ISSN 2229-5518

A Novel Approach for Image Searching using Visual Reranking Technique

Pushpanjali M. Chouragade

Abstract—Rapid advances in computers and communication technology is pushing the existing information processing tools to their limits. The past few years have seen an overwhelming accumulation of media rich digital data such as images, video, and audio. The internet is an excellent example of a distributed database containing several millions of images. Image search has become a popular feature in many search engines, including Google, Yahoo!, MSN, etc., majority of which use very little, if any, image information. Image Retrieval system is a powerful tool in order to manage large scale image databases. Retrieving images from large and varied collections using image content as a key is a challenging and important problem. Due to the success of text based search of Web pages and in part, to the difficulty and expense of using image based signals, most search engines return images solely based on the text of the pages from which the images are linked. No image analysis takes place to determine relevance/quality. This can yield results of inconsistent quality. So, such kind of visual search approach has proven unsatisfying as it often entirely ignores the visual content itself as a ranking signal. To address this issue, we present a new image ranking and retrieval technique known as visual reranking, defined as reordering of visual images based on their visual appearance. This approach relies on analyzing the distribution of visual similarities among the images and image ranking system that finds the multiple visual themes and their relative strengths in a large set of images. The major advantages of this approach are that, it improves the search performance by reducing the number of irrelevant images acquired as the result of image search and provides quality consistent output. Also, it performs text based search on database to get ranked images and extract features of them to obtain reranked images by visual search.

Index Terms—Image Searching, Visual Reranking, Text Based Image Retrieval, Content Based Image Retrieval, Image Ranking & Retrieval Techniques, Pyramid Structure Wavelet Transform, Energy Level Algorithm.

1 Introduction

Text retrieval systems satisfy users with sufficient success. Google and Yahoo! are two examples of the top retrieval systems which have billions of hits a day. The explosive growth and widespread accessibility of community-contributed media content on the Internet has led to a surge of research activity in visual search. However, it remains uncertain whether such techniques will generalize to a large number of popular Web queries and whether the potential improvement to search quality guarantees additional computational cost. Also, the fast development of internet applications and increasing popularity of modern digital gadgets leads to a very huge collection of image database. The database mentioned here can be a small photo album or can be the whole web.

In simple words, an image retrieval system is defined as a computer system for browsing, searching and retrieving images from a large database of digital images. These systems are useful in vast number of applications like engineering, fashion, travels and tourism, architecture etc. Because of the relative ease in understanding and processing text, commercial image-search systems often rely on techniques that are largely indistinguishable from text search. Thus we need a powerful image search engine which will organize and index the images available on web or large database in proper format.

Image database is increasing day by day, because searching images from large and diversified collection using image features as information is difficult and imperative problem. Image search is an important feature widely used in majority search engines, but the search engine mostly employs the text based image search. Commercial image search engines provide results depending on text based retrieval process. There is no active participation of image features in the image retrieval process; still text based search is much popular. Image feature extraction and image analysis is quite difficult, time consuming and costly process. However, it frequently finds irrelevant results, because the search engines use the insufficient, indefinite and irrelevant textual description of database images.[1]

Most research activities have been focused on image feature representation and extraction, classification, similarity measures, fast indexing and user relevance feedback mechanisms. Significantly, the ability to reduce the number of irrelevant images shown is extremely important not only for the task of image ranking for image retrieval applications but also for applications in which only a tiny set of images must be selected from a very large set of candidates.

Multimedia search over distributed sources often result in recurrent images which are manifested beyond the textual modality. To exploit such contextual patterns and keep the simplicity of the keyword-based search, we propose novel reranking method to leverage the recurrent patterns to improve the initial text search results.[2]

Unlike many classifier based methods, that construct a single mapping from image features to ranking, visual reranking relies only on the inferred similarities, not the features themselves[1]. One of the strengths of this approach is the ability to customize the similarity function based on the expected distribution of queries and bridging the gap between "pure" CBIR systems and text-based commercial search engines as a result of reranking. In order to improve the efficiency of database images pyramid-structured wavelet transform is used to obtain energy feature values.

