SCIENTIFIC OASIS

Spectrum of Operational Research

Journal homepage: www.sor-journal.org
ISSN: 3042-1470

A Framework for Extensive Content-Based Image Retrieval System Incorporating Relevance Feedback and Query Suggestion

Rabia Younas[1], Hafiz Burhan Ul Haq[1,*], Muhammad Daniyal Baig[2]

1 Department of Information Technology, Faculty of Computer Science, Lahore Garrison University, Lahore, Pakistan
2 Department of Cyber and Applied Computing, University of Gloucestershire, GL50 2RH Gloucestershire, United Kingdom

ARTICLE INFO

Article history:
Received 11 April 2024
Received in revised form 19 July 2024
Accepted 25 July 2024
Available online 27 July 2024

Keywords:
Relevance Feedback; Content Based Image Retrieval; Random Walker.

ABSTRACT

Over the last decade, there has been a huge increase in the number of images, in which visual information has become increasingly important. If images are to be utilized, they must be organized into databases, from which they can be searched based on different criteria. Content-Based Image Retrieval (CBIR) frameworks have become a mainstream subject of research due to their ability to retrieve images based on actual visual content rather than manually assigned textual descriptions. In the CBIR framework, the analysis and interpretation of image information in large and diverse image databases are evidently complex because there is no prior information about the size or scale of individual structures within the images to be analyzed. In CBIR, retrieval is based on visual image features, which can be extracted automatically from images with the help of human intervention through a technique known as relevance feedback. Nonetheless, efficient methods for differentiating the visual content of images are complex to develop. Therefore, rather than providing a perfect solution, CBIR systems must be able to exploit partial solutions to the problem of image understanding. In this paper, an implementation of a CBIR framework is introduced that not only attempts to efficiently capture user intent based on feedback but also provides query suggestions that help users formulate better queries to retrieve desired results efficiently. The proposed technique is straightforward to implement and scales efficiently to large datasets. Extensive experiments on diverse real-world datasets using image similarity measures have revealed the superiority of the proposed method over existing algorithms.

1. Introduction

Multimedia data in bulk amount is becoming available today, stirring to greater extent aspects of human activity: from online stores, sellers and resellers of store photograph to internet search engines and museums, TV channels storing their productions etc. There is an expanding interest in routines and innovations for mining and arranging rapidly and dependably these visual substances.

* *Corresponding author.*
E-mail address: burhanhashmi64@lgu.edu.pk

https://doi.org/10.31181/sor1120242

The traditional method was to append labels or keywords to each item in the multimedia databases and these annotations are used to perform searches to database. Manually annotating an extensive number of images is a complex task not only in relation to time and price but also related to quality of the annotations. There will be necessarily problems when several annotators contribute to the same database, because of differences in interpretation, stability of attention over time, errors caused by the repetitiveness of the task, etc.

Towards the close of the early part of 1980's content-based image retrieval (CBIR) was introduced to overcome the above shortcomings of the text-based retrieval. CBIR [1], aims at looking for images based on the real content of the image which is often represented features, color, texture and shape. Several systems have been proposed in recent years in the context of CBIR for still images and video sequences. Query by Browsing and Retrieval (QBIR) by [2,3] which allows the use of 'target media' both shots within videos, and images and query images to be reach by 'color, shape by texture and by sketch'. The other CBIR framework is the Virage Image Search Engine [4] where the image features are utilized to describe the database of images. These conventional approaches of CBIR are built on the computation of similarity between the query image introduced by the user and the other images available on the database using the Query by Example (QBE).

• The quality of the CBIR system can hardly be optimized within the frame of one query process.

• The hidden problem is the extracted visual features are too diverse and cannot be used to define the concept of the user's query.

One solution for reducing this gap is to divide the search session while searching more general image databases into several continuous retrieval rounds named as iterations and then the user give input with respect to the consequences of each retrieval round that is by marking images returned from the retrieval round as either relevant or irrelevant that is named as Relevance Feedback (RF). The engine gradually learns from this feedback, 'visual features' of the images the user is searching for in the present search period. CBIR framework using RF for the most part comprises of four segments, in particular 'user interface, offline learning, data storage and online learning'.

The user interface, permitting to perform the initial query generally either through sketching, 'by typing keywords' or 'by providing an example image'. Furthermore, permits the user to specify on returned images the relevance in the form of 'relevant-neutral- irrelevant'. In the 'offline learning stage' the extraction of image descriptors is carried out from all images in the database. Users' focus is towards 'high level semantic concepts' and not towards the 'low level image features' as accessible to the framework, it is extremely critical to select the most helpful 'image descriptors', to narrow down semantic gap. The 'data storage module' is in charge of loading and storing vital information and works as a virtual file system. It is key to utilize suitable 'multi-dimensional indexing' strategies for a proficient pursuit in 'multi-dimensional feature space', particularly considering that execution rapidly endures with expand in dimensionality [5,6]. The 'online learning module' carry out the real 'relevance feedback' with the collaboration of the user it learns about the 'image features' the user considers 'relevant' and which not and then accordingly by how well they adjust to the relevant features ranks the images.

In this paper, the idea proposed is a novel algorithm to Content Based Image Retrieval with relevance feedback and Query Suggestion, in which the relevant and irrelevant images that are marked by the user are taken as to find out the common images in the relevancy matrix. The common images relevancy value is incremented for those that are taken from relevant images. Likewise, from non-relevant images considered common images their relevancy values are decreased. Every unmarked image is finally ranked according to the relevancy values sorted in descending order in which images incremented will be in the top results unlike images which are decremented will be in lowest order. Lastly, the query suggestion component is incorporated in which after the raw query is carried out and obtained is the ranking array, and user is not satisfied with the relevance feedback results can simply select the image from query suggested to the user for better results. After the relevance feedback ranked results are shown then the relevant images and irrelevant images are taken as suggestion candidates' images. Lastly, these candidate queries are sorted in relation to their difficulty scores and then the image having the lowest difficulty is suggested to users. Our technique is straightforward to implement and scopes efficiently to huge datasets. Extensive experiments on diverse real datasets with image similarity measure have revealed the dominance of the proposed method over original algorithm of random walker.

2. Literature Review

It differs from the conventional similarity metrics in this technique a query image is provided, at first it induces the relationship among all the data points in the feature space using manifold ranking algorithm, and then it estimates the relevance among all the images in the database and the query eventually, based on pair-wise distance. If only positive samples are achieved in the relevance feedback, they are returned into the query set with a purpose of improving the retrieval result; if the samples of both labels can be retrieved, they allot the ranking scores of the positive and negative samples separately because of the asymmetry of these two samples of pictures. Furthermore, three active learning methods have been embedded in MRBIR to select images in each round of relevance feedback according to three different approaches to achieve the maximum increment in the rank result.

