International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

REVIEW OF IMAGE WITH SKETCH BASED SYSTEM BY RE-RANKING

***KUMAR.G**

LORDS INSTITUTE OF ENGINEERING & TECHNOLOGY

****AFTAB AMIL AHMEED**

LORDS INSTITUTE OF ENGINEERING & TECHNOLOGY

*****LOKESH.M**

LORDS INSTITUTE OF ENGINEERING & TECHNOLOGY

****ABDULLAH MOHAMMAD****

LORDS INSTITUTE F ENGINEERING & TECHNOLOGY

ABSTRACT

To retrieve image form a web, text-based image search is easy and known process in which we give image names or tags as query to search engine so that it will provide desired set of images relevant to a query from huge image collection.Web based image re-ranking is used to produce a desired way to improve the result of web based image search. Feature extraction and ranking function design are two key steps in image search reranking. The purpose of web based image search re-ranking is to reorder retrieved elements to get optimal rank list. However, existing re-ranking algorithms are limited for two main reasons: 1) the textual meta-data associated with images is often mismatched with their actual visual content and 2) the extracted visual features do not accurately describe the semantic similarities between images. A major challenge in re-ranking the web based image is that the similarity of visual features does not well correlate with image. This paper presents a detail review of comparative analysis of different Image Re-ranking approaches. The purpose of the survey is to provide an overview and analysis of the functionality and future scope of existing image re-ranking systems, which can be useful for researchers for developing effective system with more accuracy.Sketch-based image retrieval often needs to optimize the trade-off between efficiency and precision. Index structures are typically applied to large-scale databases to realize efficient retrievals. However, the performance can be affected by quantization errors. Moreover, the ambiguousness of user-provided examples may also degrade the performance, when compared with

traditional image retrieval methods. Sketch-based image retrieval systems that preserve the index structure are challenging. In this paper, we propose an effective sketch-based image retrieval approach with re-ranking and relevance feedback schemes. Our approach makes full use of the semantics in query sketches and the top ranked images of the initial results. We also apply relevance feedback to find more relevant images for the input query sketch. The integration of the two schemes results in mutual benefits and improves the performance of sketch-based image retrieval.

1. INTRODUCTION

Image Search Re-ranking is defined as the refinement of search results by employing image visual information to reorder the initial text-based search results. It comes from the observation that the noisy text-based search results still contain satisfactory images in top hundreds of search results. The ranking of images based on a textbased search is considered a reasonable baseline. Extracted visual information is then used to re-rank related images to the top of the list. Therefore, reordering of these top ones with visual cues is possible to satisfy user's search experience in both accuracy and response time. It can be viewed as a post-process of core search. Thousands of images are uploaded to the internet with the explosive growth of online social media and the popularity of capture devices [1], thus building a satisfying image retrieval system is the key to improve user search experience. Due to the success of information retrieval, most commercial search engines employ textbased search techniques for image search by using associated textual information, such as file name, surrounding text, URL, etc. Even though text-based search techniques have achieved great success in document 252 | P a g e retrieval, text information is often noisy and even unavailable. In order to improve search performance, image search re-ranking, which adjusts the initial ranking orders by mining visual content or leveraging some auxiliary knowledge, is proposed, and has been the focus of attention in both academia and industry in recent years .Most of the existing re-ranking methods utilize the visual information in an unsupervised and passive manner to overcome the "semantic gap" (the gap between the low-level features and high-level semantics). Although multiple visual modalities have been used to further mine useful visual information, they can only achieve limited performance improvements. This is because these re-ranking approaches neglect the “intent gap”

