IJCSE International Journal of Computer Science and Engineering Open Access
Research Paper Volume-2, Issue-5 E-ISSN: 2347-2693

Daubchies Wavelet transform and Frei-Chen Edge detector for Intention based Image Search Engine

Kshitija Upadhyay[1*] and Gyankamal Chhajed[2]

[1*] *Department of Computer ,VPCOE, Pune University, India,kshitija_20june@yahoo.com*
[2] *Department of C omputer,VPCOE,Pune University, India,gjchhajed@gmail.com*

www.ijcaonline.org

Received: 14/04/2014 Revised: 01/05/ 2014 Accepted: 16/05/May 2014 Published: 31/05/2014

Abstract— Image retrieval is widely used area for number of applications like journalism, medicine, art collections, scientific database .Most of existing image search engines are text query based where retrieval result is ambiguous due to multiple meanings of provided textual query. So proposed system targets at the retrieving relevant images based on user's search intention. A novel image retrieval approach uses Text query and Visual information of image for retrieval .Main objective of this system is to capture the user's search intention in just 'One Click' query image and to display most similar images to this clicked image based on its content . Firstly user's intention is captured by asking user to click one image from result of text based image retrieval .After that clusters of images are formed based on their visual content and visual query hence text query is expanded . Finally Expanded keyword and Visual query expansion are used to retrieve more relevant images. In this paper best combination techniques for important features like Color ,Texture, and shape are used to measure visual similarity between images . Mainly Daubechies' wavelet transform for better frequency resolution and Frei-Chen edge detector which is less sensitive to noise and able to detect edges with small gradients is used .Experimental results shows that our system gives good result by using above combination which is tested on multiple queries and it helps in improving the precision of top-ranked images .

Keywords— *Image reranking, Image search, Intention, Image pool expansion, Keyword expansion, Precison.,Visual features,Visual query expansion*

I. INTRODUCTION

Image retrieval systems are used to browse, search and retrieve images from a large database of images. The use of digitally produced images in areas like journalism, medicine, art collections, scientific database is increasing. For example, medical doctors have to access large amounts of images daily inorder to recognize disorders in human body, home-users often have image databases of thousands of images and journalists also need to search images by various criteria. Text based image retrieval (TBIR) and Content-based image retrieval (CBIR) are two mostly used approches to search relevant images from image database. Computational complexity and retrieval efficiency are the key objectives in the design of Image Retrieval system. However, designing a system with these objectives becomes difficult as the size of image database increases.In this project we are retrieving relevant images reflecting users search intention in just "One Click" query image.We are using both text based and content based approches to retrieve relevant images.

Most of Image search engines that we have today, use keywords as queries.Users enter query keywords to these search engines expecting that they will get images they want .The search engine returns number of images having filenames similar to query keyword or images in which query keyword appear as surrounding text. Text-based image search suffers from the ambiguity of query keywords. The keywords provided by users may be short. They cannot describe the content of images accurately. The search results are noisy and consist of images with quite different semantic meaning. Fig 1.1 shows the images retrieved by Bing image search using apple as query keyword .The result shows number of images of apple belonging to different categories, such as green apple, red apple, apple logo, and apple iphone because of the ambiguity of the word apple.

Query keyword ambiguity may occur because sometimes meaning of the query keyword may be beyond users expectations. For example, word apple have different meanings like red apple, green apple, apple computer, and apple ipod. Also sometime the user may not know how to write textual description of target images in accurate words.

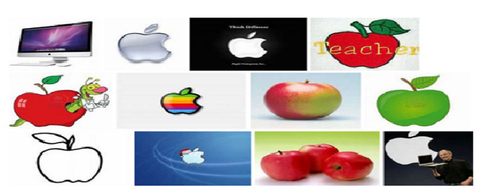

Fig 1.1 Top-ranked images returned from Bing image search using apple as query.

For example, if users do not know bat as the thing related to cricket (shown in Fig1.2) then they have to input bat as query keyword to search images of cricket bat. Sometimes user may get difficulty in describing the visual content of target images using accurate keywords.

