

Capturing User Intention to Enhance Image Retrieval and Re-Ranking Efficiency Using One Click Image Search

Dange B J^{a,1}, Yadav S K^b, Kshirsagar D B^b

^{a,b}*Dept of CSE, Shri Jagdishprasad Jhabarmal Tibrewala University, India*

Abstract. A Novel data fusion technique to support text-based and content-based image retrieval combining different heterogeneous features is proposed. The user need to give just a single click on an query image and images recovered by content based search are re-positioned dependent on their visual and texture similitudes to the query image. Textual and visual expansions are integrated to capture user intention without additional human feedback. Expanded keywords helps in extending positive model images and furthermore develop the image pool to include more relevant images. A lot of visual features which are both efficient and effective for image search are chosen. The n-dimensional feature vector for both colour and texture is reduced to single dimension each, used for comparing the similarity with query image using suitable distance metrics. Further only the images retrieved as a result of text based search and image re-ranking process are compared during run time for finding the similar images; not the entire database. This considerably reduces the computational complexity and improves the search efficiency. With improved feature extraction capturing textual and visual similarities, the proposed one click image search framework gives a productive robotized recovery of comparable images giving promising results with improvement in retrieval efficiency.

Keywords. Intent search, re-ranking, keyword expansion, visual similarities, clustering.

1. Introduction

Web-scale image a web crawler for the most part of surrounding content of images is used keywords as queries to scan for important images. They experience the ill effects of equivocallness of inquiry keywords since it is hard for users to precisely represent the visual substance of target image utilizing just watchwords. For instance, utilizing "Apple" as an inquiry watchword, the recovered images have a place with various classifications, for example, "Red Apple", "Apple I-Phone", "Apple Laptop", etc. It is critical to utilize visual data. Resolve ambiguity in text-based image retrieval. A system needs to filter images based on its content which provides better indexing and returns more accurate results. Content based image recovery (CBIR) refers to recommends a strategy that utilizes visual substance to look through a picture from wide enormous dataset of pictures as showed by the client's focal points and depends upon in this manner actuated solicitation picture highlights. In this the term 'features' can refer to hues, shapes, surfaces or whatever other data that can be acquired from the image itself. To bridge the semantic gap in the CBIR system[26], user involvement through relevance feedback (RF) is considered. This expects user to choose a few important and immaterial image cases from which visual comparability measurements

¹B. J. Dange, Dep of CSE, Shri Jagdishprasad Jhabarmal Tibrewala University, India.
Email: bapudange@gmail.com.

are found out through web based training. Images are re-positioned based on visual similarity. Be that as it may, for web-scale business frameworks, user input must be constrained to a base without web based training. Online picture re-arranging, which limits customer exertion to just a lone snap input, is a reasonable procedure to improve search things and its obligation is truly clear. Huge web picture web crawlers have grasped this framework. All of these means is automated, without additional exertion from the user. This is basically significant for any business online image internet searcher, where the UI must be incredibly straightforward. Other than this key commitment, a lot of visual features which are both powerful and proficient in image search are considered.

2. Related Work

The contributions of various authors to the development of an efficient image retrieval system can be classified into different areas. An exploratory evaluation of incalculable assembled picture descriptors for content-based picture recuperation is shown right now. For masking part extraction, used a concealing gratefulness vector to depict the photographs and the stowing away was viewed as superior to anything the histogram approach at any rate used UI-covering and two-concealing squares of concealing histograms for picture recuperation. Image retrieval effectiveness is improved by Xiang Jin et al., by posing several questions [5][6][7][8]. Xiaou Tang et al. [1][2][3][4] proposed technique to catch user expectation for a single tick Internet image search. He proposed the expectation of explicit weight blueprints to consolidate visual qualities and compute visual similitude versatile for inquiry images. Without extra human criticism, content and visual extension is coordinated to catch user aim. As of late various visual features were produced for image search. There were some global features, for example, GIST and HOG. A few measures of local features, for example, SIFT, in visual terms, and images are spoken to as sack of-visual-words (BoV). To save the geometry of visual words, spatial data was encoded in the BoV model in a few different ways. For instance, Zhang et al. [13] portray the geometry saving the visual stages that involve the local and long-extend spatial design of the visual stages.