Traditional draw and search systems require that the input sketch is colored and similar to a real photo [3]. This approach converts sketch-based retrieval to content-based image retrieval. The user must draw the sketch carefully and color it to make the sketch visually similar to the natural scene images. Then, CBIR fuses different features (such as shape, color, and texture) together to perform retrieval. However, this method will burden users by requiring detailed drawings, and most importantly, it does not solve the core problem of SBIR, i.e., matching a line-formed sketch and colored images [2]. Image retrieval must deal with the difference between the user’s desire and the query example. This difference may be even more severe in sketch-formed queries, because of the ambiguousness in the query sketch caused by a lack of semantic information such as texture attributes [1] and luminance [15]. A simple and similar image is needed for image-based retrieval. But for SBIR, results may vary dramatically if the user’s drawing skills are not sophisticated, or if the target cannot be simply depicted using only lines. For example, if a user is looking for pictures of a pyramid but they can only draw a triangle, sketch-based retrieval becomes very challenging [2]. To address this problem, researchers proposed incorporating sketches and text descriptions to disambiguate the input. Lin et al. proposed a method that does not use lines to form the query sketch [4]. The sketch is a drawing that uses different words to represent diverse objects. Their locations and sizes are represented by the words. With the help of these words, the approach first finds some corresponding exemplars, which is then used to search for objects in images. In this sense, it is like a concept-based image retrieval system instead of a sketch-based method.

Keywords: Feature Extraction, Image Re-Ranking, Image Retrieval, Image Search, Ranking Function

2. RELATED WORK

Many SBIR methods have been proposed over the past 20 years. Query by visual example [10] defines a pictorial index for each image, and computes the correlations between the corresponding indexes to retrieve the results. An image is divided into equalized blocks and the correlation is calculated by shifting these blocks. Zernike moment is a moment invariant method that has been used in SBIR [12, 13]. It can solve the rotation, scale, and translation invariant problems. The method in [13] uses Zernike orthogonal polynomials to extract the Zernike moment descriptor of an image, and uses the Manhattan distance to measure the similarity between a sketch and image. The edge histogram descriptor (EHD) and the histogram of oriented gradients (HOG) are also used to establish the SBIR system [14]. They are both global features extracted from the edges of images. Chalechale et al. proposed an angular partition approach that divides the edge into several blocks in terms of orientations [41]. An angular radial partitioning-based SBIR approach was proposed in [1], which considers the radial factor during the retrieval process. Most existing methods mainly use global features or divide images into blocks to represent the image [12–15]. These methods do not work well because of the ambiguousness of sketches and shapes. Additionally, the incompleteness of a user's drawing may also affect the results. Consequently, researchers proposed exploring the local saliency in SBIR. Chen et al. used a freehand sketch and some text labels to search for Internet images [16]. Although this method was very accurate, it was very computationally expensive. Thus, a SBIR with index structures is more appropriate for a large-scale image set, and achieves the best balance between the retrieval performance, and time and storage costs. The edgel index approach is a shape-based indexing method [2]. It solves the shape-to-image matching problem using pixel level matching. Oriented chamfer matching [17] is used to compute the distance between contours.

Different Researchers are working on the methods used to get desired performance of web search engine. Xiaopeng Yang, Tao Mei, [1] proposed click-based multi-feature similarity learning algorithm. Based on the learnt click-based image similarity measure, they conducted spectral clustering to get the final re-rank list by calculating click-based clusters typicality and within clusters click-based image typicality in descending order. ZhongJi, Yanwei Pang, Xuelong Li, [2] addressed the feature extraction and ranking function problems in image search re-ranking based on the hypersphere idea in one-class classification, they observed two things : 1. How to transfer the ISR methods to solve the outlier removal problem is an interesting research direction. 2. Deep learning has shown its promising successes in image classification and CBIR, however it has little significant influence on TBIR. How to employ it in TBIR and ISR is also a challenging direction Yongdong Zhang, Xiaopeng Yang, and Tao Mei [3] presented Image Search Re-ranking With Query-dependent Click-Based Relevance Feedback algorithm emphasizes the successful use of click-through data for identifying user search intention, while leveraging multiple kernel learning algorithm to adaptively learn the querydependent fusion weights for multiple modalities. This paper has considered only image search relevance, though image diversity is another important factor in search performance. Future work will be enhancing the diversity of re-ranked images by duplication detection or other such methods. JunjieCai, Zheng-Jun Zha, Meng Wang, Shiliang Zhang, and Qi Tian [4] proposed a visual-attribute joint hypergraph learning approach to simultaneously explore two information sources. A hypergraph is constructed to model the relationship of all images. 253 I P a g e un Yu, Member, Yong Rui and Dacheng Tao

