Proceedings of
INTERNATIONAL CONFERENCE ON COMPUTING, COMMUNICATION AND ENERGY SYSTEMS (ICCCES-16)
In Association with IET, UK & Sponsored by TEQIP-II
29th-30th, Jan. 2016

Paper ID: E&TC16

PERFORMANCE EVALUATION OF PROPOSED IMAGE RETRIEVAL SYSTEM FOR IMAGE-RICH INFORMATION NETWORKS BASED ON INTEGRATED ALGORITHM (S-CBIR)

Vishal S. Kore,

Parag P. Abhyankar,

Amar Palwankar

Abstract—Due to existence of different related information to images (tags, annotations, comments etc.) retrieval of images from image rich information networks is very challenging and complex task. Integrated algorithm (S-CBIR) issued to retrieve images from image rich information networks. The experiment is conducted to check the accuracy of proposed integrated algorithm (S-CBIR). Performance evaluation of proposed image retrieval system for image-rich information networks based on integrated algorithm (S-CBIR) is done in this paper. The results of analysis shows that the proposed image retrieval system based on integrated algorithm (S-CBIR) not only refines the search results but also improves accuracy of search results than the link based (Sim-Rank) and content based image retrieval techniques (CBIR, CEDD, SIFT, Color descriptor, texture extraction).

Index Terms— Content based image retrieval, image-rich information networks, image retrieval.

I. INTRODUCTION

SOCIAL networking and image hosting websites like Face-book and Flickr allow users to upload, tag and share images on internet, such type of sites have millions of users who actually upload millions of images daily. Famous e-commerce web sites like Amazon, flipkart, and snapdeal etc. also show product related images to customers. Images which are available on social multimedia networking, image hosting and e-commerce websites are accompanied by different annotations, tags, comments, and other related information which tend to form an image rich information networks. Due to existence of different related information to images (tags, annotations, comments etc.) retrieval of images from image rich information networks is very challenging and complex task.An integrated approach of algorithms to retrieve an image from image rich information networks is proposed in the previous work [1].Survey of image retrieval techniques and algorithms for image-rich information networks has been done in previous work [2]. The work done in paper [1] explains the detail information about the proposed image retrieval system for image-rich information networks based on integrated algorithm (S-CBIR). Performance evaluation of proposed image retrieval system for image-rich information networks based on integrated algorithm (S-CBIR) is done in this paper.

II. SYSTEM DESIGN

The proposed integrated algorithm (S-CBIR) [1] is inspired by the work done in paper [3]. The image-rich information network is developed and then image retrieval is performed. The image-rich information simulation network model is developed by using Flickr API dataset and tried to estimate the same type of nodes of a network model with the help of manual impact. The proposed and implemented integrated algorithm (S-CBIR) uses link based information and content based information to retrieve relevant images in image rich information networks. Proposed integrated (S-CBIR) algorithm is a combination of link based image retrieval system (K-SimRank) and content based image retrieval system (CEDD, Feature Extraction, Color Extraction and SIFT etc.). The proposed algorithm uses the global features like color and histogram and local features like shape and texture for feature extraction of images [1].

A. System Architecture

Fig. 1 demonstrates the system architecture of proposed image retrieval system for image-rich information networks based on integrated algorithm (S-CBIR). It has three-tier architecture where web browser, application server and database are present. User can search images through web browser, web browser sends that request to application server where the actual program runs and relevant images are searched, retrieved, indexed from database and result images are displayed to user in browser.

First Tier: The web browser is present where client can communicate with system through a web page on which search function is made available to user. User just need to write query to search link based relevantimages from database. The results i.e. retrieved images are also displayed in web

Paper ID: E&TC16

browser on result page. After that user need to select one image to give input to CBIR system.

Second Tier: In this tier application server and actual program is present. In program there are two modules first is link based image retrieval system where K- SimRank algorithm is used and second module is content based image retrieval where different feature extraction descriptors like SIFT, CEDD, Color, feature and texture extractors are used.

Third Tier: In this tier actual database is present which contains images. Different feature vectors of images are also stored in database [1].