Just type a few keywords into the Google image search engine, and hundreds, sometimes thousands of pictures are suddenly available at your fingertips. As any Google user is aware, not all the images returned are related to the search. Rather, typically more than half look completely unrelated; moreover, the useful instances are not returned first. They are evenly mixed with unrelated images. This phenomenon is not difficult to explain: current Internet image search technology is based upon words, rather than image content. These criteria are effective at gathering quickly related images from the millions on the web, but the final outcome is far from perfect.[3]

When a popular image query is fired, then search engine returns images that occur on page that contains the query term. In real sense, locating query term pictures does not involve image analysis and visual feature based search, because processing of billions images is infeasible and increases the complexity level too. For this very reason, image search engine makes use of text based search. Image searching based on text search possesses some problems like relevance, diversity and typicality. Whenever query is fired, less important or irrelevant images appear on the top and important or relevant images at the bottom of the web page.

For Example, when image query like “d80,” a popular Nikon camera is fired, it provides good image search results but when image query having diversity like “Coca Cola” is fired, searched results provides irrelevant or poor results as shown in Fig.1.

Fig.1(a): Query for “d80” a popular Nikon camera, returns good results on Google.

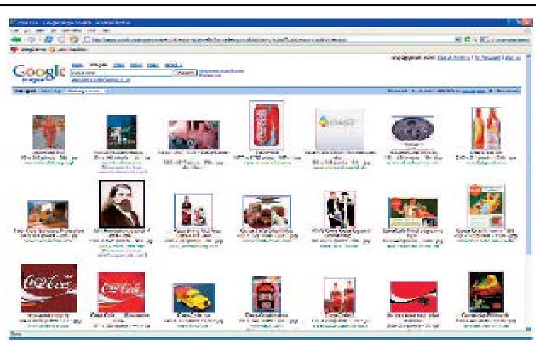

Fig.1(b): Query for (b) “Coco Cola” returns mixed results.

Here, required image of Coca Cola can/bottle is seen at the fourth position in the returned images. The reason behind it is large variable image quality [1].

2 CONTENT BASED IMAGE RETRIEVAL (CBIR)

In the last few years, several research groups have been investigating content based image retrieval. A popular approach is querying by example and computing relevance based on visual similarity using low-level image features like color histograms, textures and shapes. Image retrieval (IR) is one of the most exciting and fastest growing research areas in the field of medical imaging. There are two techniques for image retrieval. The first one uses manual annotation (Text-Based Image Retrieval) and the second one uses automatic features extracted from image larger and larger. Furthermore, it is subjective to the culture, the knowledge and the feeling of each person. The second approach uses features extracted from the image such as color, texture, shape it is independent of people. Reasons for its development are that large image databases, traditional methods of image indexing have proven to be insufficient, laborious, and extremely time consuming. These old methods of image indexing, ranging from storing an image in the database and associating it with a keyword or number, to associating it a categorized description, have become obsolete. This is not CBIR. In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image.

With the ever-growing volume of digital images are generated, stored, accessed and analyzed. The initial image retrieval is based on keyword annotation, which is a natural extension of text retrieval. There are several fundamental problems commonly associated with this approach such as Text search is language-specific and context-specific. Text search is highly error-prone, and Text is cumbersome. To eliminate problems of text-based approach, Content-based image retrieval system is proposed in which query result depend on the visual features of the image (color, texture, shape).

Image Retrieval system is an effective and efficient tool for managing large image databases [4]. The goal of CBIR is to retrieve images from a database that are similar to an image placed as a query. But the basic goal is to bridge the gap between the low-level image properties (stuff) through which we can directly access the objects (things) that users generally want to find in image databases.In CBIR, for each image in the database, features are extracted and compared to the features of the query image. It is a term used to describe the process of retrieving images form a large collection on the basis of features (such as color, texture etc.) that can be automatically extracted from the images themselves. The retrieval thus depends on the contents of images. A CBIR method typically converts an image into a feature vector representation and matches with the images in the database to find out the most similar images.