In like way, unique data, which would out and out have the option to improve visual likeness learning, was not considered in past work. s. Liu et al. [9][10][19] proposed a keyword expansion strategy in which keyword extension is utilized to grow the recovered image pool and to broaden positive cases. The Keyword expansion was utilized essentially in archive recovery. In any case, for various query images, the successful low-level view properties change. In this way, Cui et al. [11][12][14][15] Inquiry images were characterized in different eight classes which are decided in prior, and distinctive feature waiting scheme were given for various kinds of query images. Be that as it may, it was difficult to classify all images in eight classes to cover the huge range of all web images. It was additionally likely for a wrong image to be arranged in an inappropriate classification.

There is a great deal of work on utilizing visual features to re-rank images retrieved by beginning text just search, be that as it may, without the requirement for user to choose query images. Tian et al. [16][17][18][20] Image re-ranking with a Bayesian structure. Hsu et al. [22] utilized data bottleneck (IB) hypothesis to amplify the interrelations between search relevance and visual features. Lu et al. [21] proposed

further references to refine the list items. Cai et al. [23] Re-ranked images with features/highlights that were physically characterized and gained from physically named preparing tests. These methodologies expected that a query was a significant semantic classification under the keyword. Images were ranked by demonstrating this significant class with visual and content highlights. This recommends query explicit semantic signature [24][25] is additionally powerful in this application, where it is critical to limit semantic differences by figuring the similarity of images. The remaining paper is structured as follows: The system architecture with work flow of proposed system is explained in section 3. Feature extraction techniques used for CBIR with different experimentations are also discussed in this section. In section 4, different result of experimental study with standard datasets and its analysis are presented. Section 5 concludes the paper.

3. System Overview

3.1 System Architecture

Figure 1 depicts the complete design of suggested framework. As appeared, the user initially enter keyword as query keyword to perform text based search to retrieve images from large dataset based matching on keyword with stored keyword associated with each image in image index file. A pool of images is recovered by text based search. From this text based search result user select the query image for further process by clicking on one of the image as per user intention At that point the user is approached to choose an query image from the image pool retrieved as result of text based search. Framework separates the highlights of inquiry picture. Pictures in the pool are re-positioned subject to their visual similarities to the solicitation picture. These similarities are figured by contrasting features and put away features of image stored in database. Bunches of outwardly comparative images are made and group with most comparable images is chosen for visual extension. This group is again contrasted and recovered coming about pictures and pictures are re-situated. This mirrors the customer's objective at better level.

3.2 Feature Extraction

For feature extraction, colour features, we used Colour Coherence Vector (CCV) method as colour features which is an enhancement over colour histogram. For texture feature, we used texture feature extraction method called Texture Element Feature Characterization (TEFC). The n-dimensional feature vector for both CCV and TEFC are reduced to single dimension each which is used for comparing the similarity with query image using suitable distance metrics. This considerably reduces the computational and search complexity. Images in the pool (result of search based on text keywords) are re-ranked depending on their pictorial likenesses to the request image. These likenesses are figured by looking at separation between the inquiry image and images in the pool. Further only the images retrieved in pool of images are compared with offline stored features as semantic signature [24] for finding the similar images not the entire features of images to reduce computational complexity. We used distance formula given in (1) for semantic signature matching and Cityblock distance for texture features to obtain similarity between images to rerank given by

$$CD = \sum_{i=1}^n |x_i - y_i|. \quad \dots\dots (1)$$

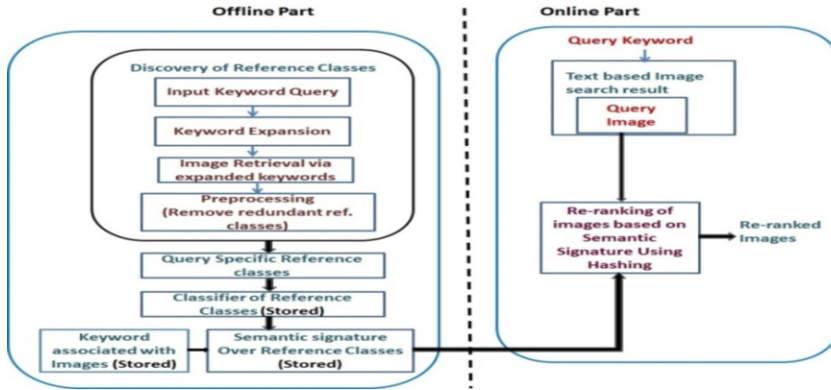


Figure 1.,Proposed Framework of One Click Image Search

3.3. Method

The step by step procedure to obtain the final image retrieval result of proposed one click image search is explored in Figure 1.

