World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:8, No:7, 2014

Performance Evaluation of Content Based Image Retrieval Using Indexed Views

Tahir Iqbal, Mumtaz Ali, Syed Wajahat Kareem, Muhammad Harris

***Abstract*—**Digital information is expanding in exponential order in our life. Information that is residing online and offline are stored in huge repositories relating to every aspect of our lives. Getting the required information is a task of retrieval systems. Content based image retrieval (CBIR) is a retrieval system that retrieves the required information from repositories on the basis of the contents of the image. Time is a critical factor in retrieval system and using indexed views with CBIR system improves the time efficiency of retrieved results.

***Keywords*—**Content based image retrieval (CBIR), Indexed view, Color, Image retrieval, Cross correlation.

Open Science Index, Computer and Information Engineering Vol:8, No:7, 2014 waset.org/Publication/9998805

I. Introduction

CONTENT based image retrieval system actually retrieves images from database upon the query image. Fig. 1 shows a typical CBIR system. Terminology CBIR was first coined by Kato in 1992 [1]. Kato has used color and shape features as features vectors for the retrieval of images from databases. Since then the term CBIR is being extensively used in the fields like Medical informatics, surveillance digital rights management [2]. A system used for storing and organizing data in a systematic way is called a database.

A database management system (DBMS) is a collection of software's that helps a user to edit, store and retrieve the desired data from database. Text, image, audio and video are multimedia data and a Multimedia Database Management System is a frame work that manages multimedia data.

The continuous work in the CBIR system has made the field mature but, difficulties still reside in locating an image in large and huge collection of images as this problem does not exist in small repository of images [3]. Retrieval time is also an issue in large database [4] i.e. Google, yahoo, Bing which contain huge repository of images and billions of users.

View defines functionality from a couple of base tables into an extracted table [5]. A view can be made based on a single table with self-join. This functionality is commonly recomputed every time the particular view is referenced, indexed views are the tuples that are stored physically in the database, indexed views can be indexed and access to the indexed can much faster as it works just like a cache [5].

Indexed view eradicates the extra execution connected with joins along with aggregations for hugeas well as important category of queries. Queries to huge data source frequently contain joins among tables, aggregations including average, sum, count or both aggregation & joins. An indexed view provides substantial changes with query processing time, in particular with regard to aggregation queries over huge tables of multimedia data [6].

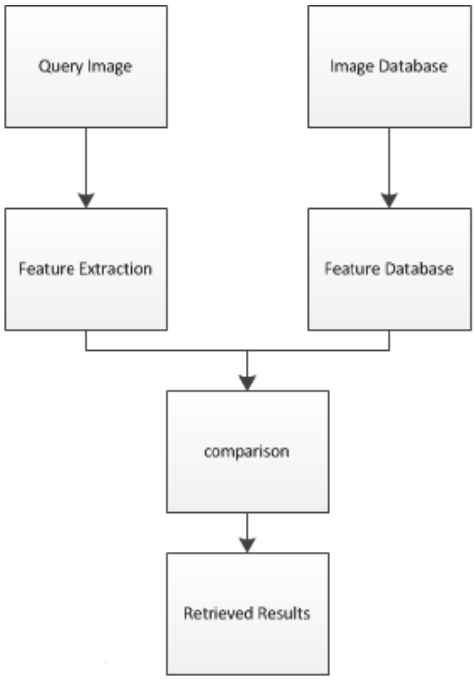

Fig. 1 Mapping Nonlinear Data to a Higher Dimensional Feature Space

II. Background

A. Content Based Image Retrieval (CBIR)

Content based image retrieval is a system that outputs the results matching the image query from multimedia database [7]. In Content Based Image Retrieval, a separate database is maintained containing the feature vectors called Feature database .There are various low level features like color,

Tahir Iqbal is MS Scholar at City University of Science and Information Technology Peshawar, Pakistan. He is currently working under Education Department Government of Pakistan in Independent Monitoring Unit Project. (Phone: +92-3459036653; e-mail: tahir.iqbal921@gmail.com).

Mumtaz Ali is PHD Scholar at COMSATS Institute of Science and information Technology Islamabad, Pakistan. He is currently working at City University of Science and Information Technology as a Lecturer in Department of Computer Science (e-mail: mumtazalikhan@gmail.com).

Syed Wajahat Kareem is MS Scholar at City University of Science and Information Technology Peshawar, Pakistan. He is currently working under Department of Education as Assistant Programmer in EMIS Section. (e-mail: syedwajahat.kareem@yahoo.com).