- “Pure” CBIR systems - search queries are issued in the form of images and similarity measurements are computed exclusively from content-based signals.
- “Composite” CBIR systems - allow flexible query interfaces and a diverse set of signal sources, a characteristic suited for Web image retrieval as most images on the Web

ISSN 2229-5518

are surrounded by text, hyperlinks, and other relevant metadata.

In general, CBIR can be described in terms of following stages:

a) Identification and utilization of intuitive visual features.
b) Features representation
c) Automatic extraction of features.
d) Efficient indexing over these features.
e) Online extraction of these features from query image.
f) Distance measure calculation to rank images.

3 FEATURE EXTRACTION & REPRESENTATION

VERY large collections of images are growing ever more common. From stock photo collections and proprietary databases to the World Wide Web, these collections are diverse and often poorly indexed; unfortunately, image retrieval systems have not kept pace with the collections they are searching. The limitations of these systems include both the image representations they use and their methods of accessing those representations to find images.

In our research work, features like energy level values are extracted for both query image and images in the database, using pyramid structure wavelet transform. The distance (ie., similarities) between the feature vectors of the query image and database are then computed. The database images that have highest similarity to the query image are retrieved and ranked.

The wavelet transform transforms the image into a multiscale representation with both spatial and frequency characteristics. This allows for effective multi-scale image analysis with lower computational cost. Wavelets are finite in time and the average value of a wavelet is zero. A wavelet is a waveform that is bounded in both frequency and duration. Examples of wavelets are Coiflet, Morlet, Mexican Hat, Haar and Daubechies. Of these, Haar is the simplest and most widely used, while Daubechies have fractal structures and are vital for current wavelet applications. So, Haar wavelets are used here.

3.1 Pyramid Structure Wavelet Transform (PSWT)

The pyramid-structured wavelet transform indicate that it recursively decomposes sub signals in the low frequency channels. This method is significant for textures with dominant frequency channels. For this reason, it is mostly suitable for signals consisting of components with information concentrated in lower frequency channels. It is highly sufficient for the images in which most of its information is exist in lower sub-bands.[4]

Using the pyramid-structure wavelet transform, the texture image is decomposed into four sub images, as lowlow, low-high, high-low and high-high sub-bands. The energy level of each sub-band is calculated. This is first level decomposition. Using the low-low sub-band for further decomposition is done. Decomposition is done up to third level in this project. The reason for this type of decomposition is the assumption that the energy of an image is concentrated in the low-low band. Energy of all decomposed images is calculated using energy level algorithm. Using Visual Rerank images similar to query image is retrieved.

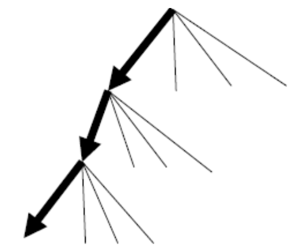

Fig.3.1:Pyramid Structure Wavelet Transform

3.1 Energy Level Algorithm[4]

Step 1: Decompose the image into *four* sub-images

Step 2: Calculate the energy of all decomposed images at the same scale, using:

$$E = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}|X(i,j)| \quad (1)$$

where$\boldsymbol{M}$ and $\boldsymbol{N}$ are the dimensions of the image, and $\boldsymbol{X}$ is the intensity of the pixel located at row $\boldsymbol{i}$ and column $\boldsymbol{j}$ in the image map.

Step 3: Repeat from step 1 for the low-low sub-band image, until it becomes third level.

Using the above algorithm, the energy levels of the subbands is calculated, and further decomposition of the lowlow subband image is also done This is repeated three times, to reach third level decomposition. These energy level values are stored to be used further.

