Image Search Using Features and Social Ranking

Nayan Nathu Sawant[1] ,Prof.(Dr).N.R.Wankhade[2]

[1] *Student, Computer Department,late G.N.Sapkal coe, Maharashtra,India*
[3]*HOD and Asst Professor, Department,late G.N.Sapkal coe, Maharashtra,India*

ABSTRACT

Researchers are gaining interest in matching the textual query with visual pictures and their encompassing texts or tags for quick image search. The system proposes a social re-ranking feature for tag-based image retrieval with the thought of images relevance and diversity. The system extracts relevant images based on given image input or tags input. For image retrieval visual features like SURF, GIST and hierarchical wavelet packet descriptor are used. The system applies re-ranking of pictures according to their visual information, semantic info and social clues. The initial results include images contributed by completely different social users. Typically every user contributes many images. Initially system sorts these images by inter-user re-ranking. Users that have higher contribution to the given query rank higher. Then it consecutively implements intra-user re-ranking on the ranked users image set. The system works in two phases: Training and testing. In training phase, image visual information, semantic info and social clues are extracted and saved. In testing phase, test image or tags are matched with training data and re-ranking algorithm is applied to extract top k matched images. The proposed system is enforced and tested using java Experimental results on Flickr dataset show that our social re-ranking methodology is effective and economical.

Index Terms—*specific Social-Visual Ranking(SVR), Social image search, Image re-ranking, Social relevance, feature extraction, tagging*

1. INTRODUCTION

Basically, user intention plays a vital role in image search. Most of conventional image search engines represent user intentions with textual query. Thus, a lot of existing research work highlights on amending the relevance between the textual query & visual images. However, there exists semantic gap between user intention and textual query. Let's take the query[1] "jaguar " as an example, Different users have different intentions when inputting the query "jaguar ". Some are expecting leopard images, while others are expecting automobile images. But the search engine gives the mixed output for the query "jaguar" as shown in fig 1. This scenario is quite common, particularly for queries with heterogeneous concepts or general (non-specific) concepts. This raises a fundamental but yet unsolved problem in Web image search: how to understand user intentions when users conducting image search?

Fig 1: (a): Jaugar as animal (b) Jaugar as car (c) Web search

In the past years, interest analysis is very difficult due to the lack of personal data. With the growth of social media platforms, such as Flickr & Facebook, the way people can get social data has been changed: users profiles, interests & their loved images are exposed online and open to public, which are decisive information sources to implicitly understand user interests.

By analyzing the user requirements, a new algorithm is proposed that exploit social data to assist image search, aiming to improve the relevance between returned images and user interests, which is termed as Social Relevance.

The system also considers the visual information of images. The visually matched images are retrieved from the dataset. The system uses social re-ranking algorithm. The social re-ranking algorithm uses visual information and users social information. The system uses inter user ranking and intra user re-ranking to achieve good tradeoff between relevance score and result diversity. Visual information and views of an images helps to define initial relevance score of image. The system also uses the co-occurrence semantics of tags.

In the following section various images search techniques based on visual information and tag information are studied.

2. REVIEW OF LITERATURE

To improve the visual relevance, several techniques has been proposed mostly based on incorporating visual factors into image ranking like, authors in [2] projected a unique and generic video/image re-ranking algorithmic rule named as IB re-ranking, that reverses results from text-only searches by exposing the salient visual patterns of relevant & irrelevant shots from the estimate relevance rendered by text results. The IB re-ranking technique, based on a strict info Bottleneck (IB) principle, detects the optimum cluster of pictures that preserves the maximal mutual info between the search connection & the high- dimensional low-level visual options of the pictures in the text search results.

In [3] the author is concerned with the matter of multimodal fusion in video search. First, an object-sensitive approach is utilized to query analysis to enhance the baseline results of text-based video search. Then, authors propose a Page Rank like graph-based approach to text-based search result re- ranking. to raised exploit the underlying relationship between video shots, the planned re- ranking scheme at the same time leverages textual connection, semantic construct connection, and low-level-feature-based visual similarity. during this Page Rank like scheme, a set of graphs is built with the video shots as vertexes, and therefore the abstract and visual similarity between video shots as “hyperlinks.” A changed topic-sensitive Page Rank algorithmic rule is then applied on these graphs to propagate the connection scores through all connected video shots.

