ISSN: 2321-7782 (Online)
Volume 2, Issue 6, June 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

Internet Image Search Based On User Intention

Raviraj Kasture[1]
Computer Engineering Department
P.V.P.I.T, Bavdhan
Pune – India

Dr. A. M. Dixit[2]
Computer Engineering Department
P.V.P.I.T, Bavdhan
Pune – India

Abstract: Image search is a particular data search used to determine images. For searching images, a user may give query terms such as keyword, image file, or click on few image, and the system will determine images "similar" to the query. The resemblance used for search criteria could be Meta tags, color distribution in images, region/shape attributes, etc. Many commercial Internet scale image search engines use only keywords as queries. The search engines (e.g. Google Image Search, Bing Image Search) mostly depend on surrounding text features. It is not easy for them to understand user's search intention only by query keywords and this leads to uncertain and noisy search results. It is important to use visual information in order to solve the ambiguity in text-based image retrieval. In this paper, we propose a novel Internet image search approach. The user needs to click on one query image with the minimum attempt and images from a pool retrieved by text-based search are re-ranked based on both visual and textual content. To capture the users' search intention from this one-click query image in four steps has been presented in this paper.

Keywords: Image Search, Intention, Visual, Web Image Search, Clustering, User Interface.

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. Users type query keywords in the hope of resulting a certain type of images. The search engine proceeds with thousands of images ranked by the keywords extracted from the adjacent text. It is familiar that text-based image search suffers from the ambiguity of query keywords. The keywords provided by users tend to be short. For example, the average query length of the top 1, 000 queries of Picsearch is 1.37 words, and 98% of them include only one or two words. They cannot explain the content of images accurately. The search outcomes are noisy and consist of images with quite different semantic meanings. Figure 1 shows the top ranked images from Bing image search using "apple" as query. They belong to different categories, such as "green apple", "red apple", "apple logo", and "iphone", because of the ambiguity of the word "apple".

Fig. 1. Top ranked images returned from Bing image search using "apple" as query

II. RELATED WORK

The purpose of content based image retrieval is that whenever a user gives a query image, it retrieves the images that are mostly related to the content. Nadav Ben et al [7] attained the content based image retrieval by introducing a new approach name ReSPEC (Re-ranking Sets of Pictures by Exploiting Consistency). ReSPEC is composed of two main methods. Firstly, based on the user query image (keyword) the image search engine (Google, Yahoo,…etc) retrieves the images then, clusters the results based on extracted image features, and returns the cluster that is inferred to be the most relevant to the search query. Secondly, ranks the results that are most relevant to the user query images.

1. Segmentation of Image

In image segmentation each images collected from the image search engine has been broken into regions of resemblance, with the intuition that each of these regions is a separate object in the image by using a graph based approach.

2. Selection of Feature

In order to obtain a measure of how similar image blobs are to one another, good features are needed to represent the blobs. Color histograms in HSV color space used to represent the image features. To form a feature vector for each blob, histograms are built for the H, S and V channels, with 15 bins each, and then concatenated together to form a 45 dimensional feature vector.

3. Mean Shift Clustering in Feature Space

The next step in the system is to cluster the blobs, according to their extracted features with the hope that the object of interest will form the largest cluster. Since some of the blobs will represent garbage, it is difficult to predict the number of clusters that are present.

4. Re - ranking of Images

After obtaining the "significant" cluster in feature space, the mean is computed. The rest of the images are then resorted based on the distance of their blobs to this mean. Since each image could potentially contain more than one blob, the closest blob in each image is used.

III. SYSTEM ARCHITECTURE

We do consider that adding visual information to image search is significant. On the other hand, the interaction has to be as simple as possible. The absolute minimum is One- Click. In this paper, we plan a novel Internet image search approach. It wants the user to provide only one click on a 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. We consider that users will believe one-click interaction which has been used by several popular text-based search engines. For instance, Google needs a user to choose a recommended textual query expansion by one-click to obtain extra outcome. In this paper, the main problem to be solved is how to capture user intention from this one-click query image. Four steps are proposed as follows.

1. Adaptive similarity

Designing a set of visual features to explain different aspects of images. How to combine different visual features to calculate the similarities between the query image and other images is a main problem. An Adaptive Similarity is introduced to deal with a user always has precise intention when submitting a query image. For instance, if the user gives a picture with a big face in the middle, most likely he/she requires images with related faces and with face-related features is more suitable. The query image is firstly categories, such as "portrait" and "scenery". Under every type, a exact pre-trained weight schema is calculated to combine visual features adapting to this type of images to improved re-rank the text-based search outcome. This correspondence among the query image and its appropriate similarity measurement reflects the user intention. This primary re-ranking outcome is not fine enough and will be enhanced by the following steps.[2]

2. Expansion of Keyword

Users entered the query keywords which tend to be short and some significant keywords may be missed since 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, this is not considered in traditional approaches. A word *w* is recommended as an expansion of the query, if a query image are visually analogous to the cluster of images and all include the same word *w* 1.The expanded keywords improve users' search intention because the consistency of both visual content and textual description is ensured.

