ISSN:0975-9646

Framework on Capturing User Search Intention Based on Duplicate Image Detection

Ankita K. Jagtap
Computer Networking Department
Flora Institute of Technology
Pune- India

Mr. B. A. Tidke
Computer Networking Department
Flora Institute of Technology
Pune- India

***ABSTRACT*- Web image search is the major aspect in present days like Google image search. Web-scale image search engines rely almost purely on surrounding text features. It is difficult for them to interpret user's search intention only by query keywords and this leads to ambiguous and noisy search result. To solve the ambiguity in text based image retrieval use of visual information is important. Proposed search approach only requires the user to click on one query image with minimum effort and images from a pool retrieved by text based search are re-ranked based on both visual and textual content. First query image is categorized into adaptive weight category in which specific weight schema is used to rerank the text based result. Keyword is expanded through visual information and image clustering. In the later steps, the keyword expansion is used to expand image pool. The proposed method uses a visual vocabulary of vector quantized local feature descriptors (SIFT) to find similarity measures to evaluate near duplicate image detection. Without additional human feedback internet image search become effective and efficient. Using this duplicate image detection technique with existing system improves precision of top ranked images as result demonstrates.**

Keywords: Image search, Intention, Visual, Weight schema, Clustering, Keyword expansion, duplicate detection

I. INTRODUCTION

The process of retrieving and displaying relevant images based on user's queries from a database is known as Image search. Mostly Internet scale image search engines manage only keywords as queries. However, image retrieval is currently far less efficient then text retrieval because images are unstructured and much more difficult to process then text. User type query keywords to find certain type of images and search engine returns thousand of images extracted from the surrounding text. Many of returned images are noisy, disorganized, or irrelevant. Keyword based image search causes ambiguity in image retrieval because keyword tends to be very short, keyword's meaning may be richer then user's expectation. They cannot describe the visual content of image accurately. The search result consists of images with different semantic meaning. To reduce the ambiguity and to capture user's search intention, one way is text based keyword expansion and another way is content based image retrieval with relevance feedback, which requires more user efforts, makes it unsuitable for commercial system. It is important to use visual information in text based image retrieval to improve performance.

Data mining is the process of analyzing data from different perspective and summarizing it into useful information. Image data mining can be done manually by slicing and dicing or it can be done automatically. Primitive features of an image used to identify and retrieve closely matched images from an image database. Clustering is the process of organizing objects into groups whose members are similar in some way. Goal of clustering is to determine the intrinsic grouping in a set of unlabeled data. Clustering algorithms can be broadly classified into two groups: hierarchical and partitional. Hierarchical algorithms recursively find nested clusters either in a top-down (divisive) or bottom-up (agglomerative) fashion. In contrast, partitional algorithms find all the clusters simultaneously as a partition of the data and do not impose a hierarchical structure. Most hierarchical algorithms have quadratic or higher complexity in the number of data points and therefore are not suitable for large data sets, whereas partitional algorithms often have lower complexity.

II. EXISTING WORK

David G. Lowe [9]: have described a technique for extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. SIFT algorithm is used to find out the matched points from two images. Object recognition proceeds by matching individual features to a database of features from known objects using a fast nearest-neighbor algorithm, followed by a Hough transform to identify clusters belonging to a single object, and finally performing verification through least-squares solution for consistent pose parameters. Feng Jing et al. [2]: proposed an IGroup clustering algorithm that organizes Web image search results into semantic clusters in efficient and effective manner. This algorithm first identifies several semantic clusters related to query based on key phrases that are keywords extracted from general web search by result clustering algorithm. It then assigns all the resulting images to corresponding clusters based on visual or textual features. N. Ben-Haim et al. [10]: proposed an approach named ReSPEC (Re-ranking Sets of pictures by Exploiting Consistency) which is combination of two methods. Image is segmented into "blobs", from each of these blob set of features are extracted. Feature vectors clustered using mean shift algorithm. Uses color histograms in HSV color space as Features. Significant cluster obtained by computing mean in feature space and rest of images are restored based on the distance of their blobs to this mean using chi-squared distance comparison algorithm. Feature extraction is limited to color only, additional discriminative features