An essential downside in these ways is to live the visual similarity[4]. As an efficient approach, VisualRank[5] deter- mines the visual similarity by the quantity of shared SIFT features[6]. when a similarity based mostly image link graph was generated, Associate in Nursing iterative computation like PageRank

[7] is employed to re-rank the pictures. Visual- Rank obtains a more robust performance than text-based image search within the measure of connection for queries with homogeneous visual ideas. However, for queries with heterogeneous visual ideas, VisualRank doesn’t work well[8]. Authors in [9]Propose a generic approach that contributes to up the informativeness of image tags by combining gener- ali zations concerning the spatial arrangement tendencies of physical objects within the universe and statistics of linguistic communication use patterns that have been mined from the online. The approach, that we have a tendency to see as Reading between the Tags, provides for every tag related to a picture, first, a prediction regarding quality , i.e., whether or not or not the tag denotes a physical entity, and, then, regarding the real-world size of that entity, i.e., large, medium or tiny. Mining takes place employing a set of Language Use Frames (LUFs) that consisted of linguistic communication neighborhoods characteristic of tag categories.In [10] authors propose a unique algorithmic rule which is scalable and dependably learns tag connection by accumulating votes from visually similar neighbors. Further, treated as tag frequency, learned tag connection is seamlessly embedded into current tag-based social image retrieval paradigms.

Studying the existing system we find that combining social relevance and visual relevance faces the following challenges:

1) Social data sparseness: In social media platform, most users only possess a small number of favored images, from which it is difficult to discover user intentions. With the hypothesis that users in the same

community share similar interests, a community-specific method is more practical and effective than a user-specific method.

2) The tradeoff between social relevance and visual relevance: Although social relevance may guarantee the interest of returned images for the user, the quality and representativeness of images, cannot be ignored. Both of which are essential for good search results. Thus, some social relevance and visual relevance are needed to be addressed and subtly balanced.
3) Complex factors: To generate the final image ranking, one needs to consider the user query, returned images from current search engines, and many complex social factors derived from social media platforms. How to integrate these heterogeneous factors in an effective and efficient way is quite challenging.

To solve this issue author in [1] proposes a social re-ranking algorithm in which user information is firstly introduced the conventional ranking method considering the semantics, social clues & visual information of images. A tag-based image search approach with social re-ranking is proposed in this paper. It systematically fuse the visual information, social users information and image view times to boost the diversity performance of the search result also propose the inter-user re-ranking method and intra-user re-ranking method to achieve a good trade-off between the diversity and relevance performance. These methods not only reserve the relevant images, but also effectively eliminate the similar images from the same user in the ranked results and in the intra-user re-ranking process. The system combines the visual, semantic and Views information into a regularization framework to learn the relevance score of every image in each users image set. To speed up the learning speed, we use the co-occurrence word set of the given query to estimate the semantic relevance matrix[13].

For visual information extraction hierarchical wavelet packet descriptor (HWVP) descriptor is used. This is texture based feature extraction technique. The search result depend on the various factors like template size, pixel sistribution, color brightness, etc. The best result can be extracted if texture based and shape based object detection techniques are combined[11]. Surf is the image processing technique that extracts effective scale and transforms invariant features. SIFT is very efficient in representing images with salient structures.[12]

Problem Formulation:

Tag based image retrieval faces problems like:

1: Tag mismatching: User has freedom to add tag to their own images and there is no ontology or taxonomy for tag assignment in social networks hence many seemingly irrelevant tags are introduced.
2: Tag ambiguity: polysemy and synonyms are important aspects in query ambiguity. E.g. Query “jaguar” may refer to animal in jungle or a car.
Tag based retrieval of images based on social aspects like user who have tag the image, User contribution in that domain, image views, Semantic Relevance between tags, etc.
Visual information matching plays vital role in finding similar images. There is need of such system that find top k similar images based on visual information and social information re-ranking.

3.SYSTEM ARCHITECTURE / SYSTEM OVERVIEW

PRILIMNARIES:

1: Hierarchical Wavelet Packet Descriptor (HWVP): This is a texture analysis approach[14]. Sub-band filtering is used in this technique. Images of different categories have different properties to the wavelet packet filters. Texture information of the sub-bands should be effective for object categorization. This is an object categorization method.

2: SURF: Speeded up robust features(SURF)[15] is one more technique. This technique is quite faster than SIFT technique. SURF is based on Hessian matrix technique. SURF evaluates the approximation of image descriptors using integral images.

