ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 3, Issue 1, pp: (352-357), Month: January - March 2015, Available at: www.researchpublish.com

Re-Ranking Web Images Based On Query-Specific Semantic Signatures

Navamani.C[1], Anantharaj[2], Chendhavarayan.R[3]

Student [1], HOD [2], Assistant Professor [3]. Computer Science Engineering,
Thiruvalluvar College of Engineering and Technology, TamilNadu, India

Abstract: **Image re-ranking, as an efficient means to progress the results of web-based image search, has been implemented by present commercial search engines such as Bing and Google. Given a query keyword, a pool of images is initially recovered based on textual information. By inquiring the user to choose a query image from the pool, the remaining images are re-ranked based on their visual matches with the query image. A chief dispute is that the similarities of visual features do not well associate with images' semantic meanings which infer users' search intention. Currently people proposed to match images in a semantic space which utilized attributes or reference classes directly connected to the semantic meanings of images as basis. Though, learning a general visual semantic space to typify highly various images from the web is tricky and ineffective. In this paper, we propose an original image re-ranking framework, which routinely offline learns dissimilar semantic spaces for diverse query keywords. The visual features of images are planned into their associated semantic spaces to obtain semantic signatures. At the online stage, images are re-ranked by comparing their semantic signatures attained from the semantic space denoted by the query keyword. The proposed query-specific semantic signatures considerably develop both the accuracy and efficiency of image re-ranking. The novel visual features of thousands of dimensions can be planned to the semantic signatures as short as 25 dimensions. Experimental results demonstrate that 25-40 percent comparative improvement has been attained on re-ranking exactness evaluated with the up to date methods.**

Keywords: **Image Search; Image Re-Ranking; Semantic Space; Semantic Signature; Neural Network; Keyword Expansion.**

I. INTRODUCTION

WEB-SCALE image search engines typically use keywords as queries and believe on adjacent text to search images. They experience from the uncertainty of query keywords, since it is inflexible for users to precisely depict the visual content of target images only exploiting keywords. For an instance, exploiting "apple" as a query keyword, the recovered images belong to diverse groups (besides called concepts), such as "red apple," "apple logo," and "apple laptop." For resolving the uncertainty, content- based image retrieval [1], [2] with relevance feedback [3], [4], [5] is extensively employed. It obliges users to choose numerous related and unrelated image examples, from which visual resemblance metrics are studied through online training. Images are re-ranked based on the studied visual similarities. Conversely, web-scale commercial systems require the users' feedback to be the least without online training.

An efficient manner to develop search results is online image re-ranking [6], [7], [8], which limits users' effort to just one-click feedback and its interaction is also easy enough. Most web image search engines have implemented this approach [8]. It is illustrated in figure 1. A group of images related to the query keyword are recovered by the search engine for a given query keyword input, according to a stored word-image index file. In general, the size of the returned image pool is preset. The user is inquired to choose a query image for reflecting the user's search intention; from the pool, the remaining images in the pool are re-ranked based on their visual resemblances with the query image. The word image index file and visual features of images are pre-calculated offline and stored. The foremost online computational cost is

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 3, Issue 1, pp: (352-357), Month: January - March 2015, Available at: www.researchpublish.com

the comparison of visual features. To attain high effectiveness, the visual feature vectors require being small and their matching requires being quick. Various well-liked visual features are in high dimensions and efficiency is not adequate if they are straightly matched.

One more main challenge is that, without online training, the resemblances of low-level visual features may not well associate with high-level semantic meanings of images which infer search intention of users. Furthermore, low-level features are occasionally not consistent with visual observation. For an instance, if images of the similar object are captured from diverse viewpoints, under dissimilar lightings or even with diverse compression artifacts, their low-level characteristics may vary considerably, even though humans assume the visual content does not alter much. To diminish this semantic hole and inconsistency with visual perception, there have been a number of studies to map visual features to a set of predefined concepts or characteristics as semantic signatures [9], [10], [11], [12]. According to our experimental study, images recovered by 120 query keywords only comprise more than 1,500 concepts. It is tricky and ineffective to devise an enormous concept dictionary to describe extremely different web images. As the topics of web images vary dynamically, it is attractive that the concepts and attributes can be involuntarily found instead of being manually described. The conventional image re-ranking framework is shown in fig. 1.