3. SKETCH-BASED IMAGE RETRIEVAL WITH RELEVANCE FEEDBACK

The framework of the proposed SBIR system is shown in Fig. 1. It consists of two parts: the offline part and the online part. Our approach can be included at the back end of any initial SBIR system (such as the edgel [2] and ARP [1] methods) using relevance feedback to improve performance. We now focus on an edgel SBIR system to illustrate our approach. In the offline part of the method, we must build an edgel index structure for each image based on the Berkeley edge detector [21]. Then, we extract SIFT features and record the SIFT descriptors with their locations and orientations. Finally, we build a contour similarity index for each image. In the online part, for a given input query sketch, we sequentially execute five stages: 1) the initial SBIR [2], which obtains the initial result shown to the left of Fig. 1; 2) relevant image grouping for the initial results, which finds the relevant images from the top R images in the top N ranked results; 3) re-rank and verify the results using SIFT matching; 4) contour-based relevant feedback to find more relevant images; and 5) re-rank the results of the relevant feedback to improve the

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

performance. A. Sketch-Based Image Retrieval In the offline system, we build a feature index structure (such as edgel) for each image, as in [2]. More details can be found in [2]. We give a brief overview of the approach, which consists of the following three steps. 1) For an image database with T images, we apply the Berkeley detector [21] to each image (resized to 200×200). This produces hit maps with six orientation channels (θ=6).

Thus, for each image, we build an index structure with 200×200×6 entries for the six orientation channels. 1) The Berkeley detector [21] extracts contours. It uses the brightness, color, and texture gradients to accurately detect and localize the boundaries of images. 2) For each point at a certain orientation, we build an inverted list for fast indexing (i.e., the edgel index structure used in [2]). For each edgel point in the contours, the position (x,y) and quantized orientation channel θ are combined to (x,y,θ). For each entry (x,y,θ), we build an inverted list of images (IDs). 3) When a query sketch Q (normalized to 200×200 entries) is input to the system, six hit maps are generated by marking the regions surrounding the sketch lines within a certain radius, and quantizing each edge orientation into six channels [2]. By comparing the edgels (x,y,θ) of the hit maps of the query sketch and the edges extracted from the database images, we can measure the similarities between the sketch and images. Each edgel marked in the hit maps is used to search the inverted list for corresponding image IDs. Finally, the similarity between the query sketch (Q) and the image (D) in the database is computed by counting how many times D appears during the search. We sort the similarity scores in descending order, and determine the initial results (the N top ranked). In the following steps, we apply re-ranking and relevance feedback schemes to these N images. B. Relevant Images Grouping For Relevant Feedback The top-ranked images obtained by the initial SBIR may contain irrelevant images. In our approach, the relevant images are the ones that occur most in the top N images. We make full use of the top R images ($R < N$) to find relevant images for CBRF. Our approach is motivated by retrieval results clustering, which improves the diversity of top-ranked results [42,43] by finding near duplicated image groups [44–46]. We apply near-duplicate image clustering to the top ranked R images to find similar images from the top N initial SBIR results [46]. This approach consists of the following steps. 1) For each image, we record the SIFT descriptors together with their locations (x,y) and orientations [30,31]. The SIFT feature extraction is carried out off-line for the dataset images. 2) We first find near-duplicated images for the top R images of the top N images returned by the initial SBIR, as shown in Fig. 1. We use the similarity measurement (i.e., near-duplicate image detection) with the existing image matching approach [37,46]. In this paper, we use binary edge-SIFT to carry out the near-duplicate image retrieval approach and find near-duplicate image groups. 3) We further cluster the detected near-duplicate images into groups for the top ranked R images. Assume that the group number is K ($K \leq R$) and we record the corresponding image numbers. 4) We use the cluster with the most near duplicate images as relevant image groups for the query sketch. At the same time, we set the initial scores of images in the relevant image group as their maximum, and the initial