In [7], attempts made to enhance the effectiveness of RF by using images not labelled to reside in the given database. More specifically, this paper integrates active learning as well as semi supervised learning into the aspects of relevance feedback. Altogether, two primitive learners are trained from the marked data extracted from each one round of RF, namely images from the user feedback and user query. Each learner then feeds to the corresponding 'blank' database some images not designated by the other learner. They again sort the images in the database after the learners have re-prepared with the extra labeled information that has been given to them and the resulting groups of similar items are shuffled. This means that the images go through inspection and those, which are retrieved with high confidence are labeled to be part of the positive retrieval result, however, the ones reviewed with low confidence are placed in a pool and are used during the subsequent round of relevance feedback.

A prototype of hierarchical graphical has been presented in [8], that extracts regions and objects from the positive images in which the user is interested which are utilized to determine features from images for better understanding of the user feedback intention that enhance the image retrieval interaction. This prototype inserts into its edges; former details about image formation, user intention, statistical and utilizes a strategy named as a max-flow/min-cut for simultaneous segmentation of images that are positive feedback retrieved into user interested regions and uninterested regions. A paramount development of this prototype is the presentation of layer that fuses user intention and builds in diverse images a bridge of linking similar objects of visual appearance prototype nodes. This structural planning not just makes it conceivable to utilize all feedback images to acquire more powerful user intention earlier accordingly enhancing the object segmentation results and thus upgrading the retrieval execution, additionally significantly diminishes the unpredictability of the graph and the computational expense.

In [9] presents a novel retrieval prototype in which an image is annotated with keywords and then after the user's feedback alteration is done to the keywords, confidence level. Positive images are given the confidence MAXCONF; if images have query keywords otherwise it is given value as MINCONF. Then for the negative images, if the confidence becomes lower than MINCONF, the keyword is considered irrelevant to the image and thus removed from the image. The presented model not just contains given feedback about the images; additionally different images with 'visual features' like the features used to recognize the 'positive images' are subjected to confidence alteration. This takes into consideration adjustment of an extensive number of images with generally 'little feedback', eventually prompting quicker and more faultless retrieval results. A technique for consequently choosing the ideal 'weighted similarity function' for content-based searches has been proposed utilizing positive and negative images that are chosen throughout the relevance feedback process.

In 2010, Auer *et al.,* [10] describe CBIR framework that endeavors implicit RF throughout a search session named as 'Pinview'. It includes a few novel techniques that derive the aim of the user. Pinview learns from RF, for instance 'eye movements or clicks, and visual features' of images, finds a 'similarity metric' between images which rely upon the present motives of the user. Using specialized reinforcement learning algorithm retrieves images that adjust the exchange among exploring new images and exploiting the already inferred interests of the user. Pinview is incorporated to the CBIR system; PicSOM' with a specific end goal to apply it to real world image databases.

Bach *et al.,* [11] presented a novel technique focused on the algorithm namely random walker, presented in the perspective of interactive image segmentation to CBIR with relevance feedback. The scheme is treating for the random walker problem, the images as relevant and irrelevant marked by the user as 'seed 'nodes at each one feedback round. The ranking score for each one unmarked image is processed as the likelihood that random walker beginning from that image will achieve a relevant seed before experiencing irrelevant one. Some of the other techniques have been implemented by authors in [12-15]

2.1 Overview of Query Suggestion Algorithms

In this paper Yangxi Li *et al.,* [16] proposed a system, that can calculate queries ranking performance regarding their difficulties and then can also apply ranking optimization techniques to query difficulty guided image retrieval. The model appraises the query difficulty by extensively investigating the data existing in the query image, retrieval results, and target database. In this paper, introduced an algorithm namely a linear multiple features embedding that handles in the large-scale image retrieval situation, the high dimensional image features and multi model image features and it learns a linear transformation by incorporating a combined subspace from a small set of data in which the neighborhood data is conserved. The conversion could be viably and effectively used to deduce the subspace features of the recently watched information in the online setting. It demonstrated the centrality of query difficulty to image retrieval by applying it to guide the conduction of three retrieval refinement applications, i.e., re-ranking, federated search, and query suggestion.

In [17], addresses the problem of suggesting only the object's queries that contain relevant matches in the dataset. This requires to first discover accurate object's clusters in the dataset (as an offline process); and then to select the most relevant objects according to user's intent (as an on-line process). It introduces a new object's instances clustering framework based on a major contribution: a bipartite shared-neighbors clustering algorithm that is used to gather object's seeds discovered by matching adaptive and weighted sampling. Furthermore, introduced two practical ways to implement this new paradigm, i.e. "Mouse-over visual objects suggestion" and "Text-aware visual objects suggestion".

In short, multiple RF methods have been proposed over time; each having its own importance. Similarly for incorporation of Query Suggestion to RF system has been looked upon recently.

2.2 Problem Statement

This research is focused around answering the following questions:

i. How can the images be ranked and retrieved?
ii. How it will be possible in least number of feedback rounds the system to produce more relevant results?
iii. What will be the threshold to update or subtract the value in the relevancy matrix for relevant and irrelevant images?
iv. What will be the procedure to obtain the intersection of images from the relevant and irrelevant images?
v. What will be mechanism used for the Query Suggestion?
vi. Did the system produce after in-corporation of the component Query suggestion more precise results?

3. Methodology

The user initiates the system with a query image, which is automatically marked relevant because it is the standard of significance with which other images in the database will be compared. The standard method in the deficiency of any feedback is k-nearest neighbor retrieval which is modeled to retrieve the first set of K images. The user then labels these K returned images as either relevant or non-relevant and these labeling, together with the query that is marked as relevant, are conveyed back to the system as feedback.

The pseudo code of our technique is presented in Prototype-1. Our method requires in input the (relevancy matrix) R abstracting the CBIR problem, where each cell represents the similarity value between one image to other image, the user function in the form of array namely feedback, which collects the user's feedback, and a scope size K, which defines the number of images that should be presented to the user at each feedback iteration.

At lines 1-3 we set up the system by initializing all the arrays and variables, n image is equal to one means that query image iteration is started with only one image as a query image. So our method sort the row that contains the relevancy values of the query image with the threshold i.e. mean- relevancy value obtained from the query image row.