- Keywords/marks accommodated pictures; various names additionally (like creature, horse). Picture recovery dependent on watchword search. Database containing Image-Id and Labels.
- Query picture is therefore purposefully picked by the client from the pool of the photographs recovered. CBIR is performed utilizing hiding and surface highlights and arranging of the photographs in the pool are finished by separating question picture consolidate vector and pool of the broad number of pictures subject to CCV and city-square segment. Pictures from the pool are re-sorted out with the target that pertinent pictures will show up at the top.
- Keyword expansion: Top k pictures are picked and the entireties of its names are disengaged and the rehash of tally of each engraving is kept up. Right now, the printed depictions of picture words are removed from the top k pictures consistently like the solicitation picture. To save computational cost, only the top m words (in setting on repeat score) are used as probability for next getting ready.
- Clustering: For every competitor word w_i , we discover every one of the pictures containing w_i , and gathering them into various groups dependent on visual substance and repeat count of all these picked names. Finally this is used to pack all the photos present in the database into two gatherings: one have a spot with pictures having the removed names which are significant pictures and other containing remaining pictures. K means clustering is used.
- Visual Query Expansion: The gatherings of significant pictures are appeared. Pictures from both the pools are seen subject to feature vector w.r.t. request picture subject to city-square partition. Pictures moreover broke down by checking imprints and picture Id for the abundance.

4. Experimental Results

The system has been implemented using Netbean IDE 8.0.2, mysql-essential-5.4.3-beta-win32, Glassfish-4.1 webserver and JDK 1.8. The system is tested on Intel Pentium Quadcore CPU N3530 @ 2.16 GHz and 4 GB RAM. The performance of our system is tested on CUHK-Search-Reranking (CUHKSR) dataset. For each query keyword around 1,000 images are available in dataset. Figure 2 (a) shows the result of text based search and semantic signatures based one click image search, Figure 2(b) shows the result of top *m* precision for different query keywords used for search. Top *m* precision is defined as the proportion of relevant images among top *m* re-ranked images. Relevant images are those in the same category as the query image. The database consists of two folders one consist of images and other consist of link of images along with keywords.

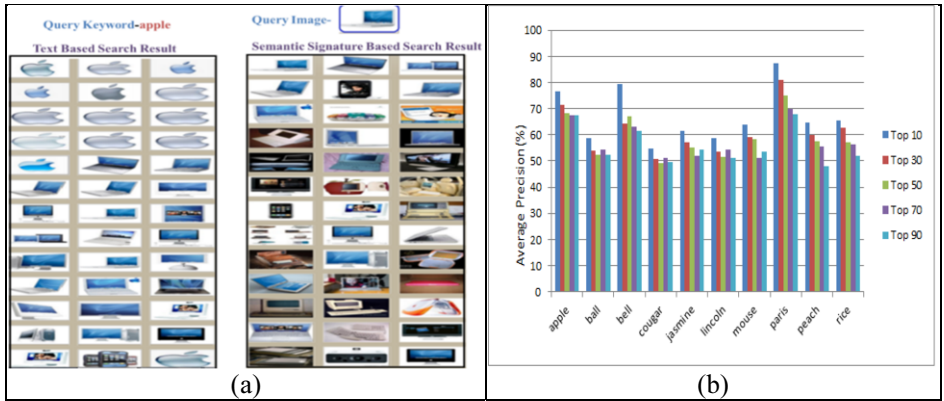


Figure 2. (a) Result of text based search and semantic signature based one click image search (b) Result of Average Top *m* Precision for 10 query Keywords