scores of the irrelevant images as their minimum. This step ranks the images in the relevant image group ahead of the other images. Using relevant image grouping, we can roughly eliminate the noisy images from the top-ranked results. Then, we further use the top N images with RVFV to obtain more relevant images. We use the duplicate image group from the top R-ranked images (denoted by top-R+top-N), rather than the top N images to eliminate noise. Generally, a higher-ranked image is more relevant to the query sketch. If we use the top-ranked N images directly in RVFV, we will include some noise. This would negatively impact the final CBRF. More discussions are given in our experiments. C. Re-ranking via Visual Feature Verification Although the relevant image grouping approach can find more relevant images for the query sketch, some irrelevant images may appear in the top N results. If we re-rank the top N results by measuring their similarities in the visual feature space, then the refined search results will be more satisfactory. Our aim is to filter out irrelevant images using content matching or spatial constraints [22,23,47,48], which are often used in retrieval result verifications [22–30]. Thus, in this paper, we leverage the advantages of both retrieval result verification and relevance feedback to improve the retrieval performance. We apply RVFV twice, as shown in Fig. 1. The first time reduces the number of false positive results, and the last time optimizes the final results. RVFV consists of two steps:

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

1) finding SIFT pairs of the standard image and other images; and 2) re-ranking using the similarity scores. 1) Feature Matching In this paper, RVFV is only applied to the top N initial results. We select some of the relevant images from the top N-ranked images to expand the query and get more relevant results. We find SIFT pairs of the standard image (the top-ranked image after relevant image grouping of the initial SBIR results, IS) and other images (the top-ranked N images, but not including duplicates of the standard image). The similarity scores are measured using matched SIFT point pairs. PA is a SIFT point in image IA, and PB is a SIFT point in image IB. We define (PA PB) as a SIFT pair, if and only if, the best-matched SIFT point of PA of image IA in image IB is PB, and vice versa. The similarity of two SIFT descriptors (d1 and d2) is measured using the L2-norm [32]. That is, $dis\ d\ (d1,d2) = \|d1 - d2\|2\ 2 = \sum |d1\ i - d2\ i|2i = \sum |d1\ i|2i|d\ 2\ i = 0 + \sum |d2\ i|2i|d\ 1\ i{=}0 + \sum |d1\ i - d2\ i|2i|d\ 1\ i{\neq}0,d2\ i{\neq}0 = \|d1\|2\ 2 + \|d2\|2\ 2 + \sum (|d1\ i - d2\ i|2 - |d1\ i|2 - |d2\ i|2)i|d\ 1\ i{\neq}0,d2i{\neq}0 = 2 - \sum d1\ id2\ i\ i|d1\ i{\neq}0,d2\ i{\neq}0$, (1) where $d*\ i$ is the value of d* in the i-th dimension, for i=1,□,128. $d*\ i$ is normalized using $d*\ i = d*\ i\ \|d*\|2\ 2⁄$. (2)

Thus, in (1), we have $\|d1\|2\ 2 + \|d2\|2\ 2{=}2$. According to [32], the similarity score between dAi of image

IA and dBj of image IB is defined as $Sim\ d\ (dAi,Bj) = \sum dAi\ l\ dBj\ l\ l|dAi\ l \neq 0,dBj\ l \neq 0$

, (3)

where dl denotes the value of the l-th dimension of the descriptor d. Based on (3), the similarity score is $SimN\ d\ (dAi,dBj) = Sim\ d\ 2(dAi,dBj)\ 1\ LB \sum Sim\ d\ (dAi,dBk)\ LB\ k{=}1 * 1\ LA \sum Sim\ d\ (dAk,dBj)\ LA\ k{=}1$, (4) where LA and LB are the number of SIFT points in image IA and IB, respectively. The denominator serves as a normalization, considering the average similarity between dAi and all other descriptors in image IB, and the average similarity between dBj and all other descriptors in image IA. 2) Similarity-based re-ranking SIFT feature matching has been extensively applied to image classification [30,33,34,35]. Considering the spatial locations, orientation, or other geometric constraints [36, 37] can improve matching performances. Sketch-based image retrieval has strong spatial constraints. Therefore, we use SIFT locations (L) and orientations (O) to add weights to matched