To solve this ambiguity, additional information has to be used to capture users search intention. One solution to this is text based keyword expansion, making the textual description of the query more detailed. Existing methods use either synonyms or other linguistic-related words from thesaurus, or words frequently co-occurring with the query keywords for keyword expansion. For example,Google image search provides an option of Related Search which suggests likely keyword expansions.Sometimes, the intention of user can be different and cannot be accurately captured by these expansions. As shown in Fig.1.2, cricket bat is not in the keyword expansions suggested by Google related searches.

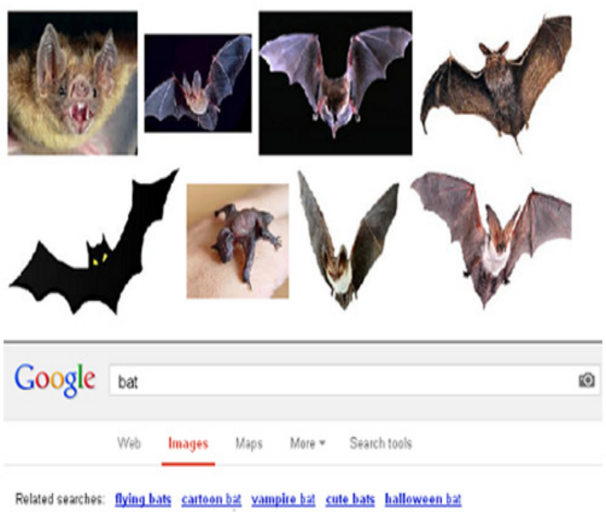

Fig 1.2 (a) Images of bat. (b) Google related searches of query bat.

Content-based image retrieval by relevance feedback can also be used to get relevant images reflecting users intention.In CBIR with relevance feedback users has to label multiple positive and negative image examples. A visual similarity metric is calculated from the selected positive, negative image examples and used to rank images.But these systems require more users effort as they have to select multiple positive and negative image examples .This makes it unsuitable for image search systems in which users feedback has to be minimized.

So to search images visual information should be used.The interaction of user to system should be simple and minimum.An Internet image search approach which captures the users search intention in just One Click is proposed in this paper.The system gives most relevant images by first applying text based image search and after that it it applies content based image search.

II. Related Works

2.1 Image Search

2.1.1. Text Based Image Retrieval (TBIR) [2,3] :

Text based image search engines use only keywords as queries. Users type query keyword and the search engine returns number of images having filenames similar to query keyword or images in which query keyword appear as surrounding text.Text based image search engines rely on text for indexing of images. As a consequence of this, the quality of an image search engine result depends on the quality of the textual information that surrounds or associated with the images (e.g. filename, nearby text, page title, or picture tags within the HTML code).

Advantages: Text based image search is easy to implement.TBIR doesnt require user to have a similar image to search.TBIR is easy to conceptualize as everything is done manually.

Limitation : Text-based image search suffers from the ambiguity of query keywords. The keywords provided by users tend to be short. Also sometime user get difficulty in describing the visual content of target images using accurate keywords. Thus text based image search result consist of noisy images and images with quite different semantic meanings.So the user gets relevant images by text based image search engine only if it is annoted correctly.Manual annotation is needed and in order to fully describe content of images human annotator must provide description of every objects characteristics.

As images contain many details, a comprehensive description of images is usually impossible.If database of images grows in size then it becomes difficult for annotator to annote each image manually.In that case annotation are made by multiple annotator and interpretation of images may vary for each annotator.The user must know exact terms the annotator used in order to retrieve images he wants.

2.1.2. Content Based Image Retrieval (CBIR)[4]:

In Content-based image retrieval (CBIR),user has to provide query image as input to search engine instead of providing text query as in case of text based image retrieval.The user can either browse query image from the hard disk or he can also select the example images provided by us to search the image of that kind.

Content-based image retrieval (CBIR) retrieves images based on features like color, texture and shape.Content-based means that the search will analyze the actual contents of the image. The term content in this context might refer color, shape, texture, or any other feature/information that can be extracted from the image itself.

