# A Survey Paper on To Retrieve the Relevant Images according to Re-Ranking of Images with Click based Similarity and Typicality

**Deepak B. Waghchaure[1], K. N. Shedge[2]**

[1]Student, Master of Engineering, Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Chincholi, Sinner

[2]Assistant Professor, Department of Computer Engineering, Sir Visvesvaraya Institute of Technology, Chincholi, Sinner

**Abstract:** ***Searching performance can be improved with image re-ranking. Search engines provide the results according to top ranked set of images. The result set is processed and rearranged by using some specific features. Search engines like Google and Bing uses this concept of re-ranking. In these search engines initially textual query is processed and then pool of images are retrieved for the text query. A query image is selected from the pool and according to visual features of the query image, the pool of images are re-ranked. To get semantic signatures, the visual features are projected with semantic space specified by query keyword. Proposed approach enhances matching efficiency of query specific semantic signatures. Similarities of visual features do not well correlate with image semantics. The visual features of images are projected into their related semantic spaces to get semantic signatures. At the online stage, images are re-ranked by comparing their semantic signatures obtained from the semantic space specified by the query keyword. Thus proposed image re-ranking framework using query specific semantic signature improves both the accuracy and the efficiency of retrieval. A semantic signature is a list associated with an administered metadata object.***

**Keywords:** Reference classes, re-ranking, Visual Features, semantic space and signature

## 1. Introduction

Nowadays search engines are used worldwide, as a result of it provides easier way to collect information associated with any topic at any time. Mostly WEB SCALE image search engines use keywords as queries and rely on surrounding text to search images. There are ambiguities in result because of query keywords, and it is hard for users to exactly describe the visual content of target images only using keywords. For example, using "apple" as a query keyword, the retrieved images belong to different categories such as "red apple", “apple logo", and "apple laptop". In order to solve the ambiguity, content-based image retrieval with relevance feedback is widely used. One drawback with all current approaches is that the reliance on visual similarity for judgment semantic similarity, which can be problematic because of the semantic gap between low-level content and higher-level ideas. It requires users to select multiple relevant and irrelevant images, from which visual similarity metrics are learned through online training. Images are re-ranked based on the learned visual similarities.

Given a query keyword is input by a user to system, a pool of images relevant to the query keyword is retrieved by the search engine according to a stored word image index file. By asking the user to select a query image, which reflects the user's search purpose, from the pool, the remaining images in the pool are re ranked based on their visual similarities with the query image. The word image index file and visual features of images are initially computed offline and then stored.

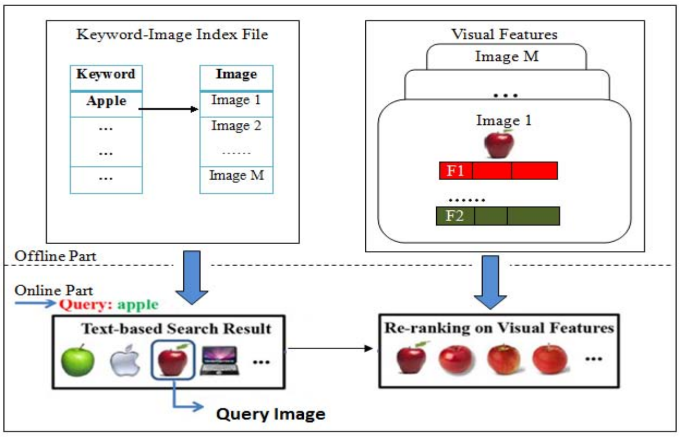



**Figure 1:** Conventional Image Reranking framework using Visual features

Once a client inputs a textual query (e.g., “Disney”) and starts to look around the text based research result, user has a search intention, which could be a particular goal image or images in a particular category (e.g., images of Cinderella Castle). When the user finds a candidate image similar to the target image or belonging to the category of interest, the re-ranking purpose is used by choosing that candidate image as a query image.

Following criteria should be considered in search scenarios:
1) Expect the top ranked images are in the same semantic category as the query image.
2) If the search purpose is to find a goal image, they suppose with the intention of images visually similar to the query image should have higher ranks.
3) If the search purpose is to look through images of a particular semantic category, diversity of candidate images

can be considered.