would be necessary to attack the more challenging problem of generic object recognition in an unconstrained set of images**.** J.Cui. et al. [1]: have proposed an approach to re-rank the text-based search results by using adaptive visual similarity. A query image is first categorized into one of several predefined intention categories, and a specific similarity measure is used inside each category to combine image features for re-ranking based on query image. Intention categorization model is used to integrate a set of complementary features adaptive to the query image. He designed attributes which makes intention categorization relatively easy. By modifying Rank Boost algorithm specific weighting schema is obtained inside each intention category by minimizes the rank loss for all query images. Different features have been designed that are effective in describing the content of images and efficient in computational and storage complexity. This approach has limitation that there are many irrelevant images in top *1.* ranked image because the visual similarity metrics learned from one query image is not robust enough.

J. Krapac et al. [8]: reported generic classifiers which are based on query-relative features. Without additional model training, generic classifier ranks images for previously unseen search queries. They combine textual features, based on the occurrence of query terms in web pages and image metadata and visual histogram representation of images. Query-relative features are used significantly to improve raw search engine ranking. Due to text based image retrieval user's intention is not accurately captured when the semantic meaning of query keyword had large diversity. W. Luo et al. [4]: examines Content based photo quality assessment framework with help of global and regional features. Subject areas are extracted having most attention. Regional features are extracted from subject area and background regions are combined with global features to assess the photo quality. Proposed system extracts several features such as subject area detection methods, new global and regional feature for content based photo quality assessment. In this approach images are manually divided into categories which are not feasible for large real time systems. X. Tang et al. [3]: have investigated an approach called IntentSearch which is an interactive system for real time web based image retrieval. Web user intention is captured by search engines based on re-ranking mechanism. Images are retrieved based on keywords, also based on query image. It involves categorization of query image, keyword expansion, visual query expansion which results into relevant images. S. Sunitha et al. [12]: To describe the content of image regions or the global content of an image used Gabor filter, which is a powerful texture extraction technique. Color histogram as a global color feature and histogram intersection as color similarity metric combined with Gabor texture give good retrieval results as that of region based retrieval systems. Provides better retrieval results for almost all semantic classes by combination of global and region based approaches. D.M. Spandhana et al. [11]: Propose an automatic annotation method by hybriding decision tree (DT) and support vector machine (SVM). A novel inverted file is used to rank the search result. Relevance model based re-ranking process utilizes global information from the image's HTML document. It is infeasible to manually label thousands of images retrieved from image search engine.

III. FUTURE SCOPE

Quality of reranked images can be further improved with photo quality assessment frame work.

IV. RESEARCH WORK

This paper proposes a fast and effective online image search algorithm based on one query image that capture user's search intention. This approach requires user to provide only one query image and images from a pool retrieved by text-based search are re-ranked based on their visual and textual similarities to the query image . Four steps are proposed to capture user search intention from this one click query image as follows:

Adaptive similarity

To describe different aspects of image uses set of visual features. The Adaptive Similarity is introduced with idea that a user always has precise intention when submitting a query image. For instance, if the user submits a picture with a big face in the middle, most likely he/she requires images with similar faces and using face-related features is more appropriate. The query image is first categorized into adaptive weight categories. Under every category, a specific pre-trained weight schema is used to combine visual features adapting to this kind of images to improved re-rank the text-based search result.

Adaptive similarity is calculated by using following formula.

$$s^q(i,j) = \sum_{m=1}^{F} s_m(i,j)$$

Where$s_m(i,j))$ is the similarity between image i and j on feature m, F is set of visual features.

The feedback function Φ_i

$$(\boldsymbol{j},\boldsymbol{k})\colon \boldsymbol{D_i}(\boldsymbol{j},\boldsymbol{k}) = \frac{\Phi_{\boldsymbol{i}}(\boldsymbol{j},\boldsymbol{k})}{\sum_{\boldsymbol{j},\boldsymbol{k}} \Phi_{\boldsymbol{i}}(\boldsymbol{j},\boldsymbol{k})}$$

Where i is query image, a real-valued feedback function $\Phi_i(j.k)$is defined to denote preference between image j and k.

This correspondence between the query image and its proper similarity measurement reflects the user intention. This initial reranking result is not good enough and will be improved by the following steps.