3: GIST: For feature extraction, color GIST[15] is a method that summarizes the image gradient information in terms of scale and orientation. This GIST technique convolves the image with 32 graber filters and generates the 32 feature map. This feature map is then divided in equal 16 sections i.e. 4X4 grid and generates the average value of each grid. Assembling all the 16 section values for each feature map 16X32 feature values are extracted.

Following figure shows the architecture of system. The architecture is mainly classified in 2 sections: Training and testing. In training phase images and its social information is given as input. And in testing phase image tags are generated and tag based similar images are retrieved.

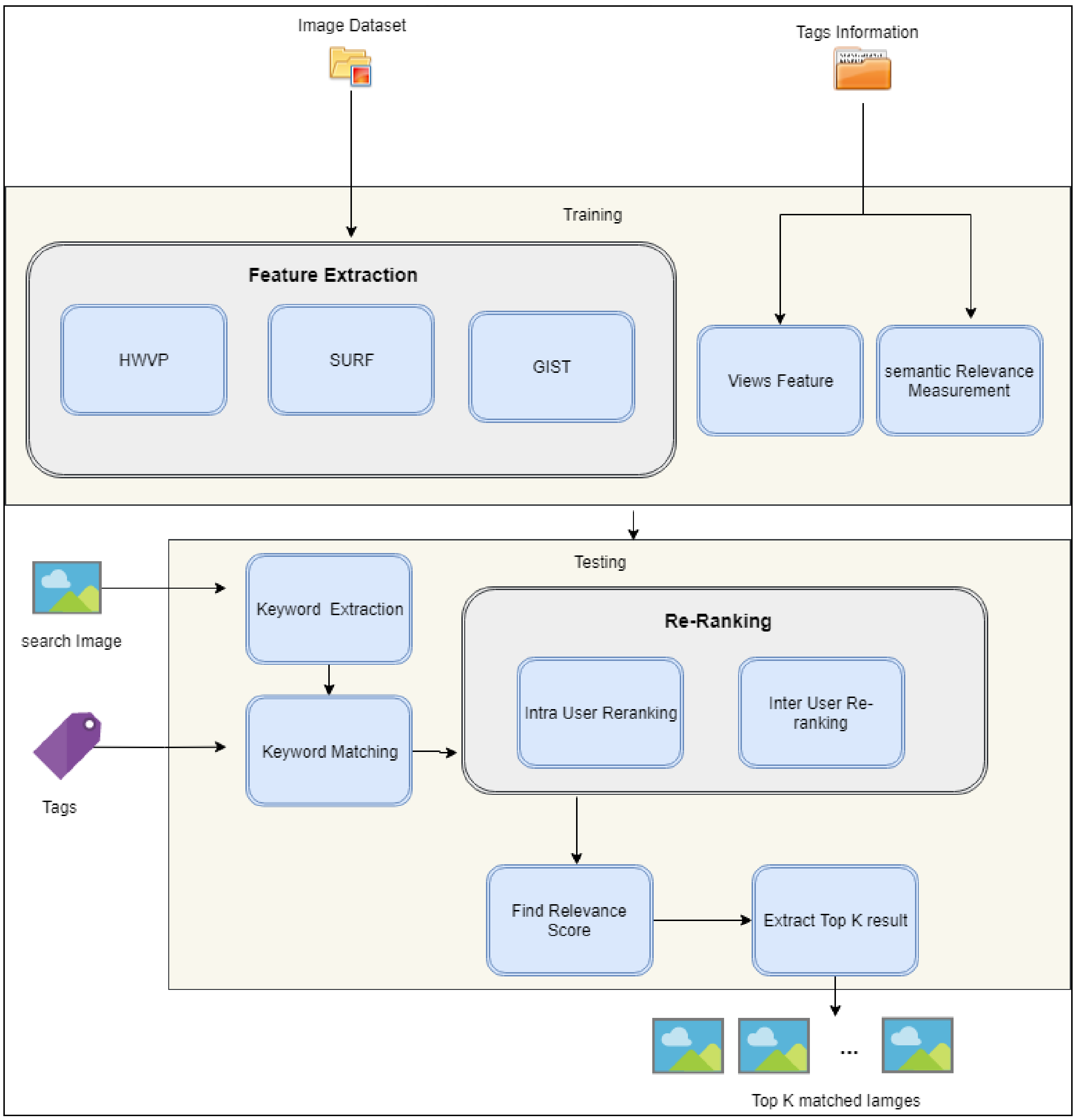

Fig:. System Architecture

3.1 The Training part includes the following 2 phases:

1: Feature Extraction:

The training image dataset is given as input to the feature extraction phase. The features of images are extracted and saved in training phase. For feature extraction HWVP, SURF and GIST feature extraction techniques are used. The HWVP extracts. To implement this technique OPENCV-3.1 java library is used.