"As per literature there is gap between low-level visual features and semantic categories to effectively overcome that gap proposed query-specific semantic signatures and also make image matching more consistent with visual perception."

## 2. Related Works

Abundant literature has been dedicated to the web based searching and tremendous progress has been made ranging from efficient and scalable algorithm for different datasets. We believe that mining research has significantly become larger the scope of data analysis and will have deep impact on mining methodologies and applications in the future.

- **[1] Dacheng Tao:** Presented new algorithms to get Relevance feedback schemes based on support vector machines (SVM). These relevance feedbacks have been widely used in content-based image retrieval (CBIR). But, if the number of labeled positive samples is small then the performance of SVM-based relevance feedback is often poor. To advance the SVM performance, this work uses bagging and a random subspace method which shows extra efficiency than conventional classifier. In [2] Xiaoou Tang presented a novel Internet image search approach. The pool of images are retrieved using query keyword, user have to click on one query image with the minimum effort and images from a pool retrieved images are re-ranked based on both visual and textual content. Key contribution is to capture the users' search intention from this one-click query image are (1)The query image is categorizing into one of the predefined adaptive weight categories.(2)Keyword expansion based on the visual content of the query image selected by the user and through image clustering. (3)Expanded keywords are used to rise the image pool to contain additional relevant images. Tentative evaluation shows that the precision of top ranked images significantly improved.

- **W. H. Hsu, L. S. Kennedy(2007):**
The another approach image search re-ranking on real time Google images is described by G. Dupret and C.Liao Generally search engines rely almost purely on surrounding text features. Text based searching leads to ambiguity and noisy results. This paper uses adaptive visual similarity to re-rank the text based search results. Initially query image is sort out into one of several predefined target category, and a precise similarity measure is used inside each category to combine image features for re-ranking based on the query image. There is lot of work in field of feature extraction of images is described in [4][5][6]. To filter out the most possible non relevant images using deep contexts, in which the extra information that is not limited in the current search results. The deep contexts for each image are set of images that are returned by searches using the queries formed by the textual context of this image. The irrelevance score is checked by comparing the popularity of this image in the current search results and the deep contexts. Then the irrelevance scores are propagating to the images whose useful textual context is missed. This work combines two schemes together to reach a Markov random field, which is successfully solve by graph cuts. The key is that this scheme does not depend on the assumption that relevant images are visually aggregated among top results and is based on the observation that an outliers under the current query is likely to be more popular under some other query. After that, graph re-ranking is performed over filtered results to reorder them. Proposed method achieves important improvement over previous approaches.