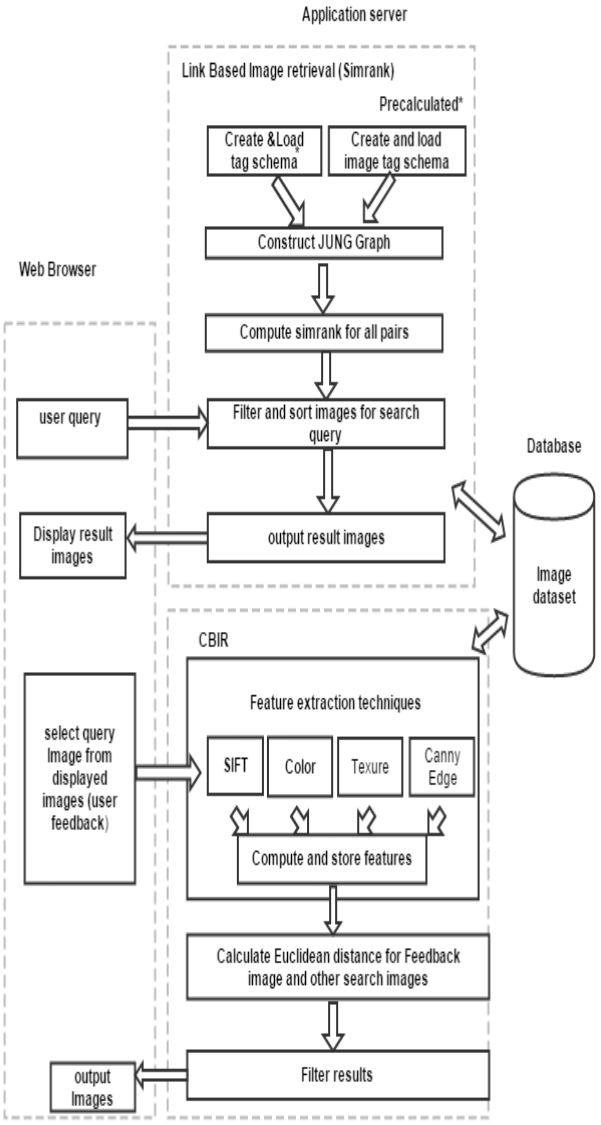

Fig. 1. System Architecture of Proposed image retrieval System based on integrated algorithm (S-CBIR) [1].

B. Proposed Integrated Algorithm (S-CBIR)

Main steps in proposed integrated algorithm (S-CBIR) are as follows

1. Preprocessing step:

1. Load common tags from image dataset file into a database table1.
2. Load image ids and tag ids from image dataset into another database table2.
3. Create JUNG graph to represent the relation between image nodes and image tag nodes.
4. Compute and apply SimRank on JUNG graph.
5. Serialize objects of SimRank in a file.
6. Extract and store features of all images in dataset.

2. Integrated algorithm (S-CBIR)

(a) Link based image retrieval: Input: - keyword query (tag)

i. Check whether entered keyword is present in table 1, if tag is present in table 1 then get tag ids for that tag.
ii. Find out image ids for retrieved tag ids.
iii. Fetch K-SimRank for all the image ids retrieved in step 3 with all other image nodes to get score. (K = 0.6).
iv. Sort and display images.
Output: images.

(b) Content based image retrieval: Input: output images from link based image retrieval.

1. User feedback (selection of an image as input query).
2. Perform basic feature extraction (SIFT).
3. Perform color extraction.
4. Perform Canny Edge detection.
5. Perform texture extraction.
6. Output of step 1 to 5 is feature vector.
7. Compare the feature vector that achieved in step 6 with the feature vectors of all output images of link based image retrieval phase.
8. Sort and display results.

Output: relevant images.

Advantages of Proposed Image Retrieval System Based on Integrated Algorithm (S-CBIR)

Proposed image retrieval system has user friendly GUI, and it is used to retrieve images in homogeneous image-rich information networks efficiently and accurately.

Proposed image retrieval system is integration of link and content based image retrieval techniques. It takes less time to retrieve relevant images from database.

Paper ID: E&TC16

1. Preprocessing step that developed and used in project helps to reduce processing time of system because SimRank and image features are precalculated. Proposed system allow user to add new images in current image database efficiently and retrieve those images in minimal time.

III. Experimental Setup

A. System Configuration

Different experiments are performed to evaluate the performance of the proposed system. All experiments are conducted using system configuration mentioned in Table 1.

B. Dataset

We have performed experiment on Flickr dataset API [4], it consist of images, tags, groups and other metadata information.

Hardware	Software
RAM Size: 4GB Processor: Intel ® core ™ i3-3110M CPU@2.40 GHz 2.40GHz Hard disk: 500GB	Operating System: Windows 8 Coding platform: java 1.8 IDE: Eclipse Ganymede Apache Tomcat 5.5 PostgreSQL 9.1 JUNG library jQuery Script library

Table 1. Hardware and software configuration

C. Statistics and Graphs

Gv Chart is jQuery plugin which renders interactive google charts from existing data tables to visualize data on website so it requires active internet connection to work.

IV. System Implementation

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.

The implementation stage involves careful planning, investigation of the existing system and its constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

A. Module Description

There are three different modules in proposed system Admin module, User module and Statistics module.

Module 1: Admin Module Admin can upload new images into the dataset through the ad- min module of proposed system, while uploading the new image user needs to enter the tags for that image. Admin needs to log in before uploading theimages.