Features of query image and images in database are extracted and are stored in feature vector.In CBIR feature vector of query image is compared with feature vector of all images stored in database and the image whose feature vector is at minimum difference will be indexed first.Therefore images will be indexed according to visual content for features like color, texture, shape or any other feature or a combination of set of visual features.

Advantages:
In CBIR there is the automatic retrieval of images based on their visual content, whereas the keyword-based approach requires time-consuming annotation of images in database.CBIR retrieves relevant images fastly and doesnt need of manual annotation of images.

Limitation: High feature similarity may not always correspond to semantic similarity .For the same image different users may have different interpretations.

2.1.3. CBIR Relevance Feedback [5]:
CBIR system doesnt retrieve relevant images that user wants in first response to user.The relevance feedback approach has been applied also to content-based image retrieval (Rui et al. 1997b, Taycher et al. 1997, Minka 1996).In CBIR with relevance feedback user has to select multiple positive and negative image examples from image pool.CBIR with relevance feedback was widely used to expand visual query examples. A query-specific similarity metric was learned from the selected examples.

Advantages: In CBIR with relevance feedback, user is allowed to interact with system to refine the results of query until he/she is satisfied.

Limitation: The CBIR with relevance feedback requires more users effort which makes it unsuitable for web-scale commercial systems(Bing image search, Google image search in which users feedback has to be minimized.

2.1.4. Pseudo Relevance Feedback [6,7]:
The most common problem of CBIR with relevance feedback is that it requires more intereaction of user with system as he has to select positive and negative image examples.This problem can be reduced by Pseudorelevance feedback.Pseudorelevance feedback takes the top N images visually similar to the query image as positive examples in order to reduce users burden. However, due to the well-known semanticgap, the top N images may not be all semantically consistent with the query image which may reduce the performance of pseudo relevance feedback.

To verify the spatial configurations of local visual features and to purify the expanded image examples Chum et al.[7] used RANSAC. However, it was only applicable to object retrieval. Users need to draw the image region for the object which he wants to get retrieved and assumed that relevant images contained the same object.
Limitation : Top images retrieved by Pseudo Relevance Feedback may not be semantically consistent with the query image.

2.2 Keyword Expansion

For document retrieval, Keyword expansion is mainly used. Keyword expansion is used to expand text query entered by user. It can also be used to expand retrieved image pool and to expand positive examples of query image. Following are existing systems for keyword expansion.

2.2.1. Thesaurus-based[11] &Corpus-based methods[12] :
Thesaurus-based methods [11] expanded query keywords with either synonyms or hypernyms.Corpus-based methods, such as well-known term clustering[12] and Latent Semantic Indexing , measured the similarity of words on the basis of their co-occurrences in documents.Words that co-occur mostly with the query keywords were chosen as textual query expansion.

2.2.2. Annotation [13]:
Some image search engines provides expanded keywords suggestion. They mostly use surrounding text. Some algorithms [13] generated tag suggestions or annotations based on visual content of input images. The performance of image reranking is not improved by this annotation.

Limitation: For annotation fixed keyword sets is considered, which are difficult to obtain for image reranking in the web environment which is open and dynamic.Supervised training is required for some annotation techniques.Instead of image annotation, method proposed in this paper provides extra image clusters during the procedure of keyword expansions, and this image clusters can be used as visual query expansions which improves the performance of image reranking.

III. METHODOLOGY

Intention based Image Search Engine is implemented using Java with four different interconnected modules:Text based Image Search, Rerank Result based on similarity with query image, Keyword Expansion and Visual Query Expansion, Image Pool Expansion and reranking final result.