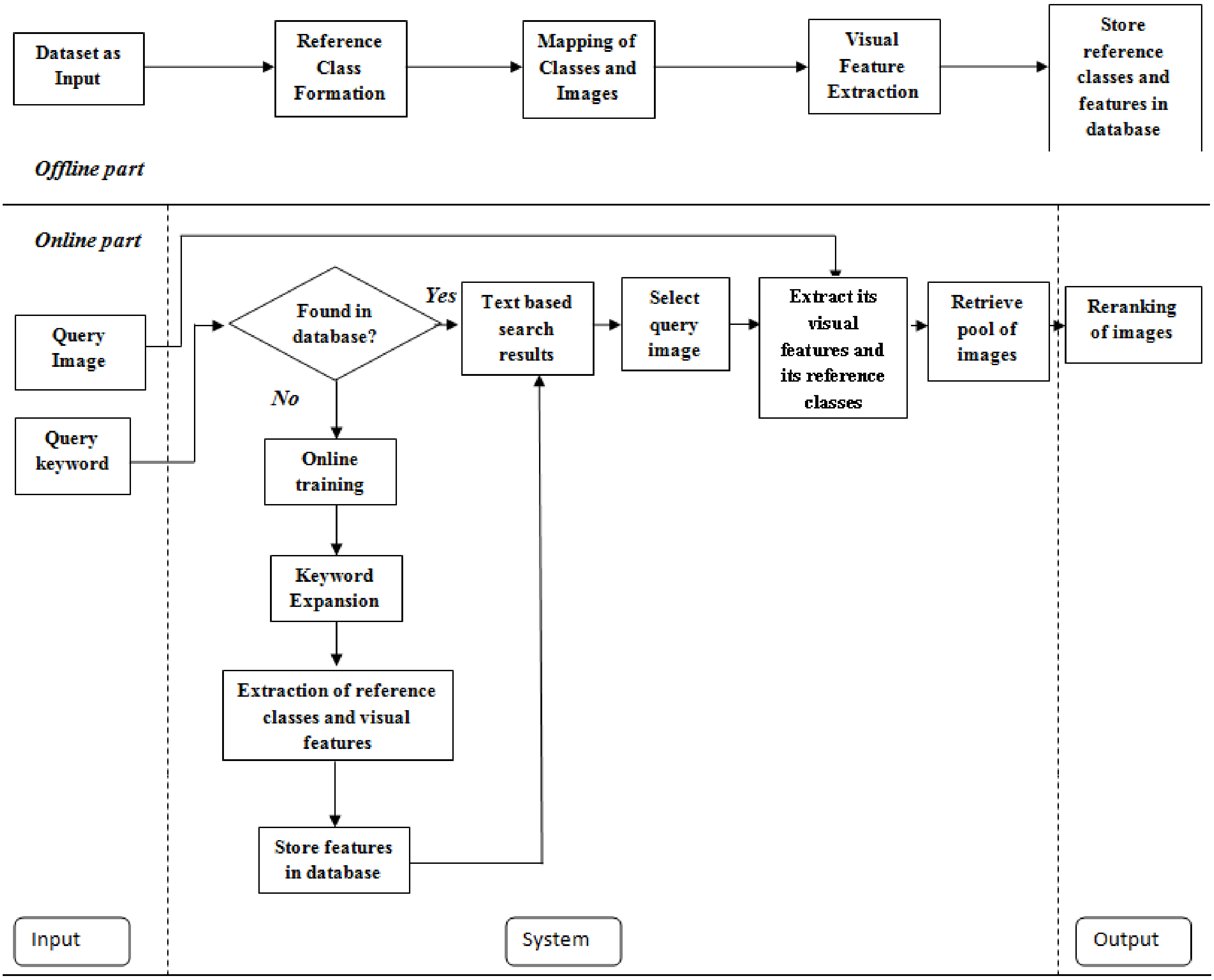
- **R. Yan, A. Hauptmann (2003):** An In [5] Qi Yin presented a new model, called "Associate-Predict" (AP) model, to overcome issues related to finding similar faces. The associate-predict representation is built on an extra standard uniqueness of data set, in which each identity contains multiple images with large intra-personal variation. By considering two faces under significantly different settings (e.g., non-frontal and frontal), first "associate" one input face with alike identities from the generic identity date set. Via the associated faces, "predict" the appearance of one input face under the setting of another input face. The two proposed prediction methods as "appearance-prediction" and "likelihood-prediction". By leveraging an extra data set ("Memory") and the "associate-predict" model, the intra-personal variation can be effectively handled. Final model can substantially improve the performance of most existing face recognition methods

- **M. Wang, H. Li, D. Tao, K. Lu** : Antonio Torralba presented technique to recognize where it is and what the main objects in the scene are. In this paper a context based vision system for place and object recognition. The goal is to spot known locations (e.g., workplace 610, meeting room 941, Main Street), to sort out new environments (office, passage, road) and to use that information to give appropriate priors for object recognition (e.g., tables are more likely in an office than a street). This paper considered a low-dimensional global image representation that provides relevant information for place recognition and categorization, and show how such contextual information introduces strong priors that simplify object recognition. The algorithm has been included into a mobile system that provides real-time feedback to the user.

## 3. Exiting techniques of Keyword Extraction

In proposed system, initially offline training for given dataset is performed. Offline part accepts dataset as input and dataset is processed to retrieve reference classes. Once reference classes are formed mapping of images with those reference classes is performed. Finally visual features are extracted and stored in database that is used in online part to retrieve relevant images to users query.