2. Keyword expansion

User's entered query keywords tend to be short and some significant words may be missed because of users' lack of knowledge on the textual description of target images. To capture users' search intention query keywords are expanded, inferred from the visual content of query images, which are not considered in traditional approaches. A word w is recommended as an expansion of the query if a cluster of images are visually similar to the query image and all contain the same word w. The expanded keywords better capture users' search intention since the consistency of both visual content and textual description is ensured.

3. Image pool expansion

Reranking images in the pool retrieved by text-based search is not very effective because image pool

accommodates images with a large variety of semantic meanings and the number of images related to the query image is small. Thus, more accurate query by keywords is needed to narrow the intention and retrieve more relevant images. Keyword expansions are suggested using both visual and textual information better capture users' intention. They are automatically added into the text query and enlarge the image pool to include more relevant images. Feedback from users is not required.

Proposed System Architecture

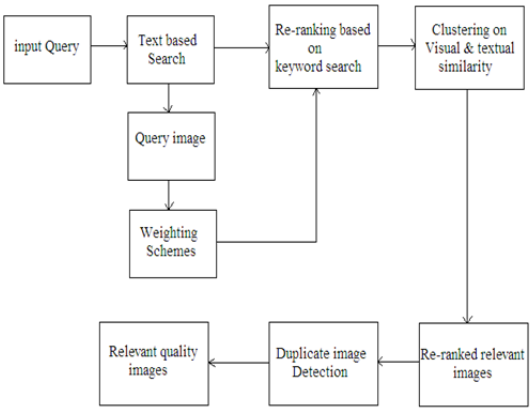

Fig. 3 Proposed System Architecture.

First User enters a keyword on the search engine, which retrieves the thousands of images, from which user clicks on one query image. Then categorize query image into predefined adaptive weight categories. Re-rank images based on keyword based expansion is performed using *tf-idf (*term frequency–inverse document frequency*)* method. Perform clustering of images based on the visual and textual similarity. Efficient re-ranking of images is done using visual and textual similarity. Finally detection of duplicate images based on similarity measures by using duplicate image detection algorithm is done and relevant quality images are obtained.

4. Duplicate Image Detection

The similarity measures are applied and evaluated in the context of near duplicate image detection. The proposed method uses a visual vocabulary of vector quantized local feature descriptors (SIFT). Following are four major stages of computation used to generate the SIFT features:

- Scale-space Extrema detection
- Keypoint localization
- Orientation Assignment
- Keypoint Descriptor

V. RESULTS

Top m precision, the proportion of relevant images among the top m ranked images, is used to evaluate the performance of image reranking. Images are considered to be relevant if they are labeled as the same class. In order to evaluate the effectiveness of different steps of our proposed image reranking framework, we compare the four approaches such as text based search; visual similarity based search, intent search, and our intent search plus duplicate removal. The averaged top m precisions are shown in Fig. 2. Using our approach, the top 70 precision of reranking using adaptive weight is improved from 41.4 to 61.4 percent.

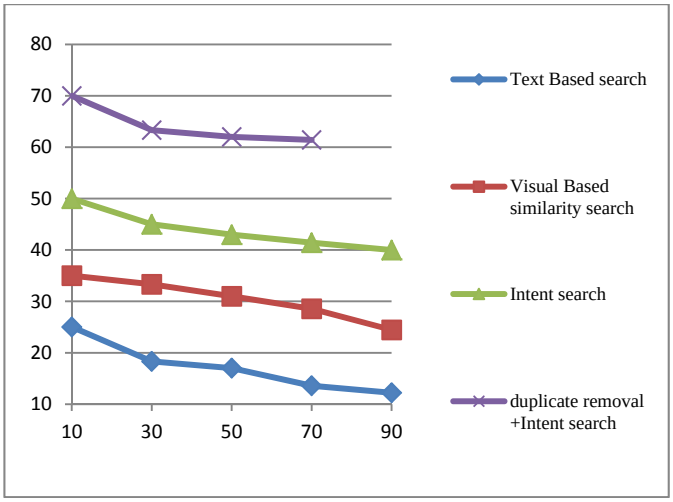

Fig.2. Comparison of averaged top m precisions with existing methods.

