation Module

In this module, the feature vector of the image query is constituted by using the SIFT descriptor. Later, this image description is utilized in the similarity matching process.

2) Query-Cluster Assigning Module

In this module and based on the obtained SIFT descriptors for both the query and the dataset image clusters, the distance between the query and each cluster is obtained. On the basis of this distance, the system decides to exclude the longest distance clusters from the processing of the Query-Cluster assigning component.

7.3 Relevance Feedback Layer:

1) Query Point Movement (QPM) Module

In this module, the obtained results of the system are refined by using one of the well known RF techniques which is QPM. QPM considers the positive examples as an initial query point at each feedback. After much iteration the contour of the query point becomes close to the region of the user interest [1].

1) Query Modified Re-Weighting (QMR) Module

In this module, the proposed technique for updating the traditional Re-Weighting RF technique is presented.

8. CONCLUSION

In this paper, a proposed framework for an integrated CBIR system is introduced. The proposed framework addresses the most common problems in CBIR systems. It counts on considering the RBIR instead of the traditional CBIR approaches. Furthermore, it considers the distributing methodology to work on reducing the consumed time for such job by cutting the image dataset into well-defined related clusters.
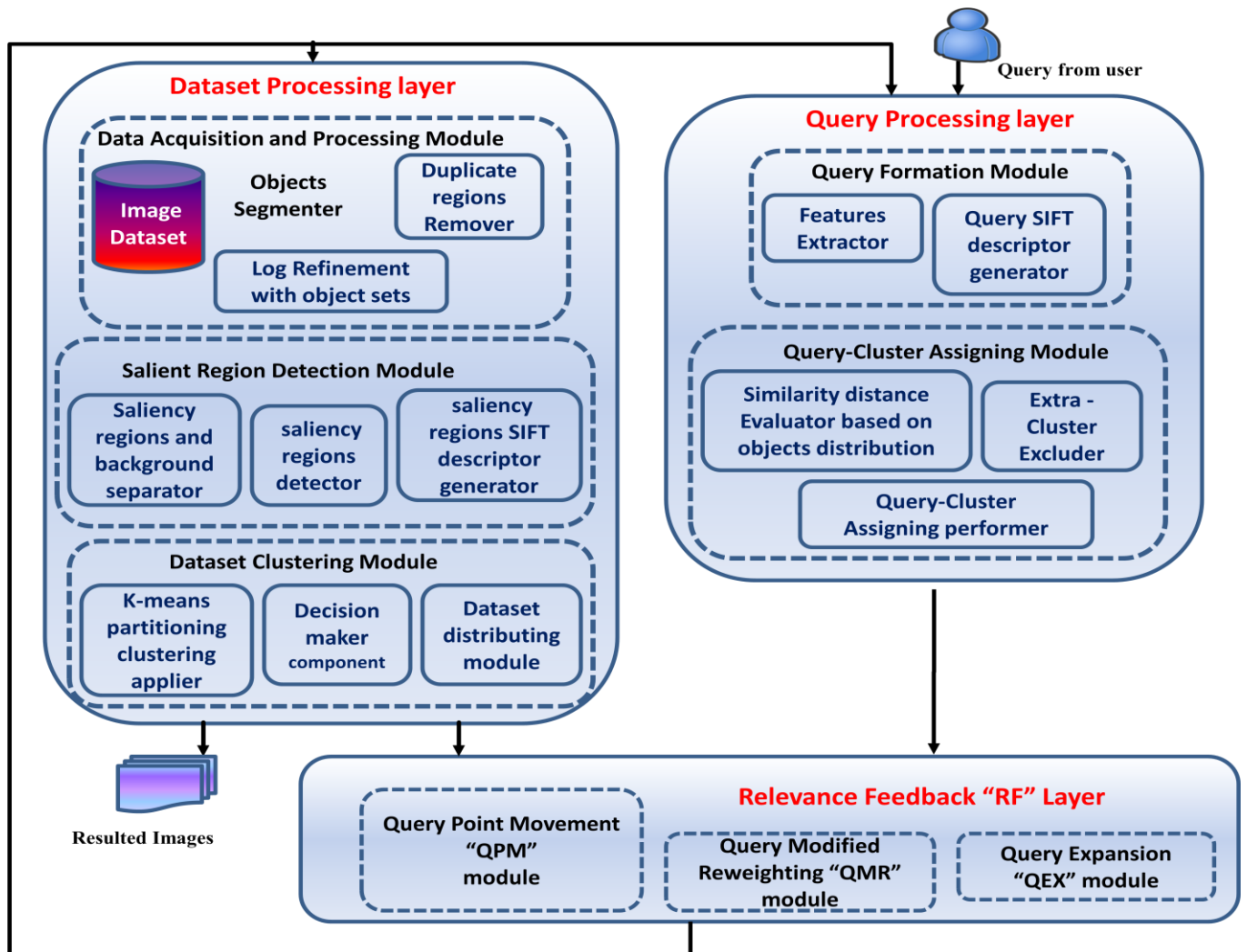