2: View Features:

The uploaded images are viewed by other users. The view feature describes the social media click count of an image. The number of click counts helps to calculate relevance score of an image in image retrieval phase. In ags re-ranking phase, view feature helps to re-rank the generated image similar image result set. The image having more click count is treated as popular one and displayed at the top position in search result.

The normalized view score Vt(i) of an image I is calculated as:

$$Vt(i) = \frac{View_i - View\_min}{view\_max - view\_min} \quad \text{------(eq. 1)}$$

where view_max and view_min are the maximum and minimum views of images present in training set.

3: Semantic relevance measurement:

In semantic relevance calculation, co-occurrence of tags are considered. This measurement is used to remove noisy tags, remove tags having influence of the seldom-used tags. The relative frequency of each tag is considered. The system finds top k co-occurrence tags for suggestion.

This is calculated using normalized Google distance NTD. This provides pair wise concept distance by computing the co-occurrence of concepts in web pages. NTD uses largest world wide web dataset. The information is provided by billions of users and it is independent of personal knowledge. This overcomes the limitations of local dataset.NGD (ci,cj) $= \exp\{-\frac{\max\{Log\ G(ci), Log\ G(cj)\} - \log G(ci,cj)}{\text{Log nG} - \min\{Log\ G(ci), Log\ G(cj)\}}\}$ ------(eq. 2)

Where,

nG : The number of pages indexed by Google

G(ci,cj) : The number of pages containing tags ci and cj.

G(c) : The number of pages containing concept c reported by Google search engine.

The testing part includes the following four phases:

1: Tags Retrieval:

For testing image is given as input. Based on the feature extraction methods image is matched with dataset images using hierarchical wavelet packet descriptor (HWVP), SURF and GIST and most relevant top k images are retrieved. Based on the matched images set tags of image is retrieved. All images in top K matched results are present in training dataset. Along with the image tags are also present in a dataset.

Let {I1, I2,..IK} are matched image set and {T1,T2,..Tn} are matched tags. Then the tags for test image is suggested if t belongs to more than one Ti set.

2: Keyword Matching

The extracted tags from uploaded image or tags given by image are compared with the training data. The image set Xh is retrieved from training image dataset X with matched tags. The image and its tag matching score is calculated. The images are further re-ranked using social user's information into account.

3: Inter user Re-ranking

After matching the keywords with the training dataset, the users are ranked by inter user re-ranking. The ranking of user varies with respect to query. Query based user ranking is defined. To calculate the user ranking, user contribution in particular domain tags is considered.

The Inter user ranking Eh is calculated as:

Eh = $\sum_{j=1}^{k} sign(Xhj)$--(eq. 3)

Where Xh is the matched keyword image set

k is the size of Xh

$sign(Xhj)$ is calculated as :

$sign(Xhj) = 1\ if$ tag q is present in tags added by user

$else$

$sign(Xhj) = 0$

4: Intra user Re-ranking

After calculating the inter-user ranking, The image set is sorted in descending order by the user contribution in tag generation. From each users image set, images are retrieved using higher relevance score.

The relevance score ri of ach tag in query is calculated using regularization framework. It combines the score of visual, semantic and views information. It can be calculated as:

Qr = $\sum_{i,j=1}^{k} Wij\ \{{}^{ri}/_{\sqrt{Dij}} - {}^{rj}/_{\sqrt{Dij}}\}^2$ + $\alpha\sum_{i=1}^{k}(ri - Ci)^2$ + $\beta\sum_{i=1}^{k}(ri - vti)^2$ -----------------------------------(eq. 4)

Where,

Wij is the visual distance of image i and j.