4 IMAGE IMAGE RANKING AND RETRIEVAL TECHNIQUES

Image ranking improve image search results on robust and efficient computation of image similarities applicable to a large number of queries and image retrieval. A reliable measurement of image similarity is crucial to the performance since this determines the extracted features. Global features like color histograms and shape analysis, when used alone, are often too restrictive for the breadth of image types that need to be handled. Image retrieval and ranking technique like PageRank, Topic Sensitive PageRank, VisualRank, VisualSEEK, and RankCompete etc. are introduced to enhance the performance of image search.

4.1 PageRank

Sergey Brin et al. ordered web information hierarchy based on link popularity. A page was ranked higher having more links to it and a page links with higher ranked page, become much highly ranked. PageRank concepts within the web pages have the theory of link structure [1]. It assigns a numerical weighting to each element of documents, which measures its relative importance within the set.

Consider a small universe of four web pages Z, Y, X and W.

Initially, PageRank is considering as 1 and it would be evenly divided between these four documents, hence each document has 0.25 PageRank. If pages Y, X and W are links to the Page Z only, then PageRank of page Z is given as,

$$PR(Z) = \frac{PR(Y)}{L(Y)} + \frac{PR(X)}{L(X)} + \frac{PR(W)}{L(W)} \quad (2)$$

Therefore, PageRank of page Z is 0.75. If Page Y is link to page X as well as page Z, page W link to all other pages and page X link to only Page Z, then PageRank of page Z is,

$$PR(Z) = \frac{PR(Y)}{2} + \frac{PR(X)}{1} + \frac{PR(W)}{3} \quad (3)$$

For M number of document, PageRank for a page is defined as follow:

$$PR(Z) = \frac{1-\xi}{M} + \xi \sum_{j=1}^{m} \frac{PR(Aj)}{L(Aj)} \quad (4)$$

where, *PR (Z)* is PageRank for page *Z*, L (*Aj*) is the number of outgoing link for page *Aj*, *m* is the number of page linked to the page being computed, ξ is the damping factor used in computation. Damping factor ξ lies between 0 and 1 typically being equal to 0.85.

Through whole web link structure, PageRank was created without small subset. The main drawback of PageRank is, a new page with very good quality and it is not a part of existing site, has limited links; as results PageRank method favours the older pages.

4.2 Topic Sensitive Pagerank

The densely connected web pages, through link structure may have higher ranking for the query for which they are not containing resources with useful information. The same web page may have different importance for different query search; it may have higher weightage in one query and less weightage for another. To overcome this, Topic Sensitive PageRank is introduced. In this approach, set of PageRank vector is calculated offline for different topics, to produce a set of important score for a page with respect to certain topics, rather than computing a rank vector for all web pages.

4.3 Visual Rank

With the explosive growth of digital cameras and online media, it has become crucial to design efficient methods that help users browse and search large image collections. The recent VisualRank algorithm [1] employs visual similarity to represent the link structure in a graph so that the classic PageRank algorithm can be applied to select the most relevant images. However, measuring visual similarity is difficult when there exist diversified semantics in the image collection, and the results from VisualRank cannot supply good visual summarization with diversity. This paper proposes to rank the images in a structural fashion, which aims to discover the diverse structure embedded in photo collections, and rank the images according to their similarity among local neighborhoods instead of across the entire photo collection. We design a novel algorithm named RankCompete, which generalizes the PageRank algorithm for the task of simultaneous ranking and clustering. The experimental results show that RankCompete outperforms VisualRank and provides an efficient but effective tool for organizing web photos.

4.4 VisualSEEk

We presented a new image database system which pro- vides for color/spatial querying. Since, the discrimination of images is only partly provided by global features such as color histograms, the VisualSEEk system instead utilizes salient image regions and their colors, sizes, spatial locations, and relationships, in order to compare images. The integration of content-based and spatial querying provides for a highly functional query system which allows for wide variety of color/spatial queries. We presented the strategies utilized by the VisualSEEk system for computing these complex queries and presented some preliminary results that indicate the system's efficiency and power. We will next extend the VisualSEEk system to extract and index regions of texture, and color and texture jointly. We will also investigate and include methods for shape comparison in order to further enhance the image region query system.[12]

4.5 RankCompete

We present a new algorithm named RankCompete, which is a generalization of the PageRank algorithm to the scenario of simultaneous ranking and clustering. The results shows that RankCompete works well for the task of simultaneous ranking and clustering of web photos, and outperformVisualRank on two challenging datasets.

