FRAME WORK FOR CONTENT BASED IMAGE RETRIEVAL (Textual Based) SYSTEM

Jalil Abbas[1], Salman Qadri[2], Muhammad Idrees[3], Sarfraz Awan[2], Naeem Akhtar Khan[1]

[1]Department of CS University of Centeral Punjab, Lahore, 54600, Pakistan.
[2]Department of CS Virtual University Lahore, Pakistan.
[3]Department of Mathematics, University of Education, Township Campus, Lahore, 54770, Pakistan.
Sjshah786@gmail.com

Abstract: Another method of image processing content based image is said to be possibility of recovery content based information retrieval (CBIR), posing question by image content (QBIC) and (CBIR). This is an application of computer vision meant to explain image retrieval problem. In large databases we have to find the required image by applying some query on the basis of content based shapes, textures colors etc.we find the required data or image. If the ability to estimate or examine the image Content does not exist, in that case search must depend upon metadata like caption or keywords. If the query doesn't match the required contents then it is implemented on some other feature of images to retrieve from the database. This research focus on the Content Based image retrieval with specific domain of Text Based image retrieval (TBIR) system. [Journal of American Science 2010; 6(9):704-707]. (ISSN: 1545-1003).

Key Words: Content based Image Retrieval, Text based Image Retrieval, Textual, Human Perception

1. Introduction

Text-based search gives results with semantic similarity, while content-based search provides results with visual similarity [1]. Owing to the independence between these two approaches, it is likely that their combination could better the performance of search system, benefiting the afore-mentioned approaches.

Work in hand highlights an image retrieval system based on a combined search of text and content. Large collections of digital images are being created in many realms of Commerce, Govt: academia & hospitals [2].

Many of such collections are the end products of digitizing the existing collections of analogue photographs, diagrams, drawings, prints & paintings etc. These collections were searched by keyword indexing or by browsing [6]. Digital images databases open up the way to content based searching. This paper attempts to survey technical aspects of current based image retrieval system. The primary object of such survey is to provide an overview of the functionality of temporary image retrieval system of technical aspects, inquiring or questioning, reliance feedback, features, matching measures, indexing data structures and result presentation. Specific systems rather than general architectures are compared and in this way a basis or defense is provided. User interface consists of a query formulation and a result presentation part.

Specification images to retrieve from the database can be done in number of ways.

There exist a number of approaches for each of these these visual features and each of them presents the feature from a different perspective [5].

Color: Color is one of the vital and widely used visual features in current based image retrieval system. Color histogram is commonly used as representation technique. It is statistically described as combined probabilistic properties of different color channels (as red, green and blue channels) by taking hold of number of pixels having specific properties.

So far as computing distance measures are concerned, it is based on color similarity which can be achieved by computing a color histogram for each image that identifies the proportion of pixels within an image which has specific values & humans express it as colors.

Texture: Texture measures for visual patterns in images, are presented by texels which are put into a number of sets, it depends on how many textures are found out in an image [3]. These sets not only define the textures it also tells location of image in the texture. Identifications of specific textures in an image is got by modeling texture as two dimensional gray level variation [8].

Shape: Shape does not refer to the shape of an image, but to a specific region which is intended to be sought out. Shapes are often

determined first by applying segmentation (parts) or edge detection to an image.

2. Material and Methods:

2.1 Image Color Selection

Since human color perception is based on memory, object identity, culture, and emotion hence it becomes difficult to produce really good color reproduction systems. Human skill at color constancy is not understood properly. Man often errs in understanding model color perception [4].Modeling human color perception becomes difficult due to a range of technical shortcomings colors on photographs are poor representation of the color when seen directly. Compression techniques further vagues the color authenticity used on images presented on internet [2].

Monitor setting (brightness and contrast) also act as color disturbing factors. Complexity of human color perception coupled with technical shortcoming are sources of variability, making color a "noisy" source of information.
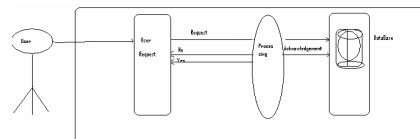

Figure 1

2.2 Text Based Image Retrieval

Content Based image retrieval (CBIR) can find the images on text, shapes, colors etc but this research focuses on textual based image retrieval(TBIR) and get the desired result .It takes the keywords from the dictionary of the images and on the basis of these keywords search the required images [1].By combining Content Based and text based image retrieval we can find the better result. Content Based provides the visual and Text Based provides the in a semantic base results. [7]
In this Text Based image retrieval image came in the form of text stores in the data dictionary and here finds the keywords if the keywords are available in the data dictionary then it shows Yes and if the data doesn't find then it shows No then goes to the image block where it is stored[10].
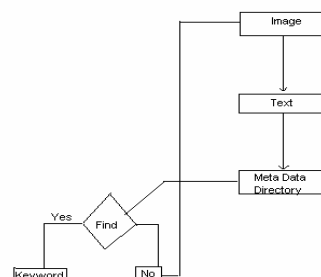

Figure 2 Text Based Image Retrieval Model

2.3 Difference between Computer and Human perception

Computer is a machine having lots of wires and circuits, any job if a command is given to it, it does it in no time and free of faults [9]. On the other hand human mind consists of mussel, nerves and vessels the work assigned to human mind takes time with a chance of errors. Task by human mind is likely to be completed and it may be completed in days, months or year.