Prototype-1: The Proposed Relevance Feedback & Query Suggestion Algorithm

```
Require: Relevancy Matrix R(x, y), feedback array, scope size K,
{Initialization} {name of variables like revInd array  to zeros etc}
1:  nimage ← 1
2:  if relevancy value of the nimage is greater than threshold then
3:      Select those values as relevant images
4:  else
5:      Discard remaining values
6:      Sort relevant images in descending order
7:  end if
8:  Ranks ← get the first K relevant images to the query image
9: Present images indexed by Ranks to the user
10:  While user is not satisfied with Ranks do
11:     for all to nimages do
12:         Sort all nimages row from R.
13:           Place the sorted values in the array of size (nimages, x)
14:           Check if the first sorted row values are common to all the other array rows
15:         if answer is yes then
16:          Place those values in the revInd/ non-revInd array
17:         else
18:            Discard all the values.
19:         end if
20:           Update/Subtract the relevancy values of the revInd/non-revInd array by certain
              threshold
21:           Threshold value ← 1 - the maximum relevancy value of query image row - 0.1
22:       Repeat for the relevant marked images and irrelevant marked images.
23:     end for
24:  if user not satisfied with Ranks, then
25:       Select the relevant images as suggestion candidates to the query image
26:       Candidate queries are sorted according to their difficulty scores/relevancy distance
          values
27:       Suggestion query image is selected on the basis of lowest difficulty score or highest
          relevancy value
28:  end if
29: Ranks ← get K top ranked vertices by merging the revInd array and array obtained from
    the query image
30: Exhibit images indexed by Ranks to the user
31: end While
```

At line 4, we store in the scope S the K images that will be then presented to the user for gathering his/her feedback. Line 6 enters a loop of relevance feedback iterations, which will be suspended the moment the user is contented with the outcome. From line 7 to 8 we start a new relevance feedback. So, we consider the relevant and irrelevant images by taking the common images from the marked images by the user. So, common images which are taken from relevant images their relevancy values are incremented. Likewise, from non-relevant images considered common images relevancy values are decreased. Consider that at any moment the user can give the user feedback on each image in array through the user function. At line 9, query suggestion component is incorporated in which after the raw query is carried out and obtains the ranking array, then the relevant images and irrelevant images are taken as suggestion candidates images. Lastly, these candidate queries are sorted in relation to their difficulty scores and then the image having the lowest difficulty is suggested to users. At line 10 computation of the ranking vector is carried out as the solution of step 7. The K images that result in the ranking vector with higher score are then accumulated in the scope S and displayed to the user for new feedback iteration.

3.1 Experimental Setup

The experiments are run with MatLab Software on a machine operational with Intel Core i-5, 2.40 GHz CPUs and 4 GB RAM. There are three different datasets used for the purpose of experiments. The 1st dataset is a subgroup of the famous benchmark Corel dataset namely Wang dataset [16,18], comprises 1000 images clustered into 10 classes (100 images per class). The images classes are distinct that are subdivided into 10 classes namely for instance ‘monuments’, ‘food’, ‘Africa’, ‘beach’ etc such that it can be perceived that the queried image which the user requests to search in the other images from the category of these ten categories. While images’ belonging to the similar category reflects to the topic of that category, it is difficult benchmark database because it shows the maximum heterogeneity regarding visual similarity within one category. This database was created at Pennsylvania State University. The size of the images is 384 × 256 dimensions and from each category few instances are showed in Figure 14.

The 2nd collection is the Oliva Dataset [19] with 2649 images divided into 8 classes. Last, many of used images belong to the Caltech-256 Database [20] in which the final data set comprises a total of 4866 images allocated across 43 classes. The datasets are diversified, and they have distinct size; they also refer to different image fields, as it is shown in the Figures 1 and 2.

Fig. 1. Mountain Class

Fig. 2. Horse Class

3.2 Feature Extraction Methods:

The computation of image similarities is carried out on Color Features and Texture features for all datasets. The datasets on which experiments are carried out following are the summarize details of the features that have been considered:

1. *Color Moments:*

On every image of the dataset the calculation of feature extraction is done by taking skewness, mean and standard deviation of every channel of the HSV color space over the image and constructing a 9-dimensional feature vector matrix.

2. *Gray Level Co-Occurrence Matrix:*

All rgb / colored images are transformed into 16 gray-scale images. For every direction such as vertical, horizontal and two diagonal directions co-occurrence is computed. Then on each direction features are extracted such as: Contrast, Entropy, Second Angular Moment and Inverse Difference Moment. Then a 16-dimensional feature vector matrix is constructed.

3. *Color Histogram:*

The HSV color space is split into 32 subspaces such as 4 ranges of S and 8 of H. The thickness of every color in the image presents the values for construction of a 32-dimensional feature vector matrix.

2.3 Similarity Measure

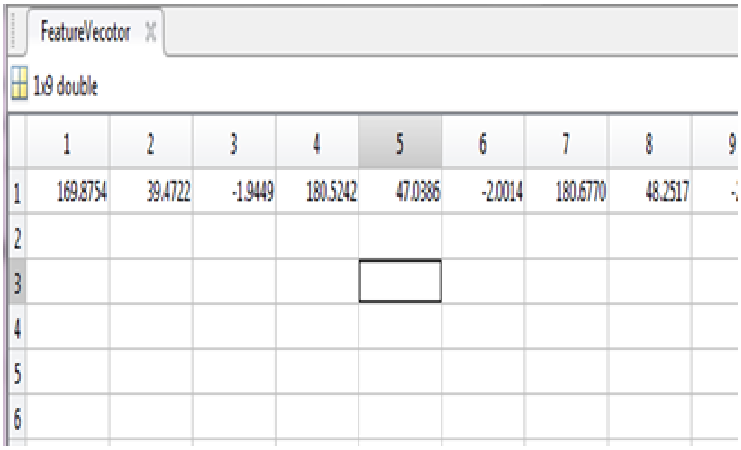
The distance measure that is used to calculate the feature vector is the Minkowski difference as written in (1):

$$D_M^{\rho}\left(\bar{a},\bar{b}\right) = \left(\sum_{i=1}^{n}|a_i, b_i|^{\rho}\right)^{\frac{1}{\rho}} \qquad (1)$$

Where a & b are 2 n-dimensional feature vectors matrix correlated to query and target images. ρ=1 is frequently used for color histogram and for texture.

2.4 Relevancy Matrix Calculation

Each cell in the relevancy matrix depicts the similarity of one image to another image. It is computed by taking Minkowski distance value between one image feature value and the next image feature.