Module 2: User Module User can search images through user module of proposed sys- tem. To search images first of all user needs to enter tag or keyword as input then user will get the search result for entered query i.e. tag or keyword after that user needs to give relevance feedback by clicking the one of the images from search results. Relevance feedback from user will be used as input for CBIR then user will get relevant and refined search result i.e. images.

Module 3: Statistics Module Statistics module is specially developed for Admin to draw graph for image count, elapsed time to get result and to com- pare the results of algorithms i.e. content based image retrieval (CEDD,SIFT etc.), Link based image retrieval (Simrank algorithm) and integrated algorithm (S-CBIR). Gv Chart is jQuery plugin which renders interactive google charts from existing data tables to visualize data on website is used in this module so it requires active internet connection to work.

V. Result Analysis

The performance of developed image retrieval system and proposed integrated algorithm (S-CBIR) is evaluated by different evaluation parameters such as precision, recall.

A. Performance Evaluation

After conducting experiment some results are obtained which are mentioned in thesis. Results are modeled into graph first graph drawn to display retrieved images count. X axis denotes dataset image count and Y axis denote number of images retrieved. We have searched images for a query or tag flower.

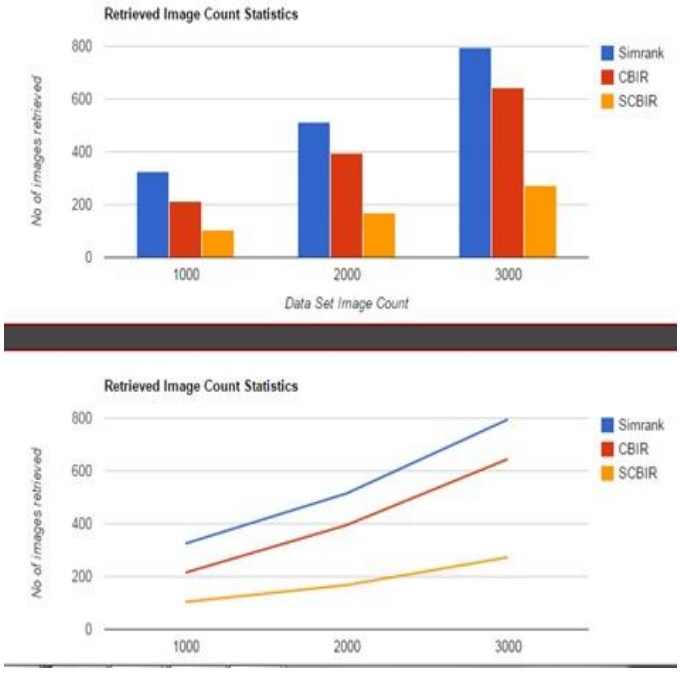

Fig. 2. Retrieved image count statistics for query "flower" [1].

Paper ID: E&TC16

The graph in Fig. 2 shows the result count of retrieved imagesby using SimRank, CBIR and proposed integrated algorithm (S-CBIR) when applied on datasets of different sizes. The results proves that proposed integrated algorithm (S-CBIR) produces refined search results [1].

The graph in Fig. 3 shows the elapsed time statistics to retrieve images by using SimRank, CBIR and proposed integrated (S-CBIR) algorithms when applied on dataset of different sizes. The results proves that the proposed integrated algorithm (S- CBIR) takes less time to retrieve relevant images as compared to SimRank and CBIR algorithms [1].

It is observed that if we increase dataset image count then the time required to search images for a given query is also increases.

The time taken by proposed algorithm to search query is still less than the time taken by the SimRank and CBIR. It is proved that the proposed algorithm is time efficient. An example to demonstrate how the proposed integrated algorithm (S-CBIR) retrieves the relevant images for semantic and visual appearances.

Fig. 3. Elapsed time statistics for query "flower" [1].

Fig. 4 shows first 10 search result for yellow rose query by using SimRank, CBIR and S-CBIR respectively. It is observed that for semantic and visual appearances S-CBIR retrieved most relevant search results.

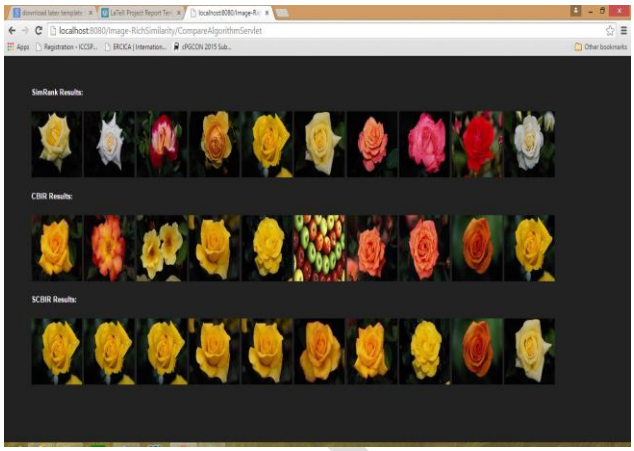

Fig. 4. Search results by using Simrank, CBIR and S-CBIR.