Ci is semantic relevance score

Dij = = $\sum_{i,j=1}^{k} Wij$

Vti = Views of image i

Algorithm work:

Processing:
Input: T: Training data containing images and xml

sm: Search Image

st: Search tags

K: No of result images required

Output:

TR: Training result

Imgt: Tags set of search image

Im: Matched image set

Processing:

/*training phase*/

1. For each image in T
   Apply HWVP and save features
   Apply GIST and save features
   Apply SURF and save features
   End For
2. For each xml in T
   Parse XML and extract usernames, tags and view count
   Find normalize view count for each image using eq. 1
   End For
3. For tag I in tag set
   For tag j in tag set
   If I <> j then
   NGD (i,j) : Calculate semantic relevance measurement between tag I and j using eq 2
   End If
   End For
   End For
4. Save NGD matrix

/*testing phase : Find tags from image*/

5. Get uploaded image and extract features using
   - HWVP(sm)
   - GIST(sm)
   - SURF(sm)
6. For features of Training image i in T
   - Match HWVP(sm) with HWVP(i)
   - Match SURF(sm) with SURF(i)
   - Match GIST(sm) with GIST(i)

   End For

7. Find top k Matched images
8. Extract tags set Ts of each image
9. imgT: Find tags tj Ɛ all images
10. Generate image tags list as imgT

/*testing phase : Find similar images from tags*/

11. S1: keyword matching and find image set containing the tags for input tag st,
12. Ut: Find user contribution for tags
13. S2: Find inter user ranking based on Ut and sort the images in S1 using eq. 3
14. Read view count and semantic information from training data for tags in st
15. S3: Find Intra-user re-ranking using eq. 4
16. Sort S3 and select top k result

4. SYSTEM ANALYSIS AND RESULT

4.1 Implementation:

Implementation:

The system is implemented in java using jdk 1.8. A desktop application is ceated using swing components. The system is tested on i5 processor with 4 gb ram on windows10 environment. For feature extraction OpenCV3.1 is used.

Dataset: A fliker[16] dataset is used for testing the system. The data set includes images and its relative social information in xml Form.

The xml includes following tags:

1: user who has tagged the image

2: Tags of image

3: views of images

4: Date of tag creation

Performance measures:

1: Time: time required for searching similar images is captured.

2: AP@n:

This is the average precision under depth n. This is the relevance performance measure.

To calculate AP@n, a system results are evaluated manually. We have evaluated the results from 5 volunteers. The result quality is observed and relevance score is given to each training images as:

0: Irrelevant,

1: Average

2: Good

3: Perfect

Using the score given by 5 volunteers, the average of relevance score is calculated for each image. The final value of AP@n is calculated as:

AP@n = $\frac{1}{n}\sum_{1}^{n}(\sum_{j=1}^{i}\frac{rel-j}{i})$---(eq. 5)

3: NDCG :

This is Normalized Discounted Cumulated Gain. A relevance measure based on the ranking position of image[14]. The normalized value 1 represents the ideal performance. The NDCG is calculated as:

NDCG = $\frac{1}{w}\sum_{i=1}^{n}\frac{2^{lev(i)}-1}{\log(1+i)}$---(eq. 6)

W is a normalization constant set as 0.1

4: System Features: The proposed system features are compared with existing systems. The feature list includes facilities provided and technique used.

Results:

1. System Results:
   Following figure 3 and 4 shows the results of system. For searching tags: Sea, water, Boat are given. The results of proposed system are more relevant than the existing system.

Fig 3: Search Result Existing System(Tags : Sea, water, Boat)

Fig 4: Search Result Proposed System(Tags : Sea, water, Boat)

Following table show the result for automatic tag generation. Image as an input is given to the system. Based on the training data Image tags are generated. The input image features are compared to the training data and annotations are retrieved.

Table 1: Image Tag generation result

Images	Annotations
	sky, water sea
	sky, road, car occluded, window occluded, building
	sky, road, car, sign, streetlight

1. Time Evaluation:
   Following table shows the time comparison for image searching for existing and proposed system. The proposed system uses HWVP, SURF and GIST feature extraction techniques. The existing system only

uses HWVP. The proposed system requires little extra time as compared to the existing system but creates more accurate results.

Table 2: Time Evaluation

Result Count(n)	Execution time for Existing System(in Sec)	Execution Time for Proposed System (in Sec)
5	1.54	1.68
10	1.75	2.03
15	1.91	2.65
20	2.19	2.79
25	2.34	2.94

The following graph shows the time comparison between existing and proposed system. The proposed system requires more time than existing system due to additional feature extraction techniques such as SURF and GIST.