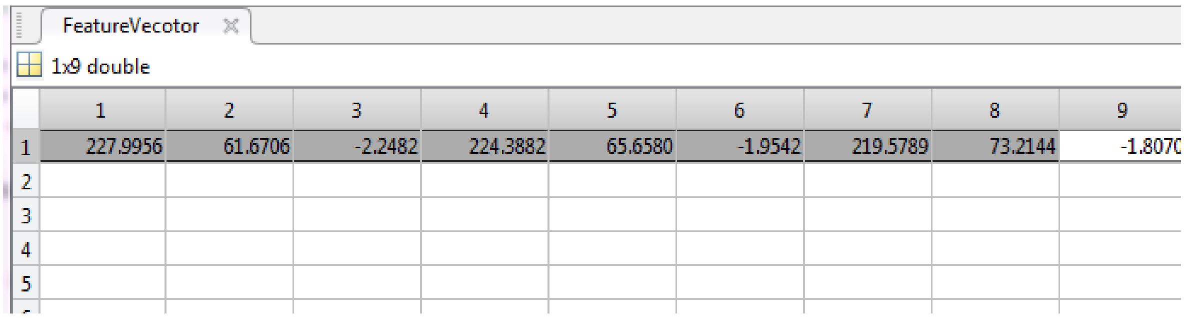


FeatureVecotor
1x9 double

	1	2	3	4	5	6	7	8	9
1	169.8754	39.4722	-1.9449	180.5242	47.0386	-2.0014	180.6770	48.2517	-
2									
3									
4									
5									
6									

Fig. 3. Color Moment Features for 1st Image

In Figure 3, features are extracted for color moments for the first image of the Dataset.



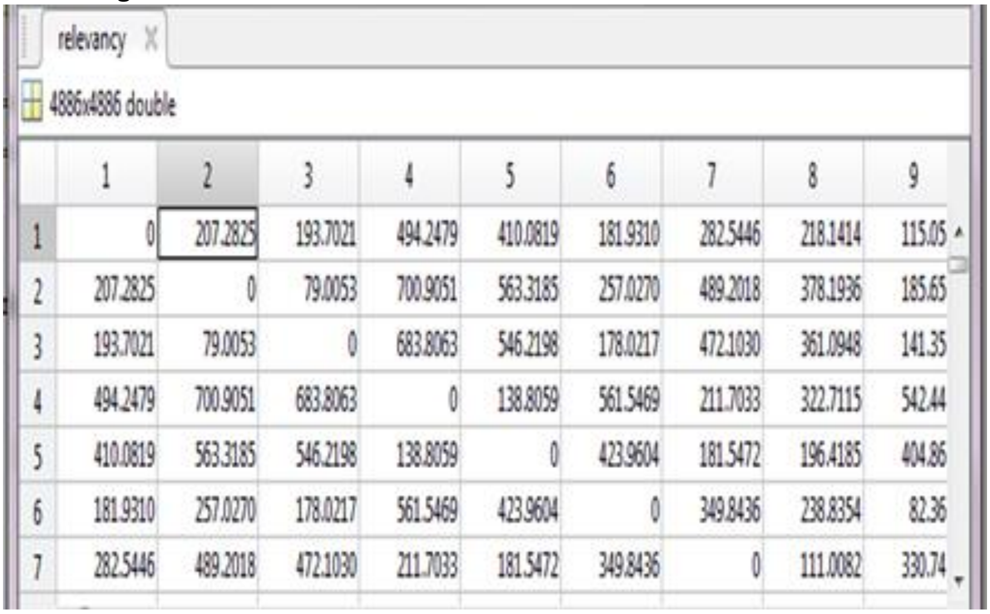
FeatureVecotor
1x9 double

	1	2	3	4	5	6	7	8	9
1	227.9956	61.6706	-2.2482	224.3882	65.6580	-1.9542	219.5789	73.2144	-1.807
2									
3									
4									
5									

Fig. 4. Color Moment Features for Second Image

In Figure 4, features are extracted for color moments for the 2nd image of the Dataset.In Figure 5, final matrix is computed by using Minkowski distance measure between one image features to other images features in the Dataset. For all diagonal values, 0 is inserted as images cannot have relevancy with each other. (1, 2) represents image 1 features similarity with image 2 based on features and so on.

Lastly, the normalization is carried out on feature vectors values so that every component must be in range i.e. 0 and 1. The normalization is carried out by dividing all distance values to the maximum gained distance.



relevancy

4886x4886 double

	1	2	3	4	5	6	7	8	9
1	0	207.2825	193.7021	494.2479	410.0819	181.9310	282.5446	218.1414	115.05
2	207.2825	0	79.0053	700.9051	563.3185	257.0270	489.2018	378.1936	185.65
3	193.7021	79.0053	0	683.8063	546.2198	178.0217	472.1030	361.0948	141.35
4	494.2479	700.9051	683.8063	0	138.8059	561.5469	211.7033	322.7115	542.44
5	410.0819	563.3185	546.2198	138.8059	0	423.9604	181.5472	196.4185	404.86
6	181.9310	257.0270	178.0217	561.5469	423.9604	0	349.8436	238.8354	82.36
7	282.5446	489.2018	472.1030	211.7033	181.5472	349.8436	0	111.0082	330.74

Fig. 5. Relevancy Matrix Calculated using Minkowski Measure

4. Results

We compared our algorithm against original Content Based Image Retrieval with Relevance Feedback using Random Walks method:We compared our algorithm against original Content Based Image Retrieval with Relevance Feedback using Random Walks method:

Relevance Feedback with Random Walk: In this paper, concerning the images screened by the user at each feedback interval, algorithm is to treat the relevant and irrelevant images as the ‘seed’ vertices for the random walker problem. In this case for each image that is not marked, the ranking score is the probability that one starting at the given unmarked image and wandering through the graph before coming across an irrelevant seed vertex will first arrive at a relevant seed vertex.

Regarding the datasets, the Wang dataset contains 1000 images the Oliva contains 2688 images and the last one is the Caltech custom that contains 4866 images, for these datasets, the methods are assessed in their performances (for every combination of feature and dataset) on 100 randomly sampled assumed queries, however for the Caltech custom the number of queries is reduced to 50 due to the large size of the dataset. Precision that is defined as in Eq.2, is evaluated on all the feedback intervals as the relative quality of the results:

$$\text{precision} = \frac{\text{number of the relevant retrieved images}}{\text{scope size}} \quad (2)$$

Also computed the mean of the precisions obtained in all the simulated queries at all the four perspectives. The size of scope introduced in experiments is given by 22, 30 and 40.