4.6 Comparative Remark

Image searching is popular after introducing PageRank algorithm because it provide good results, but image retrieval is based on text based method so that for diversifies images it provide complex results. To improve the relevancy of image retrieval results number of retrieval techniques are introduced. CBIR uses image features for image retrieval, in Topic Sensitive PageRank number of image feature vectors are calculated offline for different query. VisualSEEK improve fast indexing and provide results based on image regions and spatial outline. VisualRank provide simple mechanism for image search by creating visual hyperlink among the images and employs the way to image ranking for efficient performance. RankComplete uses clustering approach for diversified collections images.

5 VISUAL RERANKING

The basic idea of our content based reranking procedure is that an image which is visually close to the visual model of a query is more likely to be a good answer than another image which is less similar to the visual model. Visual Reranking approach requires extracting features of all images which in turn require image processing and feature creation of each image.

Image is represented by global or local features. A global

ISSN 2229-5518

feature represents an image by one multi-dimensional feature descriptor, whereas local features represents an image by a set of features extracted from local regions in the image. Though, global features has some advantages like requires a smaller amount memory, provide speed and simple to work out but provide less performance compared to local features. Local feature are extracted and represented by feature detector like Difference of Gaussian (DoG) and feature descriptor like Scale Invariant Feature Transform (SIFT), provide better results with respect to different geometrical changes and are commonly used. SIFT descriptor provides the large collection of local feature vector from an image, which does not has effect of image rotation, scaling and translation, etc.

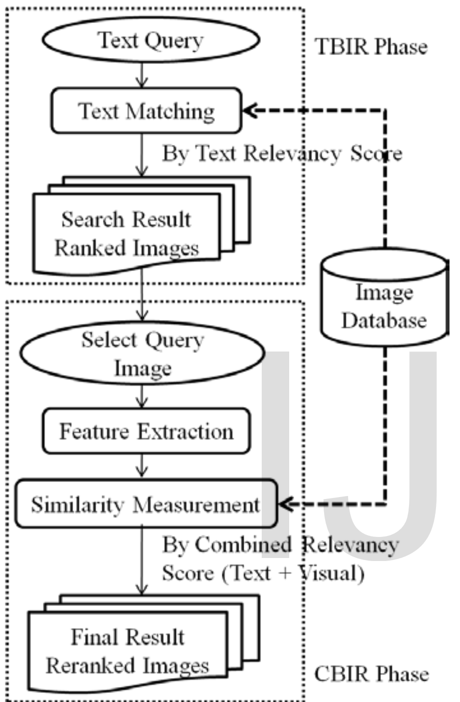

Fig.5.1: Illustration of visual reranking approach.

Both text and visual data can improve over random ranking. They do so using different data and in different situations. When there is a relatively small fraction of relevant images, then visual rerankingmethod performs good but still better than TBIR. Visual re-ranking method can improve over ranking if there is a relatively large number of an irrelevant image.

6EXPERIMENTAL RESULTS

6.1 Performance measurement

The performance measurement can be carried out using precision and recall as given below:

A. Precision:

Precision gives the accuracy of the retrieval system. Precision is the basic measures used in evaluating the effectiveness of an information retrieval system.

$$\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total number of images retrieved}} \quad (5)$$

B. Recall:

Recall gives the measurement in which how fast the retrieval system works. It also measures how well the CBIR system finds all the relevant images in a search for a query image.

$$\text{Recall} = \frac{\text{No. of relevant images retrieved}}{\text{Number of relevant images in the database}} \quad (6)$$

Table 6.1: Performance measurement for some examples.