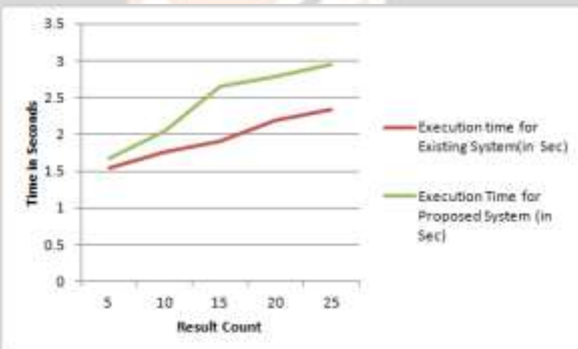

Fig 5: Time Evaluation

2. AP@n Evaluation:
   This is the accuracy measure. The accuracy of existing and proposed system is compared using average precision under depth n (AP@n) technique using eq. 5. The accuracy of proposed system is higher than the existing system.

Table 3: AP@n Evaluation

Result Count(n)	AP@n-Existing system	AP@n-Proposed System
5	2.7	2.78
10	2.6	2.71
15	2.66	2.71
20	2.62	2.69
25	2.63	2.65

Following graph shows the comparative study between existing and proposed system based on average precision under depth n(AP@n). The AP@n score of proposed system is higher than the existing system

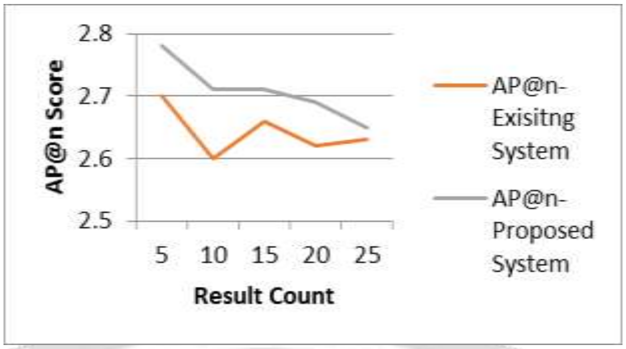

Fig 6: AP@n Evaluation

3. NDCG evaluation:
   The following table shows the comparison between existing and proposed system based on the Normalized Discounted Cumulated Gain factor eq. 6. This is the accuracy evaluation factor based on the result quality. The proposed system has higher NDCG value than the existing one.

Table 4: NDCG Evaluation

Result Count(n)	NDCG-Existing System	NDCG-Proposed System
5	0.8	0.83
10	0.85	0.87
15	0.83	0.85
20	0.86	0.87
25	0.91	0.92

Following graph shows the comparative study between existing and proposed system based on Normalized Discounted Cumulated Gain factor (NDCG). The NDCG score of proposed system is higher than the existing system

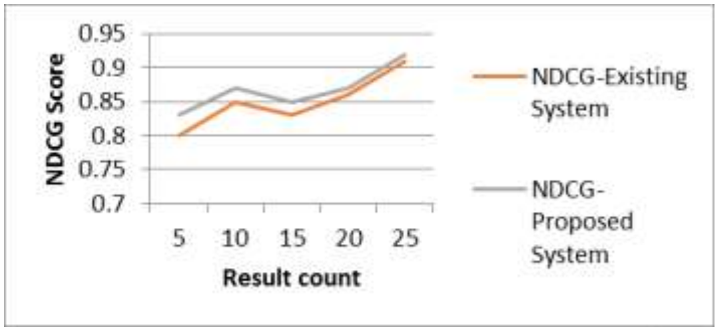

Fig 7: NDCG Evaluation

4. System Comparison:

Following table represents the comparative study between existing and proposed system in terms of techniques used and facilities provided. For feature extraction of image HWVP, SUF and GIST are used. For tag suggestions Tag co-occurrence semantics and Social information usage are compared. Based on the tags, similar images are retrieved in SR[1] and in Image Annotation[15]. The proposed system provides searching of similar images based the input image.