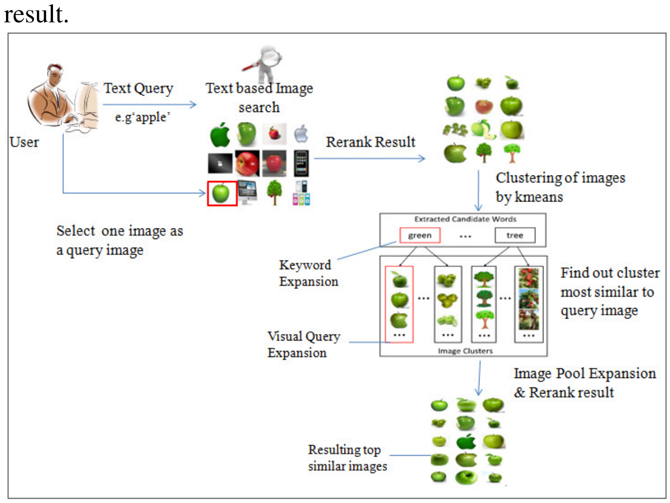

Fig 3.1 System Architecture

3.1 Text based Image Search :
For this step, user has to enter text query keyword q.This module retrieves images based on textual information that surrounds or associated with the images (e.g. filename, nearby text, page title, or picture tags within the HTML code).A pool of images is retrieved by text-based search .Then the user is asked to select a query image from this image pool.

3.2 Rerank Result based on similarity with query image :
In this step, Images retrieved by text based search are reranked based on their visual similarities to the query image.The visual similarities are computed by extracting visual features like color,texture,shape of images. Feature extractor is the mechanism which deals with the extraction of color, shape and texture using various algorithms. The feature extractor extracts the features of the query image and compares that features with the images in the database. The below given design gives the overview of it.
When an image is inserted in database or when the user uploads any query image, its low level features are extracted using various algorithms. We are extracting color, texture and shape feature for every image.

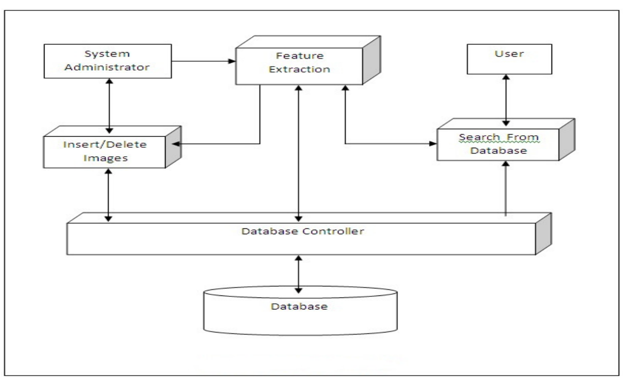

Fig 3.1 Feature Extraction

3.2.1. Color Feature Extraction [14]:

Color Histogram

The histogram of an image is a plot of the gray level values to the number of pixels at that value.Simply the histogram of image is graph of intensity values of a color channel and the number of pixels at that value. From the shape of the histogram we can get the nature of the image or subimage. For example,low contrast image has narrow histogram and bright image has a histogram skewed toward end.Mean, standard deviation, skew, energy, and entropy features are used for color feature extraction.

Mean:
The mean is the average value, it gives general brightness of the image. A bright image will have a high mean and a dark image will have a low mean.

Color Moment:
Color moment of image is used in many image retrieval systems especially when the image contains just the object. The first order (mean), the second (variance) and the third order (skewness) color moments are efficient and effective in representing color distributions of images.
Mathematically, the first three moments are defined as

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} f_{ij} \qquad \sigma_i = \left(\frac{1}{N}\sum_{j=1}^{N}(f_{ij}-\mu_i)^2\right)^{\frac{1}{2}}$$

$$s_i = \left(\frac{1}{N}\sum_{j=1}^{N}(f_{ij}-\mu_i)^3\right)^{\frac{1}{3}}$$

N is the number of pixels in the image and fij is the value of the i-th color component of the image pixel j.

Color Entropy :
The entropy is number of bits we need to code the image data.Enropy of image increases if the pixel values in the image are distributed at many intensity levels. A complex image has higher entropy than a simple image. This measure tends to vary inversely with the energy.