Sr. No.	Example	Ranking Result (TBIR) 1	Reranking Result (TBIR+CBIR) 2	Precision (3=2/1) 3	Relevant Images in Database 4	Recall (5=2/4) 5
1.	TajMahal	12	11	0.916	11	1
2.	Pyramid	12	11	0.833	10	1.1
3.	Statue of liberty	12	10	0.833	10	1
4.	Toyota	11	10	0.909	10	1
5.	Beach	14	13	0.929	12	1.08 ≈ 1
6.	Bridge	12	10	0.833	9	1.1
7.	Eiffel Tower	12	11	0.916	9	1.2
Total no. of images in database – 85						

From the table it can be observed that almost all the relevant images are retrieved from the database of some known examples.

6.2 Example - “Eiffel tower”

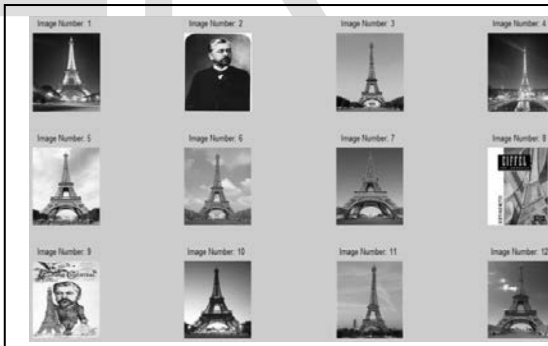

Fig 6.2(a): Database images

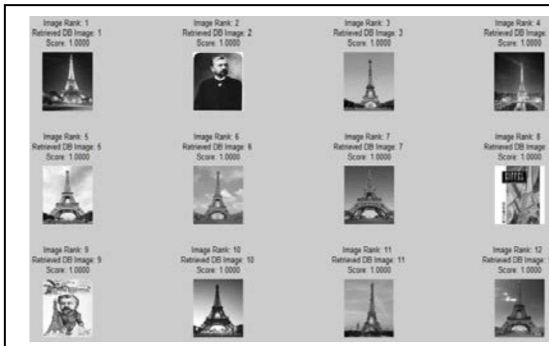

Fig 6.2(b): Ranked images (TBIR result) for text query “eiffel tower”

ISSN 2229-5518

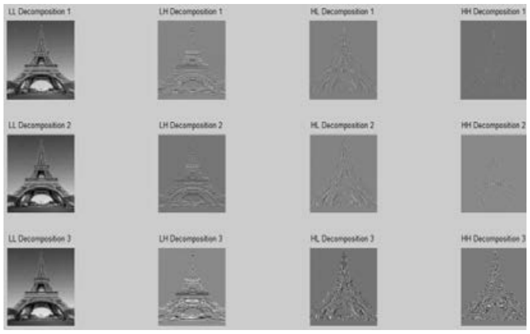

Fig 6.2(c): Pyramid Structure Wavelet Transform decomposition of the selected query image

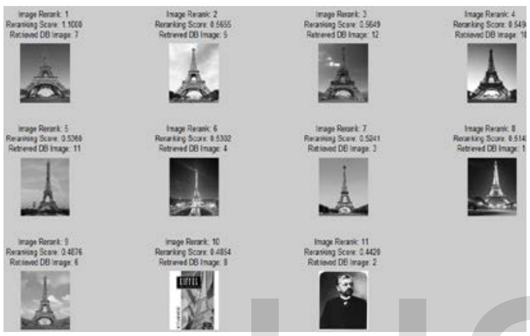

Fig 6.2(d):Reranked images (TBIR + CBIR) for query image (final result).

Figure 6.1: (a) Database images, (b) Ranked images (TBIR result) for text query "eiffel tower", (c) Pyramid Structure Wavelet Transform decomposition of the selected query image, (d) Reranked images (TBIR + CBIR) for query image (final result).

Firstly, the text-based search returns the images related to the input text query from image database and then query image is selected among the resultant images. After this the visual reranking process is applied to refine this result by similarity measurement of both textual and visual information.