3. Image pool expansion

The image collection restored by text-based search accommodates images with a large selection of semantic meanings and the number of images related to the query image is small. Here, reranking images in the group is not very efficient. Hence extra correct query by keywords is necessary to narrow the intention and restore more significant images. A naïve way is to request the user to click on one of the recommended keywords known by traditional approaches only with text information and to increase query results similar to Google "Related Searches". This enhances users' burden. Furthermore, the recommended keywords based on text information only are not precise to clarify users' intention. Expansions of keywords recommended by our approach with both visual and textual information better capture users' intention. They are automatically added into the text query and extend the image pool to contain extra important images. Opinion from users is not mandatory. Our experiments illustrate that it considerably improves the accuracy of top ranked images

4. Expansion of visual query

For Capturing user intention one query image is not particular. In Step (2), a collection of images all having the same extended keywords and visually similar to the query image are set up. For Image Reranking, images are selected as expanded optimistic examples to study visual and textual similarity metrics, which are more robust and more precise to the query. Comparing the weight schema in Step (1), these comparison metrics reflect users' intention at a higher level since each query image has dissimilar metrics. Here visual expansion does not require users' opinion.

5. Image Search and Visual Expansion

Several Internet scale image search methods are text-based and are restricted by the truth that query keywords cannot explain image content accurately. Content-based image retrieval utilizes visual features to estimate image resemblance. Many visual features were developed for image search in current years. A few were global features such as GIST and HOG. A few quantized local features, like SIFT, into visual words and described images as bags-of-visualwords (BoV). To maintain the geometry of visual words, spatial information was encoded into the BoV model in many ways. For instance, Zhang et al. proposed geometry-preserving visual stages which captured the local and long-range spatial layouts of visual words. One of the major challenges of content-based image retrieval is to study the visual similarities which well return the semantic significance of images. Image resemblance can be studied from a huge training set where the significance of pairs of images is recognized. Deng et al. studied visual similarities from a hierarchical structure defined on semantic features of training images. Because web images are extremely diversified, defining a set of features with hierarchical relationships for them is challenging. In general, learning an entire visual similarity metric for standard images is still an open problem to be resolved. A few visual features may be more efficient for definite query images than others. So as to make the visual similarity metrics more precise to the query, significance opinion was extensively used to develop visual examples. The user was asked to choose numerous significant and irrelevant image examples from the image pool.

6. Expansion of Keyword

In this approach, expansion of keyword is used to extend the retrieved image pool and to extend positive examples. Expansion of keyword was mostly used in document retrieval. Thesaurus-based methods extended query keywords with their linguistically associated words such as synonyms and hyponyms. Corpus-based methods, like well known term collecting and Latent Semantic Indexing calculated the resemblance of words based on their co-occurrences in documents. Words mainly related to the query keywords were selected as textual query expansion. A few image search engines have the characteristic of extended keywords implication. They typically use adjacent text. Some algorithms produced tag suggestions or comments based on visual content for input images. Their aim is not to develop the performance of image reranking. While they can be observed as options of keyword expansions, some difficulties avoid them from being directly used to our problem. The majority of them assumed permanent keyword sets, which are difficult to get for image re-ranking in the open and dynamic web environment. Several annotation methods need supervised guidance, which is also hard for our problem. Unlike than image annotation, this method gives more image group through the process expansion of keyword, and such image group can be used as visual expansions to further progress the performance of image re-ranking.

IV. Method

The flowchart of our approach is below in Figure 2.First query keywords *q* is given by the user. A collection of images is restored by text-based search (Figure 2a). Then the user request to choose a query image from the image pool. The query image is categorized as one of the type of predefined adaptive weight. Based on the visual similarities to the query image, collected images are reranked (Figure 2b) and the similarities are calculated using the weight schema (Figure 2c) precised by the class to combine visual feature.In the expansion of keyword (Figure 3d), words are extracted from the textual descriptions (like image file names and adjacent texts in the html pages) of top *k* images mainly analogous to the query image, and the *tf-idf* method is used to rank these words. To keep computational cost, only top *m* words are kept as candidates for more processing.