In Figure 6, shows an example of a query result that was generated through our proposed algorithm when used on the Oliva dataset with feature of Gray Level Co-Occurrence Matrix. The following results are provided which are observed at various feedback iterations. Relevant radio-buttons are Yes marked radio-buttons and non-relevant radio-buttons are marked with No. Even though the first k-NN search results in few relevant images, our approach yields good results. Precision even enhances and it goes to 100% in the 3rd round of relevance feedback.

(a) Query Image

(b) Initial results (17%)

1st iteration Relevance Feedback result:

(c) Results 1st RF round (70%)

(d) Results has been taken after 2nd RF round (87%)

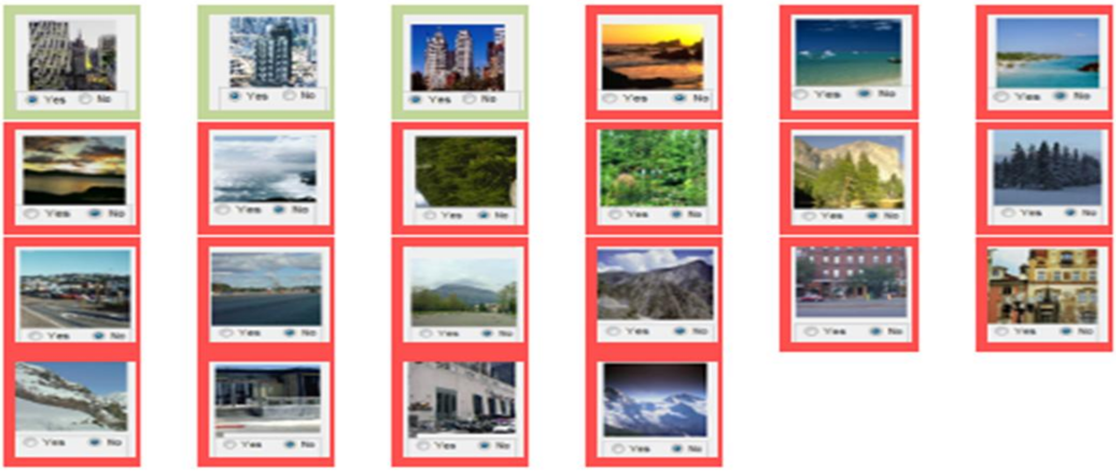
(e) Results has been taken 3rd RF round (100 %)

(f) Results after 4th RF round (100 %)

Fig. 6. Query example using our proposed algorithm on the Oliva dataset with GLCM feature: Immediately relevant images are categorized in the Yes marked radio-button, and non-relevant ones in No marked radio-button. a) The original query image that was used. b) The results that are gotten once the program is run and tested in the k-NN class. c-f) Results got after varying feedback rounds or sets of results got from the specific feedback rounds

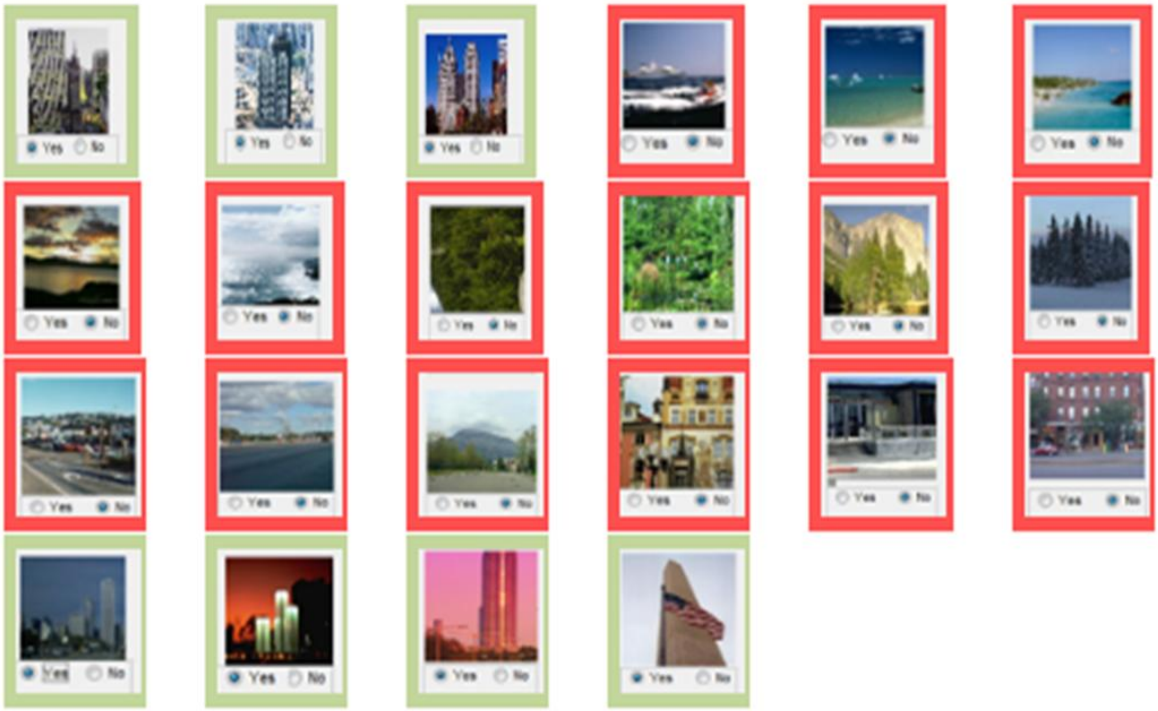
Based on the above explanation, Figure 7 shows an example of a query result achieved by the original random walker algorithm on the Oliva dataset when the Gray Level Co-Occurrence Matrix feature is used. It presents the results which, in fact, are achieved at different feedback iterations. Relevant radio-button marked as Yes and non-relevant marked as No. Approach described above gives accuracy of 0,17 on the first k-NN search. The precision increases to 100% at the 6th round of Refinement Relevant Feedback.

(a) Query Image



(b) Initial Result (17 %)

(c) Results has been taken after 1st RF round (34 %)



(d) Results has bee taken after 3rd RF round (65 %)

(e) Results has been taken after after 6th RF round (100 %)

Fig. 7. Query example using original random walker algorithm on the Oliva dataset with GLCM feature: Yes radio-button are considered as the images related to the task, and No radio-button are considered as the images are not the relevant to the task. a) The query image used. b) Outcomes derived after the resumption of K-NN algorithm computing. c-e) Outcomes at the end of different feedback loops.