Fig. 3: The proposed CBIR framework components

It introduces a proposed algorithm to assign weighted ranked degrees to the features which happen to exist in the positive examples.

2) Query Expansion (QEX) Module

QEX is a hot technique related to the area of RF in which the query is modified based on some global resources such as a resource that is not query dependent [15].

Moreover, it presents a technique to measure the distance between the image query and the dataset clusters. Eventually, it considers RF concepts by applying two well-known RF techniques which are QPM and QEX besides proposing the modified QMR technique.

9. REFERENCES

[1] Su, J., Huang, W., Yu, P. and Tseng, V., 2011. Efficient Relevance Feedback for Content-Based Image Retrieval by Mining User Navigation Patterns, IEEE Trans. On

Knowledge and Data Engineering, Vol.23, No.3, pp. 360-372.

[2] Yang, J. and Zhu, S., 2012. Narrowing Semantic Gap in Content-based Image Retrieval, International Conference on Computer Distributed Control and Intelligent Environmental Monitoring, pp.433-438.

[3] Yang, J. and Zhu, S., 2012. Narrowing Semantic Gap in Content-based Image Retrieval, International Conference on Computer Distributed Control and Intelligent Environmental Monitoring, pp.433-438.

[4] Deekshatulu, B. L., 2010. Learning Semantics in Content Based Image Retrieval (CBIR) – A brief review. 2010 Second Vaagdevi International Conference on Information Technology for Real World Problems, pp. 76-78

[5] Setia L., 2008. Machine Learning Strategies for Content Based Image Retrieval. M. Sc. Dissertation, Albert-Ludwigs-Universität, Freiburg im Breisgau.

[6] Yang, Z. and Kuo C-C. J., 1999. Semantic Classification and Composite Indexing Approach to Robust Image Retrieval", Image Processing, ICIP 99, vol. 1, pp. 134 – 138.

[7] Salton, G. 1989. Automatic text processing, Addison-Wesley.

[8] MacArthur, S. D., E.Brodley, C., and Kak, A., 2002. Interactive Content-Based Image Retrieval Using Relevance Feedback, Computer Vision and Image Understanding 88,pp. 55–75.

[9] Marques O., Mayron L., Borba G. and Gamba, H. 2006. Using visual attention to extract regions of interest in the context of image retrieval. In Proceedings of ACM annual southeast regional conference, pp. 638-643.

[10] Wang, X.-Y., Hu, F.-L. and Yang, H.-Y. , 2006. A novel regions-of-interest based image retrieval using multiple features. In Proceedings of the multi-media modeling international conference, pp. 377-380.

[11] Xiong, W., Qiu, B., Tian, Q., Xu C.-S., Ong S., Foong, K. et al.,2005. MultiPRE: A novel framework with multiple parallel retrieval engines for content-based image retrieval. In Proceedings of the ACM international conference on multimedia, pp. 1023-1032.

[12] Fan, J.-P., Gao, Y.-L., Luo, H.-Z., and Xu, G.-Y., 2004. Automatic image annotation by using concept-sensitive salient objects for image content representation. In Proceedings of the ACM SIGIR conference on research and development in information retrieval, pp. 361-368.

[13] Chen, Y., Wang, J., 2002. A region-based fuzzy feature matching approach to content-based image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24(9): pp. 1252-1267.

[14] Carson, C., Thomas, C., Belongie S., Hellerstein J., and Malik J., 1999. Blobworld: A system for region-based image indexing and retrieval". Lecture Notes in Computer Science, pp. 1614-1660.

[15] Porkaew, K.., Chakrabarti, K., and Mehrotra, S., 1999. Query Refinement For Multimedia Similarity Retrieval in MARS," Proc. ACM Int'l Multimedia Conf. (ACMMM), pp.235-238.