For sample, for inquiry “ball,” assigned labels for images are “cricket ball,” “Tennis ball,” “basketball,” etc. There are absolutely 1000 images for every one of the 10 inquiry watchwords. The search is typically founded on closeness as opposed to the definite equal. Outcome is evaluated by top *m* precision of search performed using by entering query input as text keywords, CBIR search is performed separately to search images from pool based on by providing the query image in which user is interested based on his liking and one click image retrieval framework is the proposed system designed and developed by considering the different phases as mention in the Figure 1. The performance of the proposed framework is measured in terms of its precision. Following tabulated results shows the performance and efficiency of the system. Figure 2 shows the Average Top *m* Precision for 10 Keywords. In this we have consider the semantic of images to improve the search efficiency of top *m* images which are more relevant to input image by user. To enhance the image search efficiency we have selected the similar classes of query keywords. Based on similarity classes we have constructed query dedicated reference classes which help to reduce the computational complexity of system. As reference classes are form based on different query keywords and if some reference classes contain the similar images then we are removing these reference classes to reduce the burden of system. In this

we are storing the classifier of reference classes as well as keyword associated with image in separate files as per provided in dataset. Semantic signature of images along with their reference classes is stored in database. The system is tested on 10 query keyword with approximately 1000 images for each keyword. The systems mainly consist of two parts like online and offline part. In offline part of system is trained and images with keyword and semantic signature is calculated and stored offline with respect to each reference classes. The semantic signature is nothing but the probability value of each image with respect to each reference class. This semantic signature is one dimensional vector which consists of probability value of each image how similar the image with each reference classes. In online part user input keyword to system as per user text search intention. The systems performs text search operation on stored database of keyword and images retrieved from stored database which matches keyword of query inputted by user with stored keyword associated with image database. After text based search images are displayed to user as result of text search. Displayed images may not more related to user search intention so user will select one image from result of text based search result by one click on that image which is more similar to user search intention image. For selected image semantic signature is calculated in online stage and compare with offline stored vector of semantic signatures of images. By comparing the semantic signature of query image and stored signatures distance is calculated using distance formula and the result is sorted in ascending order to rank more similar image at top as per the ascending sequence of distance value. The final result consists of images which are more similar to query image which user has selected by one click image search method.

5. Conclusion

For the development of an efficient CBIR system, a new data fusion technique to support text-based and content-based retrieval combining different heterogeneous features is proposed. The user is required to choose just a single tick on the inquiry image and images got by content based search are re-positioned dependent on their visual and printed comparability to the request image. Further, the n-dimensional feature vector for both CCV and TEFC are reduced to single dimension as semantic signature which is used for comparing the similarity with query image using suitable distance metrics. This considerably reduces the computational and search complexity. As per the comparison of average top m precision of Text Base Search, CBIR System and One Click Image Search System, we conclude that the One Click Image Search System gives improved result over all other systems. With improved feature extraction combining text based and content based retrieval approach, the framework provides an effective automated recovery of related images giving better results. In future the proposed system can be extended by using additional keywords for expansion also previous search result can be used as log for query keyword expansion to improve search efficiency.

References

- [1] Xiaoou Tang, Ke Liu, Jingyu Cui, Fang Wen. IntentSearch: Capturing User Intention for One-Click Internet Image Search. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 34, no. 7, 2012.
- [2] K. Porkaew, K. Chakraborty, and S. Meherotra. Query Refinement over Multimedia Similarity Retrieval in MARS. *Proc. ACM Int'l Multimedia Conf. (ACMMM)*, pp. 235-38, 1999.