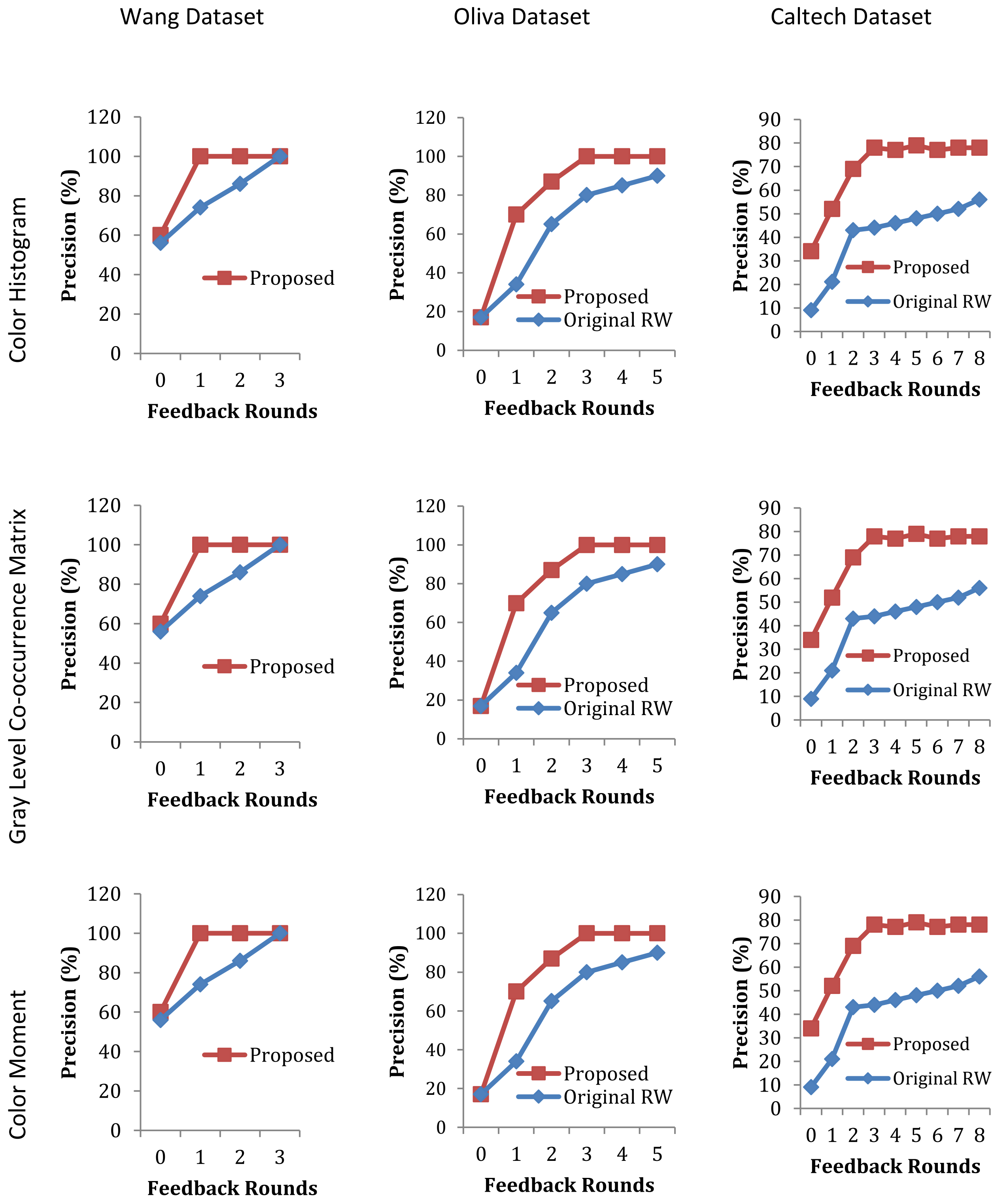


Fig. 8. The graphs demonstrate average precision of the compared CBIR approaches on the proposed algorithm to search on Wang, Oliva and Caltech datasets with features and a scope of 22