3.2.2. Texture Feature Extraction:[15]Daubechies' Wavelet Transform :
We use Daubechies' wavelet transform for the generation of texture features from each image. Mathematically, wavelet transform is a convolution operation, which can be considered equivalent to passing the pixel values of an image through a low pass filter and a high pass filter.For a two dimensional signal, wavelet transform decomposes it into 4 frequency bands, namely, the LL, HL, LH and the HH band. Haar wavelets do not have sufficiently sharp transition and hence are not able to separate different frequency bands appropriately.On the other hand Daubechies' wavelets have better frequency resolution properties because of their longer filter lengths. We therefore, choose Daubechies' wavelets for the extraction of texture features in our system.

In our application, we use Daubechies' wavelet transform on the pixel intensity values of the complete image for texture feature generation. We apply two-dimensional. Daubechies wavelet transform to each image in the database.

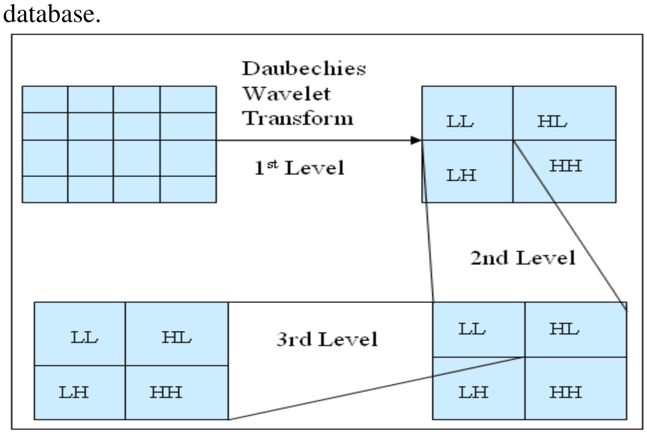

Fig 3.2 Texture feature extraction using Daubehies wavelet Transform.

After the first level wavelet transform, we retain the three high frequency bands, namely, the HL, LH and the HH bands. Standard deviations of the coefficients of these three bands form three features. We then decompose the LL band into four second level frequency bands and get three more features by calculating standard deviations of the three high frequency bands at this level. By decomposing the second level LL band to one more level the last three features are generated and calculating the standard deviations of the higher frequency bands at the third level of decomposition. Since the LL band at the lowest level contains mostly low frequency information (corresponding to image layout), use of the LL band further, adversely affects the retrieval performance. We therefore do not extract any feature from the level 3 LL band. Thus, by calculating the standard deviations of each of the three high frequency bands at each level, we generate a 9-component texture feature vector for an entire image.

Tamura Features:

Coarseness, Contrast, directionality, regularity, linelikeness, and roughness are the Tamura features used to calculate texture weight of image.

3.2.3. Shape Feature Extraction:

Edge detection is mostly used for finding the boundaries of objects within an image. Any abrupt change in frequency of image over the small area within an image is called as an edge.The places in image where the amplitude of the object abruptly changes to the amplitude of the background or amplitude of another object are boundaries of object.

The Gradient Magnitude operation is an edge detector that computes the magnitude of the image gradient vector in two orthogonal directions.It performs two convolution operations on the source image. Edges in one direction are detected by One convolution and edges in orthogonal direction are detected by other convolution. These two convolutions yield two intermediate images.

It squares all the pixel values in the two intermediate images, yielding two more intermediate images.

It takes the square root of the last two images forming the final image.

The result of the Gradient Magnitude operation may be defined as:

$$dsc[x][y][b] = \sqrt{(SH(x,y,b))^2 + (SV(x,y,b))^2}$$

where SH(x, y, b) and SV(x, y ,b) are the horizontal and vertical gradient images generated from band b of the source image by correlating it with the supplied orthogonal (horizontal and vertical) gradient masks.

The Gradient Magnitude operation uses two gradient masks; one for passing over the image in each direction. The Gradient Magnitude operation takes one rendered source image and two parameters.

Parameter	Type	Description
Mask1	Kernel JAI	A gradient mask.
Mask2	Kernel JAI	A gradient mask orthogonal to the first one.