VI. CONCLUSION

Image search is a particular data search used to find images. To search for images, a user may give query terms such as keyword, image file/link, or click on some image, and the system will return images "similar" to the query. In this, a novel Internet image search approach which only needs one-click user response. Intention specific weight schema is proposed to combine visual features and to compute visual similarity adaptive to query images. Without additional human feedback, textual and visual expansions are integrated to capture user intention. Expanded keywords are used to extend positive example images and also enlarge the image pool to include more relevant images. This framework makes it possible for industrial scale image search by both text and visual content. One shortcoming of the Intent search system is improved by including duplicate image detection in the framework. Finally, to further improve the quality of re-ranked images, we intent to combine this work with photo quality assessment work in to re-rank images not only by content similarity but also by the visual quality of the images.

REFERENCES

[1]. J. Cui, F. Wen, and X. Tang, (2008), *"Real Time Google and Live Image Search Re-Ranking,"* Proc. 16th ACM Int'l Conf. Multimedia.

[2]. F. Jing, C. Wang, Y. Yao, K. Deng, L. Zhang, and W. Ma, (2006), "Igroup: Web Image Search Results Clustering," Proc. 14th Ann. ACM Int'l Conf. Multimedia.

[3]. Xiaoou Tang, Fellow, IEEE, Ke Liu, Jingyu Cui, Student Member, IEEE, Fang Wen, Member, IEEE and Xiaogang Wang, Member, IEEE, (2012), " IntentSearch:Capturing User Intention for One-Click Internet Image Search".

[4]. W. Luo, X. Wang, and X. Tang, (2011), "Content-based photo quality Assessment," Proc .IEEE Int'l Conf. Computer Vision.

[5]. YiwenLuo and Xiaoou Tang," Photo and Video Quality Evaluation: Focusing on the Subject", Department of Information Engineering

The Chinese University of Hong Kong, Hong Kong ¡ywluo6, xtang¡@ie.cuhk.edu.hk

[6]. Yan Ke1, Xiaoou Tang2, Feng Jing2," The Design of High-Level Features for Photo Quality Assessment", 1School of Computer Science, Carnegie Mellon; 2Microsoft Research Asia yke@cmu.edu; {xitang, fengjing}@microsoft.com.

[7]. Jingyu Cui, Fang Wen, Xiaoou Tang, (2008) "IntentSearch: Interactive On-line Image Search Re-ranking."

[8]. J. Krapac, M. Allan, J. Verbeek, and F. Jurie, (2010), "Improving web image search results using query-relative classifiers," in Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition.

[9]. David G. Lowe, (January 5 2004)," Distinctive Image Features from Scale-Invariant Keypoints", Computer Science Department University of British Columbia Vancouver, B.C., Canada lowe@cs.ubc.ca.

[10]. N. Ben-Haim, B. Babenko, and S. Belongie, (2006), "Improving Web- Based Image Search via Content Based Clustering," Proc. Int'l Workshop Semantic Learning Applications in Multimedia.

[11]. D.Manasa Spandhana, B RajendraPrasad Babu (2014), "Interactive Online Image Search Re-Ranking", IJESC.

[12]. S.Sunitha, A.RamaSatish, (2014),"Extended Image Features for User Intention Refined Image Search", International Journal of Advanced Research in Computer Science and Software Engineering.

[13]. J. Philbin, M. Isard, J. Sivic, and A. Zisserman, (2010), "Descriptor Learning for Efficient Retrieval," Proc. European Conf. ComputerVision.

[14]. Y. Zhang, Z. Jia, and T. Chen, (2011), "Image Retrieval with Geometry- Preserving Visual Phrases," Proc. IEEE Int'l Conf. Computer Visionand Pattern Recognition.

[15]. G. Chechik, V. Sharma, U. Shalit, and S. Bengio, (2010), "Large Scale Online Learning of Image Similarity through Ranking," J. MachineLearning Research, vol. 11, pp. 1109 1135.

[16]. J.Deng,A.C.Berg,andL.Fei-Fei, (2011), "HierarchicalSemanticIndexing for Large Scale Image Retrieval," Proc. IEEE Int'l Conf.Computer Vision and Pattern Recognition.