In online part user can send query either in form of text or directly as image. If user is sending in the form of image then visual features of that query image are extracted, after that pool of images related to query are retrieved and reranking is performed to extract relevant result. If user is sending query in the form of text then initially searching of query keyword is performed on database. If searching is successful then text based results are retrieved and from that result user have to select one query image that query image is processed in similar way as that of directly query as image. If query keyword is not found in database then online training is performed to get relevant result for user query.




## 4. Implementation Details

Algorithm for Face Detection:
Steps:
1. Input: Input image I1
2. Initialized: opencv library. Set no of face nf=0;
3. Create HaarCascade classifier object face-cascade. Load train xml file into object.
4. Create HaarCascade classifier object eye-cascade. Load train xml file into object.
5. Get input image I1.
6. Convert it into Gray scale Gi.
7. Detect HaarCascade of Gi.
8. Detect faces and increment nf.
9. Output: Return detected faces nf.

**Algorithm for picture analysis or get frequently used colors:**
**Steps:**
1. Input: Input image I1
2. Initialized: Create empty dictionary dctColorIncidence <int, int> ();
3. for each pixel Pi of I1 do
   Get pixel value.
4. if(dctColorIncidence.Keys contains(pixelValue)) then
5. Increment index of that pixel value in dictionary
6. else
   Assign 1 to at index pointed by pixel value in dictionary.
7. end for
8. Sort dictionary
9. Output: Return first element of dictionary as mostUsedcolor.

**Re-ranking of images:**
Steps:
1. Input: get Query image Qi
2. Extract faces from Qi
3. Extract faces visual feature of Qi
4. Get references class of Qi.
5. Retrieve pool of images of that reference class if faces=0;
6. Else
7. Retrieve images of that reference class and no of faces=extracted faces;
8. Match visual features of retrieve images with Qi
9. Output: return web rerank images

## 5. Mathematical Model of Proposed System

Set theory:
Let, W= {E, V}
Where,
W= Set representing edges and vertices.
E=Represents set of input parameters
V=Represents functions to be performed of E
E= {I1,I2,I3,RC,IR,VF1,VF2,SI,PI}
Where,
I1= Input as dataset for offline processing.
I1={I11, I12 ...I1j}
I2= Input as Keyword.
I3= Input as Image.
RC= Set of reference classes
RC= {RC1, RC2,… RCn}
IR= mapping of dataset images and RC
IR= {IR1,IR2,…IRn}
IR1={IR11,IR12,….IR1m}
IR2={IR21,IR22,….IR2m}
IRn={IRn1,IRn2,….IRnm)
VF1=Visual features of dataset images
VF2=Visual features of Query image
PI=pool of images as a result of F5
V= {F1, F2, F3, F4, F5, F6, F7}
Where,
F1== Function for reference class (RC) formation from the input dataset I1having j images.
F2= Mapping of reference classes and images and generate set IR
F3= Visual Feature Extraction i.e. VF1
F4= {P1,P2}
P1=Select query image and extract visual faetures.
P2=Perform online training if result is not available in database.
F5= Comparison features of query image and features of images in database.
F6=To process I3
F7= Generate final reranked output.

## 6. Conclusion

In re-ranking scheme, click information is fully adopted to guide the image similarity learning and image typicality learning. we have studied the issue of leveraging click through data to reduce the intent gap of image search. We propose a novel image search re-ranking approach, named spectral clustering re-ranking with click-based similarity and typicality (SCCST). In this proposed With the detection of click-based triplets, we present a novel image similarity measurement, named click-based multi-feature similarity learning (CMSL), which integrates multiple kernel learning into metric learning to learn similarity measure for each feature in a unified space. Based on the learnt similarity measure, SCCST performs spectral clustering to group visually and semantically similar
images into same clusters. The final re-rank list is obtained by calculating clusters typicality and within-clusters image typicality in descending order. Experiments conducted in this paper have demonstrated the availability and superiority of our proposed SCCST compared with several existing reranking approaches. Metric adaptive fusion weights are not considered in SCCST and CMSL due to the optimization dificulty, and we will figure this out in our future work..