- [3] Greg Pass, RaminZabih, and Justin Miller .Comparing Images Using Colour Coherence Vectors .Computer Science Department Cornell University Ithaca, NY 14853, pp. 1-14, 1998.
- [4] K. Jalaja, ChakravarthyBhagvati, B.L. Deekshatulu, and Arun K. Pujari .Texture Element Feature Characterizations for CBIR . IEEE, Dept. of Computer and Information Sciences University of Hyderabad, 2005.
- [5] D.C. He and Li Wang .Textures Filters Based on Texture Spectrum .Pattern Recognition, 24(1:1187-1195), 1991.
- [6] Benjamin Bustons, Daniel Kein, and Tobias Schreak .Content Based 3-D Object Retrieval.IEEE Computer Society, pp. 22-27, July 2007.
- [7] James French, A Chapin, and Worthy Martin .An Application of Multiple Viewpoints to CBIR . Proceedings of Joint Conference on Digital Libraries, JCDL 03,IEEE Computer Society, 2003.
- [8] Xiang Jin and James French .Improving Image Retrieval Effectiveness via Multiple Queries . MMDB, USA, pp. 86-93, 2003.
- [9] H. Pour and E. Kabir .Image Retrieval using Histograms of Uni-color and Bi-color Blocks .Pattern Recognition Letters, Elsevier, pp. 1547-1556, 2004.
- [10] ThomasDeselaers, Daniel Keysers, and Hermann Ney, "Features for Image Retrieval: an Experimental Comparison," Information Retrieval, DOI 10.1007/s10791-007-9039-3, pp. 77–107, Springer Science + Business Media, LLC, 2007.
- [11] J. Cui, F. Wen, and X. Tang, "IntentSearch: Interactive On-Line Image Search Re- Ranking," Proc. 16th ACM Int. Conf. Multimedia, 2008.
- [12] Lixin Duan, Wen Li, Ivor Wai-Hung Tsang, and Dong Xu .Improving Web Image Search by Bag-Based Re-ranking . IEEE Trans. Image Processing, vol. 20, no. 11, November 2011.
- [13] Y. Zhang, Z. Jia, and T. Chen .Image Retrieval with Geometry- Preserving VisualPhrases . Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, 2011.
- [14] J. Krapac, M. Allan, J. Verbeek, and F. Jurie .Improving Web Image Search Results Using Query-Relative Classifiers .Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2010.
- [15] X. Wang, K. Liu, and X. Tang .Query-Specific Visual Semantic Spaces for Web Image Re-Ranking .Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2010.
- [16] Jun Yu, Yong Rui, and Dacheng Tao .Click Prediction for Web Image Re-ranking Using Multimodal Sparse Coding , IEEE Trans. Image Processing, vol. 23, no. 5, May 2014.
- [17] N. Ben-Haim, B. Babenko, and S. Belongie, Improving Web- Based Image Search via Content Based Clustering.Proc. Int. Workshop Semantic Learning Applications in Multimedia, 2006.
- [18] L. Yang and A. Hanjalic .Supervised Re-ranking for Web Image Search . Proc. ACM Int. Conf. Multimedia, 2010.
- [19] S. Liu, F. Liu, C. Yu, and W. Meng .An Effective Approach to Document Retrieval via Utilizing WordNet and Recognizing Phrases .Proc. 27th Ann. Int. ACM SIGIR Conf. Research and Development in Information Retrieval, 2004.
- [20] X. Tian, L. Yang, J. Wang, X. Wu, and X. Hua .Bayesian Visual Re-ranking . IEEE Trans. Multimedia, vol. 13, no. 4, pp. 639-652, Aug. 2011.
- [21] J. Lu, J. Zhou, J. Wang, X. Hua, and S. Li .Image Search Results Refinement via Outlier Detection Using Deep Contexts.Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2012
- [22] W. Hsu, L. Kennedy, and S.F. Chang .Video Search Re-ranking via Information Bottleneck Principle . Proc. 14th Ann. ACM Int. Conf. Multimedia, 2006
- [23] J. Cai, Z. Zha, W. Zhou, and Q. Tian .Attribute-Assisted Reranking for Web Image Retrieval .Proc. 20th ACM Int'l Conf. Multimedia, 2012.
- [24] B.J. Dange, D.B. Kshirsagar .Hashing based re-ranking of web images using query-specific semantic signatures . Proc. 1st International Conference on Intelligent Systems and Information Management (ICISIM), DOI: 10.1109/ICISIM.2017.8122168, pp. 164-170, 2017.
- [25] B.J.Dange, D.B.Kshirsagar,A Survey of Various Image Retrieval and Re-ranking Methods to Search Web Images. International Journal Of Advance Research And In-innovative Ideas In Education (IJARIE) ,Volume 1 - Issue 4. (2015) Page 486-489.
- [26] V.D.Ambeth Kumar (2017) .coalesced global and local feature discrimination for content-based image retrieval", International Journalof Information Technology (Springer), December 2017, Volume 9, Issue 4, pp 431–44