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

SIFT pairs. The weight is defined as $W(m) = \exp(-\alpha\times(W L(m)+\beta \times WO(m))$, (5) where m denotes the m-th SIFT pair between IA and IB. α controls the convergence of the exponential function, and β balances the two parts. $WL(m)$ and $WO(m)$ are the location and orientation weights, respectively. They are defined as $WL(m) = \|L(Am)-L(Bm)\|2\ 2$ (6) and $WO(m) = \min(|O(Am)-O(Bm)|,|O(Am)+O(Bm)|)$, (7) where L(.) and O(.) are the location and orientation of a SIFT point, and (Am, Bm) is the m-th SIFT pair of IA and IB. We use the minimum of the difference and the sum of orientations so that $W(m)$ is in the range [−π, π]. Then, the similarity between two images can be determined by summing the weighted scores of the matched SIFT point pairs. That is, $IM(IA,IB) = \sum Sim\ d\ (dAm,dBm)\ m\ W(m)$. (8) For the top N results of the initial retrieval (N=100 in our experiments), we compute the similarity of image Ik to the standard image IS using $k = SIM(Is,Ik)$, k= 1~N. (9) When k=1, we have Sk =1. Sk indicates how similar an image in the initial result is to the standard image. We evaluate if it satisfies a minimum matching requirement (i.e., Sk is larger than a cut-off threshold), or we sort Sk in descending order and select the top M images. The selected images are used for the contour-based relevance feedback. D. Contour-Based Relevance Feedback It is useful to expand the query for image-based retrieval to improve the final result [22]. A sketch is a description of contours. The contour of a top-ranked image can also be regarded as a sketch and used to return more relevant images. Our relevance feedback algorithm contains the following

steps. 1) The contours of the verified images are used as new query sketches. 2) Each image in the corpus is given a score based on each of the new query contours. 3) The final similarity score of each image in the corpus is obtained by combining the scores of the initial and expanded retrievals. 4) The final ranked list is generated using the initial system for each new query. These ranked lists are combined and used to add weight to the initial result and obtain the final ranked list. Assume that M relevant images are obtained through the first RVFV (N≥M). Then, CBRF finds more relevant images using the contours of the M images as new query sketches. After the above query expansion, we get ranked lists for the M-expanded query sketches.

4.CONCLUSION

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

In this paper we have given different techniques to search images from web. To refine the quality of retrieved images, various postprocessing methods have been adopted after the initial search process. Various experimental results on image re-ranking suggest that above method can improve the results returned by commercial search engines. We only take image search relevance into consideration, though image diversity is another important factor in search performance. In future work, Diversity of re-ranked images can be enhanced by duplication detection or other such methods.

REFERENCES

[1] A. Chalechale, G. Naghdy, and A. Mertins, "Edge image description using angular radial partitioning". IEEE Proceedings-Vision, Image and Signal Processing, vol. 151(2): 93–101, April, 2004.

[2] Y. Cao, C. Wang, L. Zhang, L. Zhang. "Edgel Index for Large-Scale Sketch-based Image Search". CVPR , IEEE Conference, 2011

[3] E. D. Sciascio, G. Mingolla, M. Mongiello, "Content-based image retrieval over the web using query by sketch and relevance feedback". VISUAL'99, London, UK, 1999, pp. 123–130.

[4] C. Liu, D. Wang, X. Liu, C. Wang, L. Zhang, B. Zhang, "Robust semantic sketch based specific image retrieval". ICME, 2010 IEEE.

[5] R. Datta, D. Joshi, J. Li, and J. Wang. "Image retrieval: Ideas, influences, and trends of the new age". ACM, Computing Surveys, 2008.

[6] G. Salton and C. Buckley. "Improving retrieval performance by relevance feedback". Journal of the American Society for Information Science, 41(4): 288–297, 1999

[7] I. J. Cox, M. L. Miller, T. P. Minka, T. V. Papathomas, P. N. Yianilos, "The Bayesian Image Retrieval System, PicHunter: Theory, Implementation and Psychological Experiments", IEEE Transactions on Image Processing, 9(1), pp. 20–37, 2000.

[8] E. Cheng, F. Jing and L. Zhang, "A unified relevance feedback framework for web image retrieval", IEEE Trans. Image Process., vol. 18, no. 6, pp.1350–1357, 2009.

[9] P. Salembier, F. Marqués, Region-based representations of image and video:

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

segmentation tools for multimedia services. Circuits and Systems for Video Technology, IEEE Transactions on, 1999, 9(8): 1147–1169

. [10] K. Hirata and T. Kato, “Query by visual example - content based image retrieval,” in Proc. Adv. Database Technol. 1992, pp. 56–71.