The Frei and Chen edge enhancement operation uses the two masks. This operation, when compared to the other edge enhancement, operations, is more sensitive to a configuration of relative pixel values independent of the brightness magnitude. The following is a listing of how the two masks are constructed.

```
float[] freichen_h_data = { 1.0F   , 0.0F ,  -1.0F,
                            1.414F , 0.0F  ,  -1.414F,
                            1.0F, 0.0F, -1.0F };
float[] freichen_v_data = { -1.0F   , -1.414F , -1.0F,
                            0.0F    ,  0.0F   ,  0.0F,
                            1.0F,  1.414F    , 1.0F  };
KernelJAI kern_h = new KernelJAI (3,3,freichen_h_data);
KernelJAI kern_v = new KernelJAI (3,3,freichen_v_data);
```

3.3 Keyword Expansion and Visual Query Expansion :

In the keyword expansion step, from the textual descriptions (such as image filenames, surrounding texts in the html pages) of the top n images most similar to the query image, words are extracted. These words are then ranked using tf-idf.To save computational cost, only the top m words are reserved as candidates for next processing. However, because the initial image reranking result is still ambiguous and noisy, the top n images may be of different semantic meanings and cannot be used as visual query expansion.

The word having highest tf-idf score computed from the top n images is not the correct word to be chosen for keyword expansion. The method proposed in this paper do keyword expansion by image clustering. For each candidate word wi, all the images containing wi are found and they are grouped into different clusters ci,1 , ci,2 . . . , ci,ti based on their visual content. Images with the same candidate word may have a large diversity in visual content. Images within the same cluster have more semantic consistency as they have more visual similarity to one another and also contain the same candidate word. The visual distance between query image and cluster c is calculated as mean of distance between the query image and images in cluster c. Among all the clusters of different candidate words, cluster c i,j with the largest visual similarity to the query image is selected for visual query expansion and its corresponding word wi is selected for keyword expansion q1 = q + wi.If the distance between the query image and the closest cluster is larger than a threshold , it indicates that there is no suitable image cluster and word to expand the query image and text query and thus query expansion will not be used.

3.4 Image Pool Expansion and Reranking final result :

From the query image and the visual query expansion, a query specific visual similarity metric and a query specific textual similarity metric are computed.The original image pool retrieved by the query keywords q provided by the user plus the image pool retrieved by the expanded keywords q1 forms the new enlarged image pool.Images in the enlarged pool are reranked using the query-specific visual and textual similarity metrics. The size of the image cluster selected for visual query expansion and similarity of image cluster to the query image gives confidence with which system captures the users search intention. If similarity between query image and image cluster is below certain thresholds, then expansion is not used in image reranking.

Do not over-explain common scientific procedures. For example, you do not need to explain how PCR or Western Blotting work, just that you used the techniques. If you are using a novel technique, then you need to explain the steps involved. Use third person passive tense. For example, “RNA was extracted from the cells.” Compare this with, “We extracted RNA from the cells.”Be sure to mention from which companies you purchased any significant reagents for your experiments. When in doubt about how to report your materials and methods, look to papers published in recognized journals that use similar methods and/or materials. Do not mention sources of typical lab ware (beakers, stripettes, pipet tips, cell culture flasks, etc)

IV. Result & Discussion/Experimental/ Analysis/Implementation

In this section the results obtained by Intention based Image Search Engine for different user queries are presented. This section also deals with performance analysis and comparison of Intention based Image Search Engine with Text based Search, Crossmedia, PRF (Pseudo relevance Feedback),NPRF(Negative Pseudo Relevance Feedback) systems. In proposed system images from Corel Database are used as data set.In the set of experiments, random queries were selected from different context.
The performance of Intention Based Image Search Engine can be evaluated using two terms namely precison and recall.

Table 4.1 shows precision and recall values of Intention based Image Search Engine for sample user queries.Figure 4.1 show the average precision of the five techniques. Graphically, precision and recall of Intention based Image Search Engine can be shown as in figure 4.2 and 4.3 for different user queries. The precision and recall values of Intention based Image Search Engine are higher as compared to Text based,Pseudo Relevance Feedback(PRF),Negative Pseudo Relevance Feedback (NPRF),Crossmedia System.This demonstrates that Intention based Image Search Engine can better capture the users search intention with minimum interaction of user with system i.e, just 'One Click'.