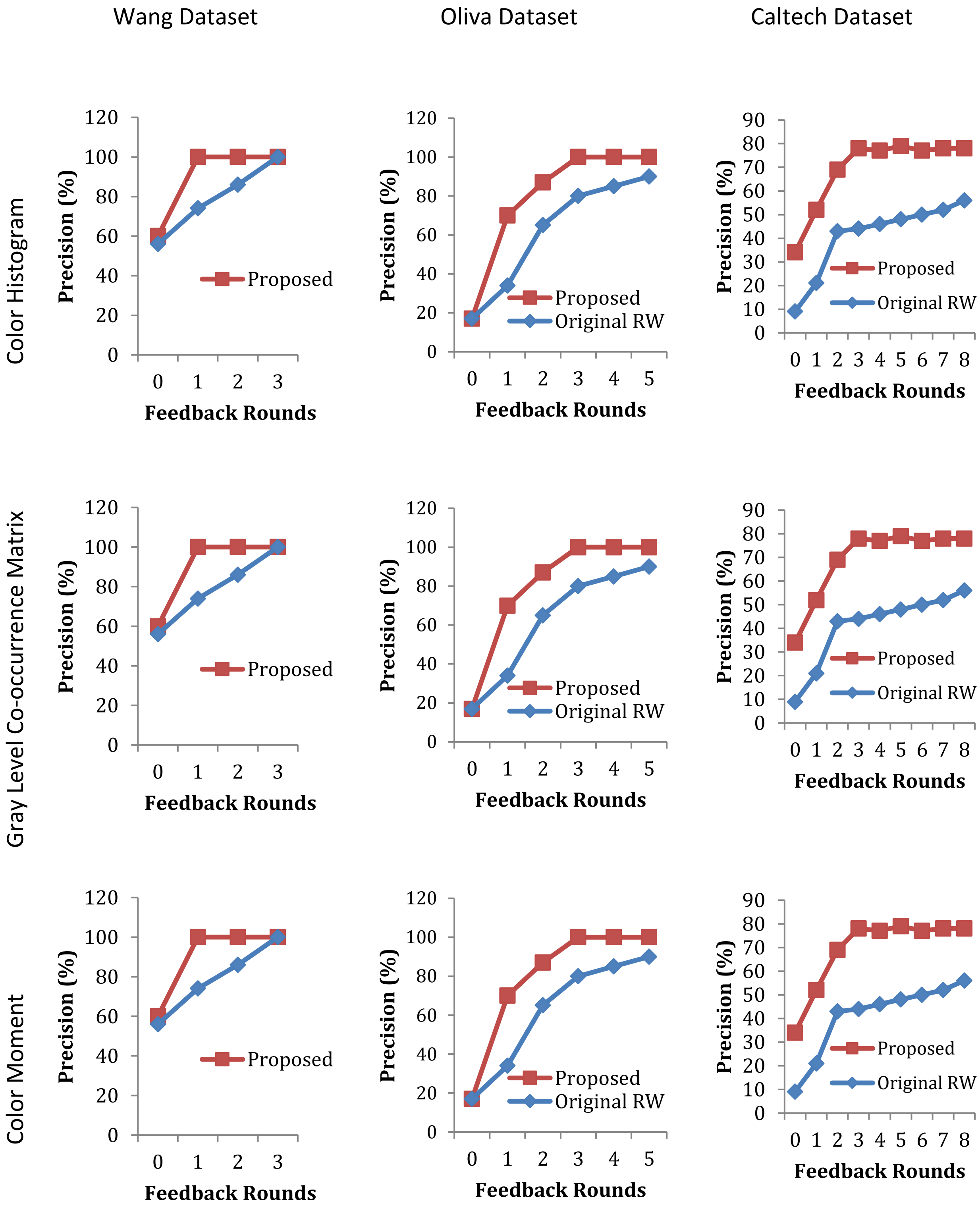


Fig. 9: depicting the average precision of various types of CBIR methodology over the Wang, Oliva and Caltech databases with features and scope of 30.

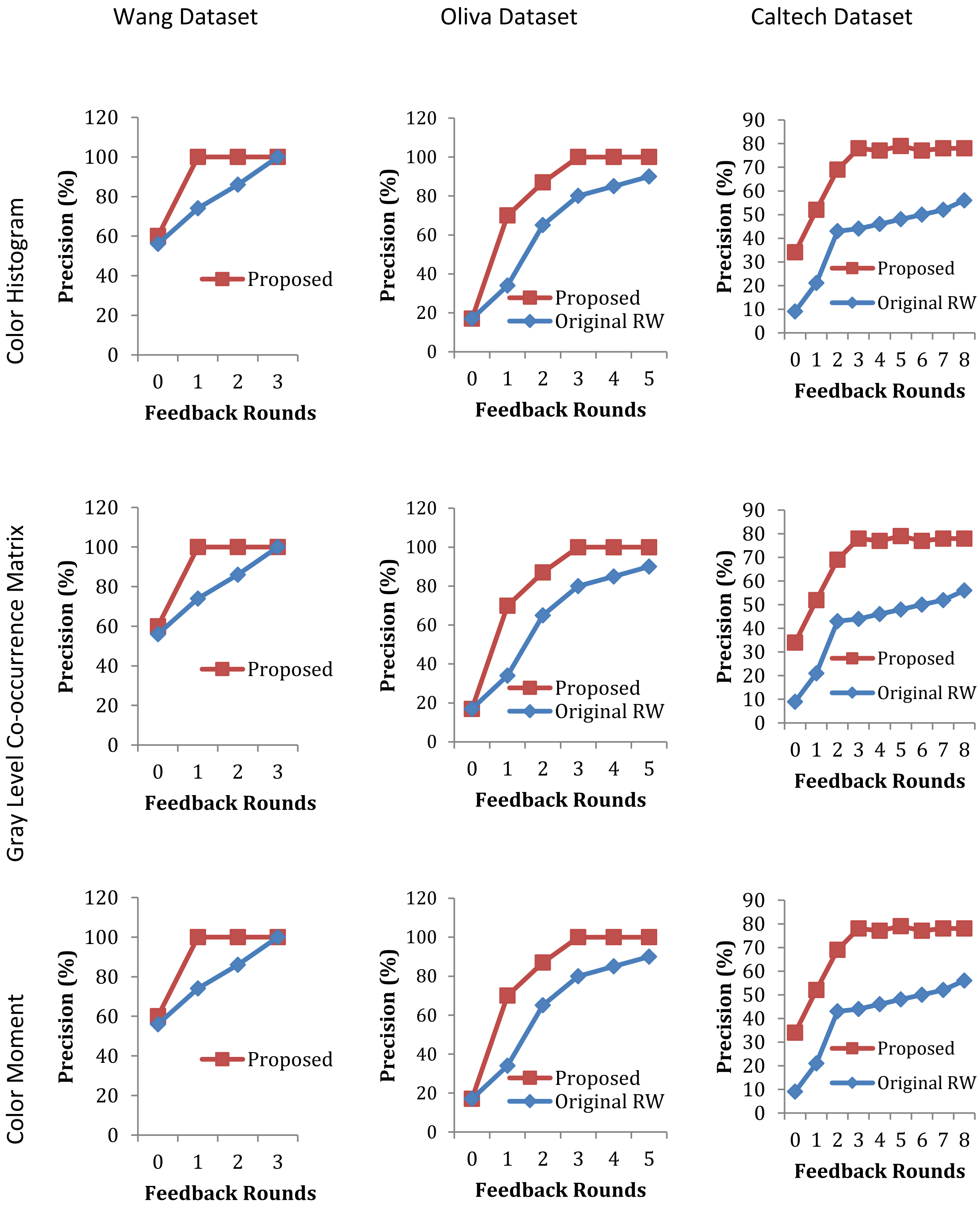


Fig. 10: Wang, Oliva and Caltech datasets average precision of the effectual CBIR approaches with variant carrying's and scope size of 40

In Figures 8, 9 and 10, some general outlines of results attained regarding precision for all the datasets with all the analyzed techniques and with all the features in scope sizes 22, 30 and 40, correspondingly. In general, based on the comparison with all the calculated plots, it can be stated that for all the combination of the features and datasets, as well as sizes of the scopes, our improved algorithm is superior to the initial algorithm.

The results are generally enhanced as one would expect on large image domains like in case of the Caltech dataset compared to the Wang dataset with less categories and rather narrow image domain. Of course, our approach, which yields outstanding results, does not exceed 80% of precision at the most. Moreover, the choice of the Image Features does influence the performances; that were used to portray total image. Specifically, regarding the visual representations in the Oliva Dataset, it can also be particularly noticed.