Muhammad Harris is MS Scholar at City University of Science and Information Technology Peshawar, Pakistan. He is currently working at Home & Tribal Affairs Department as Network Administrator. (e-mail: Muhammadharris1@gmail.com)

Open Science Index, Computer and Information Engineering Vol:8, No:7, 2014 waset.org/Publication/9998805

shape, texture, locality on the basis of which the matching process is performed and the results are output in response to any query. In recent decade some commercially launched products are QBIC [8], Photobook [9], Virage [10], Visual Seek [11], Netra [12] and Simplicity [13]. Inside the aforementioned methods, Retrieval time is concerned with size of the database. Within these systems, query image is comprehensively compared with all the images throughout the database. This strategy does not fit to huge databases due to time consumption. For efficiency purpose only a section of a database should be matched to the query. Clustering techniques are used which for this reason, which groups same images beyond the search action takes place [14].

B. Clustering Techniques in Image Retrieval System

Clustering approaches are labeled into supervised (Semi-supervised included) and unsupervised schemes.

Fuzzy and c-mean are density based clustering which lie under the category of unsupervised clustering. Splitting criteria for clustering needs human intervention i.e. having a hierarchical approach and is under the umbrella of supervised clustering [4].

1. Supervised Clustering

Keywords-based image retrieval systems discover correspondence through coordinating keywords that are input to the retrieval system to the image tags inside databases [2]. Sometimes the images in the database are no properly tagged due to which it is easy to deceive the system e.g. an image of a car tagged with name of bird or a bird's image tagged with the name of a musical instrument may easily result in wrong output of a retrieval system [4].

2. Unsupervised Clustering

To enhance the accuracy of image retrieval system, dependencies should be reduced; a simple yet effective system should be applied where the contents of the image regardless of the keywords for retrieval [4]. In [15] an effective conceptual system is proposed where the resultant images are presented as results on the basis of geometrical shapes, color, texture, and co-related objects.

III. Proposed Approach

Retrieval of multimedia data in a huge database is a very time consuming process. To reduce the time, indexed Views (Materialized Views) are used inside database. Proposed model consist of 3 Indexed views which are created on the basis of basic color i.e. RGB. Each view contains the images from the main table based on their maximum color percentage. Figs. 2 and 3 show the practical implementation of indexed views inside the SQL Server Database and Internal Structure of indexed view is shown below.

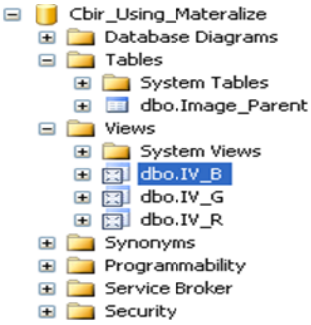

Fig. 2 Internal Structure of Indexed View

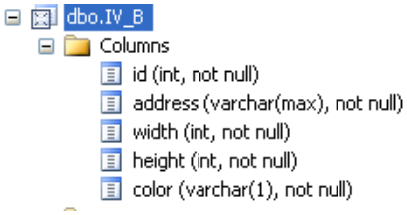

Fig. 3 Internal Structure of Indexed View

Current working model (Fig. 4) takes the image from the user. The image is first parsed and checked according to the given color group of indexed view i.e. Red, Green or Blue (Named as IV_R, IV_B and IV_G). Indexed view according to color is then selected and queried for the images inside it. Query image and fetched images from Indexed View is then passed from the comparison function to find the relevant image. The relevant images are then ordered according to their matching value. Most near images based on threshold value are given to GUI to display the result.

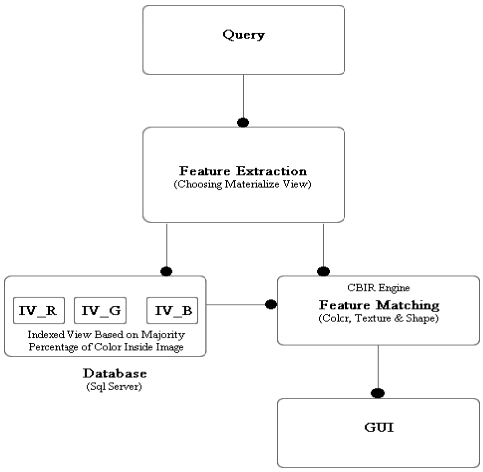

Fig. 4 Proposed Model

IV. Results

Proposed system performed better than traditional database in practical domain where there is huge multimedia database. Current system have repository of 265 images with different sizes (Large, Medium and Small). Current System utilizes cross correlation for image comparison. Above results show the overall improvement in image fetching time in large

database. As shown below in Figs. 5 and 6.