B. ACCURACY MEASURE:

Precision and Recall are most widely used as basic measures to evaluate search strategies [5].

Precision: it is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved. It is usually expressed as a percentage. Say, A=No. of relevant records retrieved C= No. of irrelevant records retrieved Precision = A/ (A + C) ∗100%.

Recall: it is the ratio of the number of relevant records retrieved to the total number of relevant records in the database. It is usually expressed as a percentage. Say, A=No. of relevant records retrieved B= No. of relevant records not retrieved Recall = A/(A + B)∗100.

The data mentioned in Table 2 represents precision and recall score of SimRank algorithm, CBIR algorithm and

Query	SimRank algorithm		CBIR		S-CBIR	
	*Pre.	*Rec.	*Pre.	*Rec.	*Pre.	*Rec.
Cake	80%	90%	6%	37%	87%	100%
Elephant	49.50 %	74%	10%	20%	100%	80%
Rose	53.17 %	85%	16%	36%	95%	90.90 %
Beach	88%	91.66 %	14%	8.10%	100%	74%
Horse and Baby	35%	93%	4%	34%	85%	100%

integrated algorithm (S-CBIR) for respective queries. From Fig. 5 it is observed that proposed integrated algorithm (S-CBIR) has high Precision which means that analgorithm returned substantially more relevant images than irrelevant. It also has high Recall which means an algorithm returned most

Paper ID: E&TC16

of the relevant images. It is proved that the proposed integrated algorithm (S-CBIR) gives high precision and recall score as compared to SimRank and CBIR algorithms.

Table 2. Precision and Recall Score.
*Pre. – Precision and *Rec. – Recall

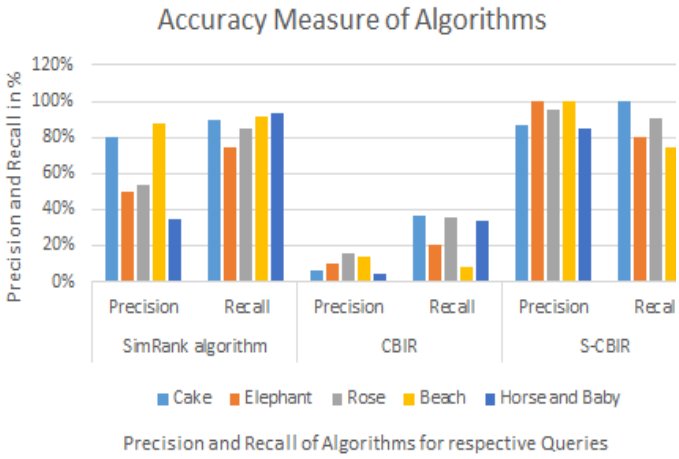

Fig. 5. Accuracy Measure of Algorithms.

VI. CONCLUSION

This paper presents an efficient way of searching relevant images by modeling a social image sharing website as image rich information network.

1. The S-CBIR algorithm which includes link based image retrieval and content based image retrieval is developed.
2. Both global and local features are considered while retrieving images by using S-CBIR algorithm.
3. The experiment is conducted on Flickr dataset and results are achieved which shows proposed algorithm improves the accuracy of search results than the link-based images retrieval and content based image retrieval techniques.
4. The proposed integrated algorithm gives high precision and recall score as compared to other algorithms and strategies like SimRank algorithm and CBIR. The proposed integrated algorithm (S-CBIR) improves accuracy of image retrieval system by retrieving semantically and visually similar images as output.

REFERENCES

[1] Vishal S. Kore, Pankaj Chandre and Parag P. Abhyankar, "Integrated algorithm (S-CBIR) for image retrieval in image-rich information networks", In proc. 4th IEEE international conference on communication and signal processing, pp. 1345-1350, April 2-4 2015, India.

[2] Vishal S. Kore, B. A. Tidke and Pankaj Chandre, "Survey of image retrieval techniques and algorithms for image-rich information techniques", International journal of computer applications, Volume 112-no.6, pp. 39-42, February 2015.

[3] X. Jin, J. Luo, J. Yu, G. Wang, D. Joshi, and J. Han, "Reinforced similarity integration in image-rich information networks", Knowledge and Data Engineering, IEEE Transactions on, vol. 25, pp. 448–460, Feb 2013.

[4] Jure Lescovec, Stanford Large Network Dataset collection, Stanford University, 2014. https://snap.stanford.edu

[5] R Jizba, Measuring Search Electiveness, 2007. https://www.creighton.edu