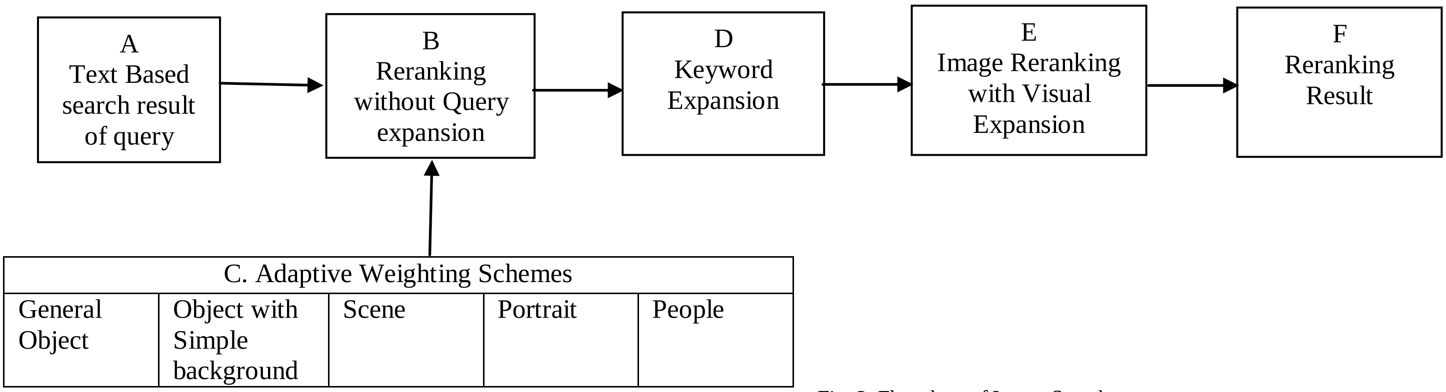

Fig. 2. Flowchart of Image Search

The Shark-search algorithm:

We will assume that such an engine is available and that for any pair query, ***(q,d)***, it returns a similarity score ***sim(q,d)*** between *0* and *1*

1. **Compute** the inherited score of *child_node*, *inherited_score(child_node)*, as follows:
   - **If** *relevance(current_node) > 0* (the current node is relevant)
     Then *inherited_score(child_node) = d * sim(q,current_node)*
     where *d* is a predefined decay factor.
     Else *inherited_score(child_node) = d * inherited_score(current_node)*
2. **Let** *anchor_text* be the textual contents of the anchor pointing to *child_node*, and *anchor_text_context*, the textual context of the anchor (up to given predefined boundaries)

3. **Compute** the relevance score of the anchor text as *anchor_score = sim(q,anchor_text)*

4. **Compute** the relevance score of the anchor textual context as follows:

   - **If** *anchor_score > 0,*

     Then *anchor_context_score = 1*

     Else *anchor_context_score = sim(q,anchor_text_context)*

5. **Compute** the score of the anchor, that we denote *neighborhood_score* as follows:

   *neighborhood_score = b * anchor_score + (1-b) * anchor_context_score*

   where *b* is a predefined constant

6. **Compute** the potential score of the child as

   *potential_score(child_node) = g * inherited_score(child_node) + (1-g) * neighborhood_score(child_node)*

   where *g* is a predefined constant.

V. Evaluation Results

To simplify the evaluation, we therefore propose to use a measure that reflects what most users expect from a "good" map: namely, getting as many relevant images in the shortest delays. We define this measure as the sum of similarity scores of all images forming the maps, which we will refer to as the "*sum of information"* measure in the rest of this paper. Note that we do not use an average measure (for instance dividing the sum by the total number of nodes in the map) since irrelevant images that were identified in the same process do not hurt the quality of the map. On the contrary, they give context information and can be pruned from the map if so desired.

So in summary, if we define as ***sim(d,q)***, the similarity score (between *0* and *1*) between any image **d** in the map and query **q** directing the mapping, and **M** the set of images in a given map, we have

sum_of_information(M,q) = □$_{(d\ in\ M)}$ sim(d,q)

Note that ***sim(d,q)*** is the actual similarity score as computed by the similarity engine for images whose contents was fetched by the crawler. Potential scores, which are speculative, do not affect this value.

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. The proposed new image reranking framework consists of multiple steps, which can be improved separately or replaced by other techniques equivalently effective. In the future work, this framework can be further improved by making use of the query log data, which provides valuable co-occurrence information of keywords, for keyword expansion. One shortcoming of the current system is that sometimes duplicate images show up as similar images to the query. This can be improved by including duplicate detection in the future work. 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, "Real time google and live image search re-ranking," in Proc. ACM Multimedia, 2008.
2. Xiaoou Tang, Fellow, IEEE, Ke Liu, Jingyu Cui "Intent Search: Capturing User Intention for One-Click Internet Image Search", IEEE
3. Y. Zhang, Z. Jia, and T. Chen, "Image retrieval with geometry preserving visual phrases," in Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, 2011.
4. S. Aarif Ahamed, B. A. Vishnupriya, V. Venkateshwaradevi, "Survey On Novel Techniques for Effective Image Search Based On Users Intention" International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 4, April – 2013 ISSN: 2278-0181
5. J. Philbin, M. Isard, J. Sivic, and A. Zisserman, "Descriptor Learning for Efficient Retrieval," in Proc. European Conf. Computer Vision, 2010.
6. N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. in Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, 2005.
7. N. Ben-Haim, B. Babenko, and S. Belongie, "Improving web-based image search via content based clustering," in Proc. Int'l Workshop on Semantic Learning Applications in Multimedia, 2006.