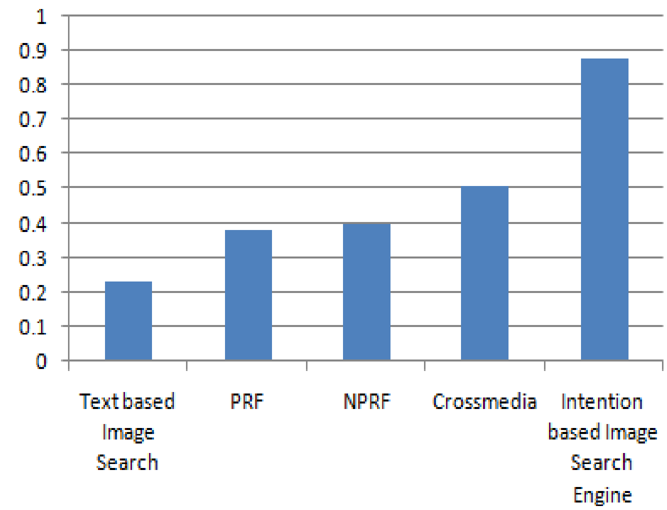

Fig 4.1 Precison Analysis

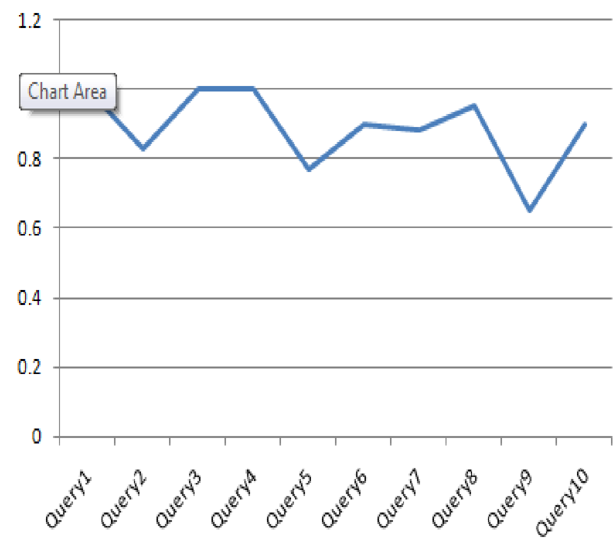

Fig 4.2 Precison Analysis for sample user queries

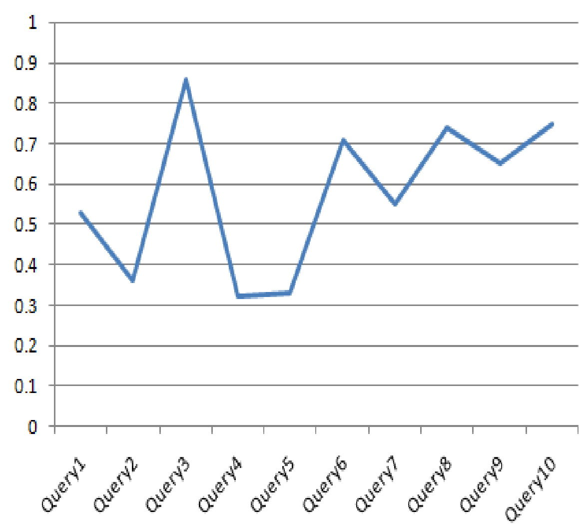

Fig 4.3 Recall Analysis for sample user queries

Text Query	Clicked Query Image	No. of Relevant images in DB	Total No. of images retrieved	No. of relevant images retrieved	Precison
Bat	Cricket bat.jpg	17	9	9	1
Bat	Vampirebat.jpg	14	6	5	0.83
Bear	Blackbear.jpg	14	12	12	1
Bear	Teddybear1.jpg	22	7	7	1
Palm	Palmreading1.jpg	21	9	7	0.77
Palm	Palmreading4.jpg	27	21	19	0.90
Palm	Palmreading.jpg	27	17	15	0.88
Palm	Palmtree.jpg	27	21	20	0.95
apple	Redapple.jpg	15	23	15	0.65