2.4 Query Based Image Retrieval

Query provides the CBIR system with an image and then search is based upon search algorithms which vary according to its application; result images must all share common elements explain in the diagram 1. A pre existing image may be supplied by user or taken from stored database [11].

3. Results and Discussion

3.1 Statistical Analysis Text Based image retrieval (Results)

Table-1

No:	Textual Based	Frequency Based
1	X	10
2	Y	12
3	Z	9
4	W	13

This is the Histogram which shows the results of image retrieval from the data base frequency of image retrieval by taking different sampling of data shows that Textual Based image revival is fast as compared to Content Based Image Retrieval.
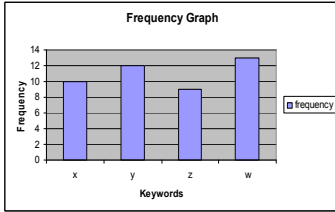

Figure 3 Histogram of (TBIR)

3.2 Pie Chart Representation (Text Based)

Pie Chart shows the results of image retrieval from the data base frequency of image retrieval by taking different sampling of data shows that Text Based image retrieval is fast as compared to Content Based image retrieval.
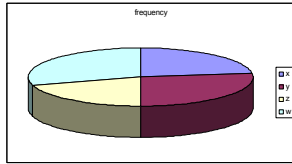

Figure 4 Pie Chart of (TBIR)

3.3 Statistical Analysis Content Based Image Retrieval (Results)

Table-2

No:	Textual Based	Frequency Based
1	X	8
2	Y	10
3	Z	7
4	W	10

It shows the results of image retrieval from the data base frequency of image retrieval by taking different sampling of data shows that Content Based image retrieval is slow as compared to text based image retrieval.
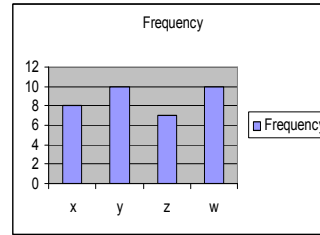

Figure 5 Histogram of (CBIR)

This Pie Chart shows the results of image retrieval from the data base frequency of image retrieval by taking different sampling of data shows that Content Based image retrieval is slow as compared to Text Based image retrieval.
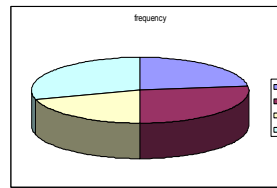

Figure 6 Pie Chart of (CBIR)

4. Conclusion

This research focuses on the content based image retrieval and mainly focused on Text Based image retrieval (TBIR) Content based is visual and Text Based is semantic by comparing above results Text Based image retrieval is fast as compared to Content Based image retrieval (CBIR).

5. Future Work

In this research main work is done about Text Based image retrieval which gives better results as compared to other but this is not the final more research can be done in this field also like Laser Image detection, sementic pattern recognition and unified query image retrieval system.

Corresponding Author:
Jalil Abbas
Department of CS University of Centeral Punjab, Lahore.
E-mail: sjshah786@gmail.com

References:
[1] R. A. Baeza-Yates and B. Ribeiro-Neto. "Modern Information Retrieval". Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.
[2] B. Bustos and T. Skopal. "Dynamic similarity search in multimetric spaces". In *Proc. 8th ACM SIGMM International Workshop on Multimedia Information Retrieval (MIR'06)*, pages 137–146. ACM Press, 2006.
[3] F. Falchi, M. Kacimi, Y. Mass, F. Rabitti, and P. Zezula. "SAPIR: Scalable and distributed image searching". In *SAMT (Posters and Demos)*, vol. 300 of *CEUR Workshop Proceedings*, pages 11–12, Genoa, Italy, November 2007.
[4] M. F. Porter. "An algorithm for suffix stripping". *Program*, 14(3):130-137, July 1980. http://tartarus.or gmartin/ PorterStemmer/156 158 157
[5] The AceMedia project, available at http://www.acemedia.org.
[6] The Google image search page, available at http://images.google.com.
 [7] F. Cacheda, V. Plachouras, and I. Ounis. Performance analysis of distributed architectures to index one Terabyte of text. In *Lecture Notes in Computer Science*, volume 2997, pages 394 { 408, March 2004.
[8] E. Cooke, P. Ferguson, G. Gaughan, C. Gurrin, G. J. F. Jones, H. L. Borgne, H. Lee, S. Marlow, K. McDonald, M. McHugh, N. Murphy, N. E. O. Connor, N. O'Hare, S. Rothwell, A. F. Smeaton, and P. Wilkins. TRECVID 2004 experiments in Dublin City University. In *Proceedings of TRECVID 2004*, November 2004.
[9] C. Gurrin, G. Jones, H. Lee, N. O'Hare, A. F. Smeaton, and N. Murphy. Digital photos: Where and when. In *ACM Multimedia 2005, Singapore, 6-12 November 2005*, 2005.
[10] H. Heaps. *Information Retrieval: Computational and Theoretical Aspects*. Academic Press, 1978.
[11] P. Howarth and S. RÄuger. Trading precision for speed: Localised similarity functions. In *Int'l Conf on Image and Video Retrieval (CIVR, Singapore, Jul 2005)*, pages 415{424. Springer LNCS 3568, 2004.
[12] H. Joho and M. Sanderson. The SPIRIT collection: An overview of a large web collection. *SIGIR Forum*, 38(2):57{61, December 2004.

8/1/2010