3.1 Implementation of the Component Query Suggestion

In Figure 11, Graphical User Interface of the component Query Suggestion is depicted. Formerly when users are disappointed with the preliminary ranking list, it is not convenient for them to find & upload a renewed image query, whereas choosing a substitution from suggested query list is more recommendable.

(a)Query Image

(b) Results after 1st RF iteration

(c) Suggested Image from the relevant marked Image results

(d) Results retrieved after using Suggested image

Fig. 11. Shows the overall working of the Query Suggestion Results

4. Conclusions

In this research paper, proposed is a novel algorithm to Content Based Image Retrieval with relevance feedback and Query Suggestion, in which the relevant and irrelevant images that are marked by the user are taken as to find out the common images in the relevancy matrix. The common images relevancy value is incremented for those that are taken from relevant images. Likewise, from non-relevant images considered common images their relevancy values are decreased. Every unmarked image is finally ranked according to the relevancy values sorted in descending order in which images incremented will be in the top results unlike images which are decremented will be in lowest order. Lastly, the query suggestion component is incorporated in which after the raw query is carried out and obtained is the ranking array, and user is not satisfied with the relevance feedback results can simply select the image from query suggested to the user for better results. After the relevance feedback ranked results are shown then the relevant images and irrelevant images are taken as suggestion candidates' images. Lastly, these candidate queries are sorted in relation to their difficulty scores and then the image having the lowest difficulty is suggested to users. Our technique is straightforward to implement and scopes efficiently to huge datasets. Extensive experiments on diverse real datasets with image similarity measure have revealed the dominance of the proposed method over original algorithm of random walker.

Acknowledgement

This research was not funded by any grant.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. International journal of computer vision, 42, 145-175. https://doi.org/10.1023/A:1011139631724

[2] Hamzaoui, A., Letessier, P., Joly, A., Buisson, O., & Boujemaa, N. (2014). Object-based visual query suggestion. Multimedia tools and applications, 68, 429-454. https://doi.org/10.1007/s11042-012-1340-5
[3] Böhm, C., Berchtold, S., & Keim, D. A. (2001). Searching in high-dimensional spaces: Index structures for improving the performance of multimedia databases. ACM Computing Surveys (CSUR), 33(3), 322-373. https://doi.org/10.1145/502807.502809
[4] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., & Yanker, P. (1995). Query by image and video content: The QBIC system. computer, 28(9), 23-32. https://doi.org/10.1109/2.410146
[5] Griffin, G., Holub, A., & Perona, P. (2007). Caltech-256 object category dataset (Vol. 10). Pasadena: Technical Report 7694, California Institute of Technology.
[6] Wang, J. Z., Li, J., & Wiederhold, G. (2001). SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. IEEE Transactions on pattern analysis and machine intelligence, 23(9), 947-963. https://doi.org/10.1109/34.955109
[7] Guan, J., & Qiu, G. (2007). Modeling user feedback using a hierarchical graphical model for interactive image retrieval. In Advances in Multimedia Information Processing–PCM 2007: 8th Pacific Rim Conference on Multimedia, Hong Kong, China, December 11-14, 2007. Proceedings 8 (pp. 18-29). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-77255-2_3
[8] He, J., Li, M., Zhang, H. J., Tong, H., & Zhang, C. (2004). Manifold-ranking based image retrieval. In Proceedings of the 12th annual ACM international conference on Multimedia (pp. 9-16). https://doi.org/10.1109/TIP.2006.877491
[9] Niblack, C. W., Barber, R., Equitz, W., Flickner, M. D., Glasman, E. H., Petkovic, D., Yanker, P., Faloutsos, C., & Taubin, G. (1993). QBIC project: querying images by content, using color, texture, and shape. In Storage and retrieval for image and video databases (Vol. 1908, pp. 173-187). Spie. https://doi.org/10.1117/12.143648
[10] Auer, P., Hussain, Z., Kaski, S., Klami, A., Kujala, J., Laaksonen, J., Leung, A. P., Pasupa, K., & Shawe-Taylor, J. (2010). Pinview: Implicit feedback in content-based image retrieval. In Proceedings of the First Workshop on Applications of Pattern Analysis (pp. 51-57). PMLR.
[11] Bach, J. R., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R. C., & Shu, C. F. (1996). Virage image search engine: an open framework for image management. In Storage and retrieval for still image and video databases IV (Vol. 2670, pp. 76-87). SPIE. https://doi.org/10.1117/12.234785
[12] Bulo, S. R., Rabbi, M., & Pelillo, M. (2011). Content-based image retrieval with relevance feedback using random walks. Pattern Recognition, 44(9), 2109-2122. https://doi.org/10.1016/j.patcog.2011.03.016
[13] Haq, H. B. U., & Saqlain, M. (2023). Iris detection for attendance monitoring in educational institutes amidst a pandemic: A machine learning approach. Journal of industrial intelligence, 1(3), 136-147. https://doi.org/10.56578 /jii010301
[14] Baig, D., Akram, W., Haq, H. B. U., & Asif, M. (2022). Cloud gaming approach to learn programming concepts. Artificial Intelligence and Applications. https://doi.org/10.47852/bonviewAIA32021378
[15] Nawaz, S. (2023). Cloud computing services and security challenges: A review. Lahore Garrison University Research Journal of Computer Science and Information Technology, 7(2), 17-28. https://doi.org/10.54692/lgurjcsit.2023.0702459
[16] Li, Y., Geng, B., Tao, D., Zha, Z. J., Yang, L., & Xu, C. (2012). Difficulty guided image retrieval using linear multiple feature embedding. IEEE Transactions on Multimedia, 14(6), 1618-1630. https://doi.org/10.1109/TMM.2012.2199292
[17] Kim, W. C., Song, J. Y., Kim, S. W., & Park, S. (2008). Image retrieval model based on weighted visual features determined by relevance feedback. Information Sciences, 178(22), 4301-4313. https://doi.org/10.1016/j.ins.2008.06.025
[18] Ul Haq, H. B., & Saqlain, M. (2023). An implementation of effective machine learning approaches to perform sybil attack detection (SAD) in IoT network. Theoretical and applied computational intelligence, 1(1), 1-14. https://doi.org/10.31181/taci1120232
[19] Zhang, J. (2011). Robust content-based image retrieval of multi-example queries. http://ro.uow.edu.au/theses/3222
[20] Zhou, Z. H., Chen, K. J., & Dai, H. B. (2006). Enhancing relevance feedback in image retrieval using unlabeled data. ACM Transactions on Information Systems (TOIS), 24(2), 219-244. https://doi.org/10.1145/1148020.1148023