7 CONCLUSION

A number of applications are there in which images play a very vital role and some of them are: Education and Training, Travel and Tourism, Fingerprint Recognition, Face Recognition, Surveillance system, Home Entertainment, Fashion, Architecture and Engineering, Historic and Art Research, etc.. The image retrieval system should thus facilitate all these users to locate images that satisfy their demands through queries.

This paper presents a image retrieval system which implements Visual Reranking approach that allows reordering of visual images based on their visual appearance to improve the search performance. Also, it improves the search accuracy by reordering the images based on the multimodal information extracted from the initial text based search results, the auxiliary knowledge and the query example image. The auxiliary knowledge can be the extracted visual features from each images or the multimodal similarities between them. Addition of supplementary local and sometime global feature may offer better image retrieval results.

Visual reranking incorporates both textual and visual cues. As for textual cues, we mean that the text-based search result provides a good baseline for the "true" ranking list. Though noisy, the text-based search result still reflects partial facts of the "true" list and thus needs to be preserved to some extent. In other words, we should keep the correct information in it. The visual cues are introduced by taking visual consistency as a constraint that visually similar samples should be ranked closely and vice versa. Reranking is actually a trade-off between the two cues. It is worth emphasizing that this is actually the basic underlying assumption in many reranking methods, though not explicitly stated.

Visual Rerank approach is one where image get higher ranking, because their similarities matches are more than others, based on common visual similarities present.In the future, we'll develop new methods to speed the reranking processes in large-scale visual search systems. Beyond the visual features used in this work, we'll also explore the use of a large set of generic concept detectors in computing shot similarity or multimedia document context.

REFERENCES

[1] Y. Jing and S. Baluja, "VisualRank: Applying PageRank to Large-Scale Image Search", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 11, 2008.

[2] W. Hsu, L. Kennedy, and S. Chang, "Video Search Reranking through Random Walk over Document-Level Context Graph," *Proc. 15th Int'l Conf. Multimedia*, pp. 971–980, 2007.

[3] R. Fergus, P. Perona, and A. Zisserman, "A Visual Category Filter for Google Images," *Proc. Eighth European Conf. Computer Vision*, pp. 242–256, 2004.

[4] L. Xavier, T. Mary.andN. David, "Content Based Image Retrieval Using Textural Features based on Pyramid–Structure Wavelet Transform", *IEEE Transactions on Image processing*, Vol. 4, No. 5, pp. 79–83, 2011.

[5] Y. Liu, T. Mei,M. Wang, X. Wu and X. Hua, "Typicality–Based Visual Search Reranking", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 5, 2010.

[6] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-Based Image Retrieval at the End of the Early Years," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 22, No. 12,pp. 1349–1380, 2000.

[7] H. Kekre, S. Thepade, T. Sarode and V. Suryawanshi, "Image Retrieval using Texture Features extracted from GLCM, LBG and KPE", *International Journal of Computer Theory and Engineering*, Vol. 2, No. 5, pp. 1793–8201, 2010.

[8] I. Simon, N. Snavely, and S. Seitz, "Scene Summarization for Online Image Collections," *Proc. 12th Int'l Conf. Computer Vision*, 2007.

[9] K. Mikolajczyk and C. Schmid, A performance evaluation of local descriptors, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 27, No.10, pp. 1615–1630, 2005.

[10] R. Datta, D. Joshi, J. Li, and J. Wang, "Image Retrieval: Ideas, Influences, and Trends of The New Age," *ACM Computing Surveys*, Vol. 40, No. 2, pp. 1–60, 2008.

[11] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image Segmentation Using Expectation–Maximization and Its Application to Image Querying," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 24, No. 8, pp. 1026–1038, 2002.

ISSN 2229-5518

[12] J. Smith and S. Chang, VisualSEEK: A fully automated content-based image query system, ACM Multimedia, Boston, MA, pp. 1–12, 1996.

[13] X. He, W. Ma, and H. Zhang, "Imagerank: Spectral Techniques for Structural Analysis of Image Database," *Proc. Int'l Conf. Multimedia and Expo*, Vol. 1, No.1, pp. 25–28, 2002.