Table 5: System Comparison

	Techniques					Facilities Provided		
	Feature Selection using HWVP	Feature Selection using SURF	Feature Selection using GIST	Tag co-occurrence semantics	Social information usage for tag generation	Tag Based Image Search	Automatic Tag generation	Image based Search
Image Search[11]	YES	YES						YES
Image Annotation[15]		YES	YES	YES			YES	
SR[1]	YES			YES	YES	YES		
Proposed System	YES	YES	YES	YES	YES	YES	YES	YES

Table 5: System Comparison

5. CONCLUSION

The proposed system uses a social re- ranking methodology for tag-based image retrieval. Along with visual features, system uses social information to re-rank the tags. Feature extraction techniques such as HWVP, SURF and GIST increases the accuracy of image extraction based on visual features. During this social re-ranking methodology, inter-user re-ranking and intra- user re-ranking are applied to get the retrieved results. Besides views of social image is additionally foremost fused into a conventional regularization framework to enhance the relevancy performance of retrieved results. The system also proposes a technique of automatic tag retrieval from given image. Discussions and experiments have incontestable that our projected methodology is effective and time saving.

In future, tag generation process can be more refined based on additional information like date of tag, user's social network, etc. Individual object detection strategy [15] can be integrated for more precise tag generation based on visual features.

6. REFERENCES

[1] Liu, Shaowei, et al. "Social visual image ranking for web image search." International Conference on Multimedia Modeling. Springer Berlin Hei- delberg, 2013.

[2] W. H. Hsu, L. S. Kennedy, and S.-F. Chang."Video search reranking via information bottleneck principle." In Proceedings of the 14th annual ACM international conference on Multimedia, MULTIMEDIA 06, pages 3544, New York, NY, USA, 2006. ACM.

[3] Liu, Jingjing, et al. "Video search re-ranking via multi-graph prop- agation." Proceedings of the 15th ACM international conference on Multimedia. ACM, 2007.

[4] R. I. Kondor and J. Lafferty. "Diffusion kernels on graphs and other discrete structures." In In Proceedings of the ICML, pages 315322, 2002.

[5] Y. Jing and S. Baluja. "Visualrank: Applying pagerank to large-scale image search." Pattern Analysis and Machine Intelligence, IEEE Trans- actions on, 30(11):18771890, nov. 2008.

[6] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata,
A. Tomkins, and J. Wiener. "Graph structure in the web." Computer Networks, 33(1C6):309 320, 2000.

[7] L. Page, S. Brin, R. Motwani, and T. Winograd. "The pagerank citation ranking: Bringing order to the web." Technical Report 1999-66, Stanford InfoLab, November 1999. Previous.

[8] Y. Liu, T. Mei, and X.-S. Hua."Crowdreranking: exploring multiple search engines for visual search reranking." In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, SIGIR 09, pages 500507, New York, NY, USA, 2009. ACM.

[9] Larson, Martha, Christoph Kofler, and Alan Hanjalic. "Reading between the tags to predict real-world size-class for visually depicted objects in images." Proceedings of the 19th ACM international conference on Multimedia. ACM, 2011.

[10] Li, Xirong, Cees GM Snoek, and Marcel Worring. "Learning tag relevance by neighbor voting for social image retrieval." Proceedings of the 1st ACM international conference on Multimedia information retrieval. ACM, 2008.

[11] Ravi Kumar, Kallakunta & Bhanu Prakash, V & Shyam, V & Anil Kumar, M. (2016). Texture and Shape based Object Detection Strategies. Indian Journal of Science and Technology. 9. 10.17485/ijst/2016/v9i30/98709.

[12] Qian, Xueming & Guo, Danping & Hou, Xingsong & Li, Zhi & Wang, Huan & Liu, Guizhong & Wang, Zhe. (2014). HWVP: Hierarchical wavelet packet descriptors and their applications in scene categorization and semantic concept retrieval. Multimedia Tools and Applications. 69. 10.1007/s11042-012-1151-8.

[13] D. Lu, X. Liu and X. Qian, "Tag-Based Image Search by Social Re-ranking," in *IEEE Transactions on Multimedia*, vol. 18, no. 8, pp. 1628-1639, Aug. 2016.

[14] K.Jarvelin and J.Kekalainen. Cumulated Gain-Based Evaluation of IR Techniques. In ACM Transactions on Information System, 2002.

[15] Linan Feng, Bir Bhanu,"Semantic Concept Co-Occurrence Patterns for Image Annotation and Retrieval," IEEE transaction on Pattern Analysis and machine intelligence, VOL. 38, NO.4, pp:785-799,APRIL 2016

[16] Flickr. http://www.Flickr.com/.