A. Practical Implementation of Indexed Views

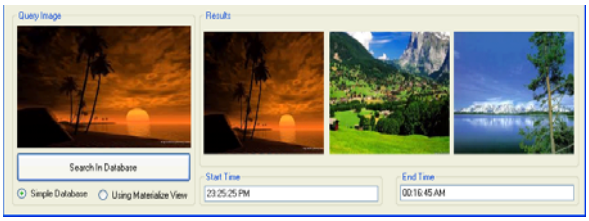

Fig. 5 Retrieval without Using Indexed Views

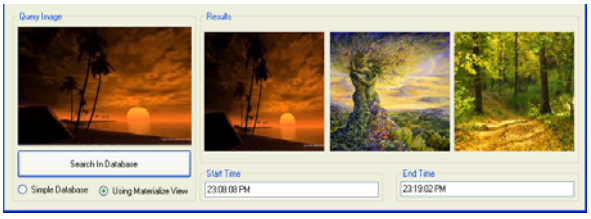

Fig. 6 Retrieval Using Indexed Views

B. Time Complexity Pseudo Code

Time complexity pseudo code of classic approach and materialized views are given below.

_____Classic Approach _____

Procedure: Classic Approach (Image)
Define threshold value
for each image in database
compare image with Image
if comparision value is equal or
greater then threshold
update comparison value
save image in temporary image
end if
end for loop
end
Display Image from Temporary Image

_____CBIR Using Materliazed View_____

Procedure: CBIR using matview (Image)
Define threshold value
Define RGB Groups for Image
Check image RGB Value
If R Then Select IndexviewR
else if G Then Select IndexviewG
else B Then Select IndexviewB
for each image in selected IndexView
compare image with Image (correlation)
if comparision value is equal or greater then
threshold
update comparision value
save image in temporary image
end if
end for loop
end
Display Image from Temporary Image

C. Time Complexity Analysis

Time complexity analyses of both approaches are given below:
For Classic Approach = N Times
For CBIR Using Materialize View = N, N/3, 1 (N/3 is average case)

TABLE I
PROCESSING TIME COMPARISON OF CLASSIC VS. INDEXED VIEW APPROACH (PER UNIT TIME)

Number of Images	Classic Approach (Unit of Time)	Using Materialize View (Unit of Time)
100	100	33.3
500	500	166.7
1000	1000	333.3
2000	2000	666.7
4000	4000	1333.3
5000	5000	1666.7

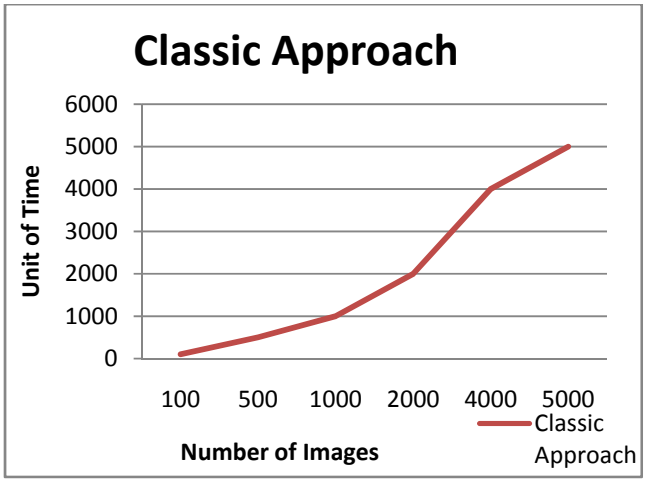

Fig. 7 Time Complexity Analysis for Classic Approach

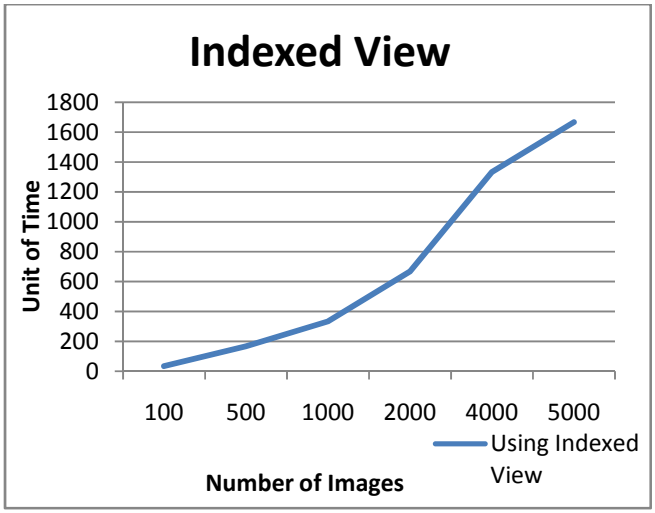

Fig. 8 Time Complexity Analysis Using Indexed View

Open Science Index, Computer and Information Engineering Vol:8, No:7, 2014 waset.org/Publication/9998805