TABLE 4.1 PRECISON VALUES OF INTENTION BASED IMAGE SEARCH

V. CONCLUSION

In this paper,We propose an approach which targets at the retrieving relevant images based on user's search intention . An approach proposed in this paper uses Text query and Visual information of image for retrieval .Visual information in image is measured by extracting color,texture,shape features of image.We observed that Daubechies' wavelet transform that we used for texture feature extraction gives better frequency resolution than other transform. We used Frei-Chen edge detector for shape feature extraction as it is less sensitive to noise and able to detect edges with small gradients.

From the set of experiments that we carried to test the functionality of Intention based Image Search Engine. we observed that with only one time user interaction relevant images reflecting user's intention is retrieved to the satisfaction of user. From the values of precison and recall obtained for sample user queries, we conclude that combining text based(TBIR) and content based(CBIR) approach helps users to easily find relevant images to clicked query image.

REFERENCES

[1] Xiaoou Tang,Ke Liu, Jingyu Cui, "Intent Search: Capturing User Intention for One-Click Internet Image Search", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 34, NO. 7, JULY 2012

[2] N. Ben-Haim, B. Babenko, and S. Belongie "Improving Web-Based Image Search via Content Based Clustering", Proc. Intl Workshop Semantic Learning Applications in Multimedia, 2006.

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

[4] G. Chechik, V. Sharma, U. Shalit, and S. Bengio "Large Scale Online Learning of Image Similarity through Ranking J. Machine Learning Research", vol. 11, pp. 1109- 1135, 2010.

[5] J. Deng, A.C. Berg, and L. Fei-Fei "Hierarchical Semantic Indexing for Large Scale Image Retrieval, Proc. IEEE Intl Conf. Computer Vision and Pattern Recognition,2011

[6] K. Tieu and P. Viola, " Boosting Image Retrieval," Intl J. Computer Vision, vol. 56, no. 1, pp. 17-36, 2004.

[7] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, " Pseudo Relevance feedback:A power tool in interactive content-based image retrieval" IEEE Trans. Circuits Syst. Video Technol., vol. 8, no. 5, pp. 644655, Sep. 1998.

[8] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman, " Total Recall: Automatic Query Expansion with a Generative Feature Model for Object Retrieval", Proc. IEEE Intl Conf. Computer Vision,2007.

[9] Z. Zha, L. Yang, T. Mei, M. Wang, and Z. Wang, "Visual Query Expansion", Proc. 17th ACM Intl Conf. Multimedia, 2009.

[10] A. Frome, Y. Singer, F. Sha, and J. Malik,"Learning Globally-Consistent Local Distance Functions for Shape-Based Image Retrieval and Classification", Proc. IEEE Intl Conf. Computer Vision, 2007

[11] S. Liu, F. Liu, C. Yu, and W. Meng,"An Effective Approach to Document Retrieval via Utilizing WordNet and Recognizing Phrases", Proc. 27th Ann. Intl ACM SIGIR Conf. Research and Development in Information Retrieval, 2004.

[12] K. Sparck Jones, " Automatic Keyword Classification for Information Retrieval", Archon Books, 1971.

AUTHORS PROFILE

Ms.Kshitija Upadhyay obtained masters Degree degree (M.E.) in Computer from Pune University.She has qualified GATE Exam in 2011. Her research interests include Data mining and Information Retrieval, Image Processing.

Prof. Gyankamal Chhajed obtained Engineering degree (B.E.) in Computer Science and engineering in the year 1991-95 from S.G.G.S.I.E.T, Nanded and Postgraduate degree (M.Tech.) in Computer from College of Engineering, Pune (COEP) in the year 2005-2007 . She is approved Undergraduate and Postgraduate teacher of Pune university and having about 16 yrs. of experience. Gyankamal authored a book and has 14 publications at the national, international level for Conferences and Journal. She is life member of the ISTE & International Association IACSIT. Her research interests include Steganography and Watermarking ,Specially for Black and White images ,Data mining and Information Retrieval