## 7. Acknowledgement

I would like to express my profound gratitude and deep regard to my Project Guide Prof. K.N.Shedge, for his exemplary counsel, valuable feedback and constant fillip throughout the duration of the project. His suggestions were of immense help throughout my project work. Working under him was an extremely knowledgeable experience for me.

## References

[1] B. Wu, T. Mei, W.-H. Cheng, and Y. Zhang, “Unfolding temporal dynamics: Predicting social media popularity using multi-scale temporal decomposition,” in Thirtieth AAAI Conference on Artificial Intelligence, 2016.
[2] T. Mei, Y. Rui, S. Li, and Q. Tian, “Multimedia search reranking: A literature survey,” ACM Computing Surveys (CSUR), vol. 46, no. 3, p. 38, 2014.
[3] W. H. Hsu, L. S. Kennedy, and S.-F. Chang, “Reranking methods for visual search,” MultiMedia, IEEE, vol. 14, no. 3, pp. 14–22, 2007.
[4] W. H. Hsu, L. S. Kennedy, and S.-F. Chang, “Video search reranking through random walk over document-level context graph,” in Proceedings of the 15th international conference on Multimedia. ACM, 2007, pp. 971–980.
[5] R. Yan, A. Hauptmann, and R. Jin, “Multimedia search with pseudorelevance feedback,” in Image and Video Retrieval. Springer, 2003, pp. 238–247.
[6] M. Wang, H. Li, D. Tao, K. Lu, and X. Wu, “Multimodal graph-based reranking for web image search,” Image Processing, IEEE Transactions on, vol. 21, no. 11, pp. 4649–4661, 2012.
[7] T. Yao, C.-W. Ngo, and T. Mei, “Circular reranking for visual search,” Image Processing, IEEE Transactions on, vol. 22, no. 4, pp. 1644–1655, 2013
[8] J. Cui, F. Wen, and X. Tang, “Real time google and live image search re-ranking,” in Proceedings of the 16th ACM international conference on Multimedia. ACM, 2008, pp. 729–732
[9] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, “Relevance feedback: A power tool for interactive content-based image retrieval,” Circuits and Systems for Video Technology, IEEE Transactions on, vol. 8, no. 5, pp. 644–655, 1998..
[10] X. Tian, D. Tao, X.-S. Hua, and X. Wu, “Active reranking for web image search,” Image Processing, IEEE Transactions on, vol. 19, no. 3, pp. 805–820, 2010.
[11] X. S. Zhou and T. S. Huang, “Relevance feedback in image retrieval: A comprehensive review,” Multimedia systems, vol. 8, no. 6, pp. 536–544, 2003.
[12] B. Carterette and R. Jones, “Evaluating search engines by modeling the relationship between relevance and clicks,” in Advances in Neural Information Processing Systems, 2008, pp. 217–224.

[13] G. Dupret and C. Liao, "A model to estimate intrinsic document relevance from the clickthrough logs of a web search engine," in Proceedings of the third ACM international conference on Web search and data mining. ACM, 2010, pp. 181–190.
[14] G. E. Dupret and B. Piwowarski, "A user browsing model to predict search engine click data from past observations," in Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2008, pp. 331–338.
[15] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay, "Accurately interpreting clickthrough data as implicit feedback," in Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval. Acm, 2005, pp. 154–161..
[16] Y. Zhang, X. Yang, and T. Mei, "Image search reranking with querydependent click-based relevance feedback," Image Processing, IEEE Transactions on, vol. 23, no. 10, pp. 4448–4459, 2014.
[17] T. Mei, Y. Wang, X.-S. Hua, S. Gong, and S. Li, "Coherent image annotation by learning semantic distance," in Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008, pp. 1–8.

## Author Profile

**Deepak Waghchaure** received the B.E. degree in Computer Engineering from K.K.Wagh Enggineering College, Nashik in 2013. He is currently pursuing Master's Degree in Computer.

**K. N. Shedge** is working as Assistant Professor in Sir Visvesvaraya Institute of Technology.