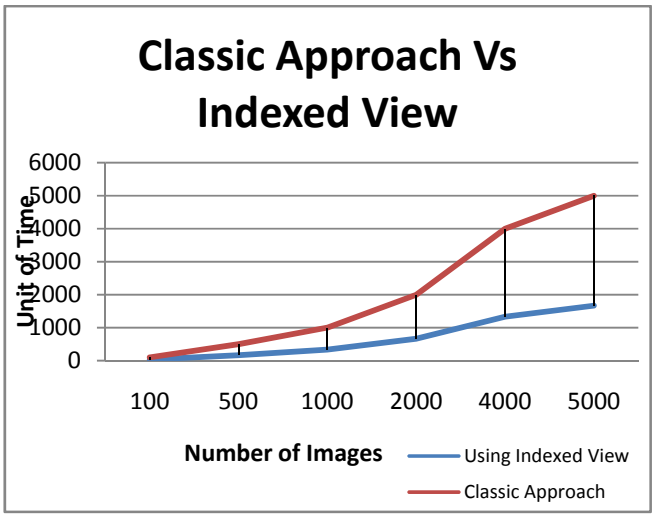

Fig. 9 Time Complexity Analysis for Classic Approach vs. Indexed View

V. Conclusion

As it is shown that proposed model works much faster than traditional multimedia database. It divides the multimedia content based on feature vector (Low level Features (Color)) and stores them in their appropriate indexed views, from where the imagesare fetched on query. This approach is efficient in large databases i.e. Yahoo, Microsoft, Bing etc. which have huge repository of images and many users querying 24/7. It improves the system efficiency by saving system resource cost and utilizing the current system on its maximum.

VI. Future Work

In future work the current indexed view can utilize the other low level features (Texture and Shape) and high level features (Segmentation, Image Position etc.) for finding the better indexed view model which will further increase the current system efficiency and saves a lot resources in an environment of many user and large content.

References

[1] Kato T, Kurita T, Otsu N, Hirata K. A sketch retrieval method for full color image database-query by visual example. In Proc. IntConfon Pattern Recognition Aug-3, 1992; 530-533W.-K. Chen, *Linear Networks and Systems* (Book style).Belmont, CA: Wadsworth, 1993, pp. 123–135.

[2] Huiyu Zhou, Abdul H. Sadka, Muhammad R. Swash, Jawid Aziz and Abubakar S. Umar Content Based Image Retrieval and Clustering: A Brief Survey.

[3] Jain R. Workshop report: NSF workshop on visual information management system. In storage and retrieval for image and video databases. In: Niblack WR, Jain RC, Eds. Proc. SPIE 1908, 1993; 198-218.

[4] Content Based Image Retrieval (CBIR) Using Materialized Views Mumtaz Ali, Awais Adnan, MuhammdSaqib and Zahidullah.

[5] D. Warehousing and J. Notices, "Data Engineering," vol. 18, no. 2, 1996.

[6] Optimizing Queries Using Materialized Views: A Practical, Scalable Solution Jonathan Goldstein and Per-Åke Larson ACM SIGMOD 2001 May 21-24, Santa Barbara, California USA Copyright 2001 ACM 1-58113-332-4/01/05.

[7] John Eakins, Content-based Image Retrieval Margaret Graham University of Northumbria at Newcastle.

[8] Flickner M, Sawhney H, Niblack W, et al. Query by image and video content: The QBIC systComput 1995; 28: 23-32.

[9] Pentland A, Picard R, Scaroff S. Photobook: Content-based manipulation for image databases. Int J Comput Vision 1996; 18: 233-254.

[10] Gupta A, Jain R. Visual information retrieval. Commun ACM 1997; 40: 70-79.

[11] Smith J, Chang S-F. Visualseek: A fully automated content-based image query system. In multimedia'96: Proc. of the Fourth ACM IntConf on Multimedia 1996; 87-98.

[12] Ma W, Manjunath B. Netra: A toolbox for navigating large image databases. In Proc. IntConf on Image Process 1997; 1: 568-571.

[13] Wang JZ, Li J, Wiederhold G. Simplicity: Semantics-sensitive integrated matching for picture libraries. IEEE Trans Pattern Anal Machine Intell 2001; 23: 947-963.

[14] Jin J, Kurniawati R, Xu G, Bai X. Using browsing to improve content based image retrieval .J Visual Commun Image Represent2001;12:123-135.

[15] Awais Adnan, Muhammad Nawaz, Sajid Anwar, Tamleek Ali, Muhammad Ali. Object Identification with Color, Texture, and Object Correlation In CBIR System, World Academy of Science, Engineering and Technology 64 2010

Open Science Index, Computer and Information Engineering Vol:8, No:7, 2014 waset.org/Publication/9998805