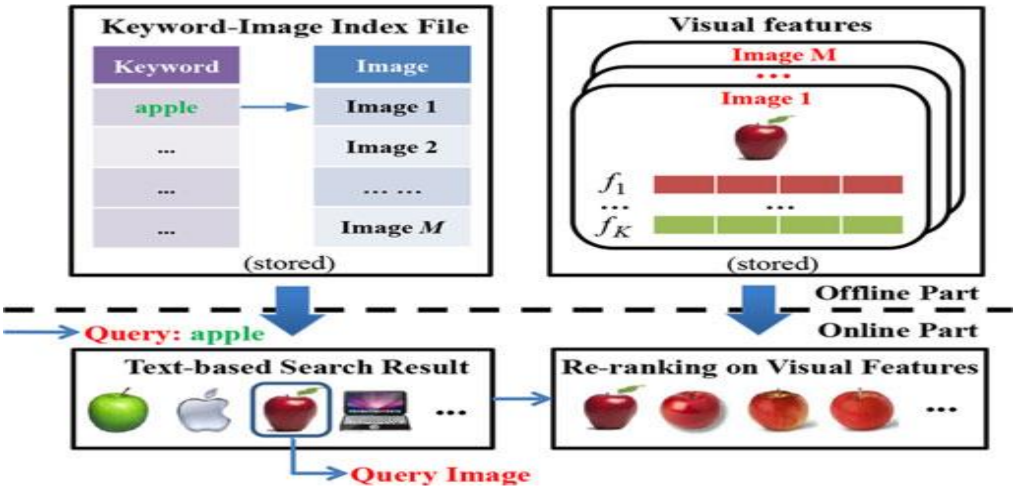


Fig. 1. The conventional image re-ranking framework

II. RELATED WORK

The chief part of image re-ranking is to calculate visual similarities revealing semantic relevance of images. Numerous visual features [13], [14], [15] have been built up in current years. Yet, for diverse query images, the efficient low-level visual features are dissimilar. For reducing the semantic gap, query-specific semantic signature was initially proposed in [16]. Kuo et al. [17] newly expanded each image with appropriate semantic features through transmission over a visual graph and a textual graph which were interrelated. One more way of learning visual similarities exclusive of adding users' trouble is pseudo relevance feedback [18], [19]. It gets the top N images most visually similar to the query image as enlarged positive instances to learn a similarity metric. As the top N images are not essentially semantically-consistent with the query image, the studied similarity metric may not consistently reveal the semantic relevance and may even depreciate re-ranking performance.

Krapac et al. [20] proposed generic classifiers based on query-relative features which could be exploited for original query keywords not including extra training. Cai et al. [21] re-ranked images with attributes which were physically defined and learned from manually tagged training samples. These schemes believed that there was one main semantic class under a query keyword. Images were re-ranked by forming this leading group with visual and textual features. Due to the uncertainty of query keywords, there may be numerous semantic categories under one keyword query. Without query images chosen by users, these schemes cannot precisely detain users' search intention.

The classifiers of concepts, attributes, and reference classes are prepared from well-known classes with labeled cases. However the knowledge gained from the known classes can be moved to distinguish samples of novel classes which have little or even no training samples. As these concepts, attributes, and reference classes are described with semantic

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 3, Issue 1, pp: (352-357), Month: January - March 2015, Available at: www.researchpublish.com

meanings, the projections over them can well detain the semantic meanings of novel images even devoid of additional training. Rasiwasia et al. [9] mapped visual features to a common concept dictionary for image retrieval. Lampert et al. [10] pre-described a set of attributes on an animal database and sensed target objects based on a mixture of human-indicated attributes as an alternative of training images. Sharmanska et al. [22] augmented this illustration with extra dimensions and permitted a smooth transition between zero-shot learning, unsupervised training and supervised training. Parikh and Grauman [23] proposed relative attributes to specify the strength of an attribute in an image with respect to additional images. Parkash and Parikh [60] employed attributes to direct active learning.

Torresani et al. [24] proposed an image descriptor which was the output of a number of classifiers on a set of identified image classes, and exploited it to match images of extra not related visual classes. In the present schemes, every concepts/attributes/reference-classes are commonly applied to all the images and they are physically defined. For modeling all the web images, an enormous set of concepts or reference classes are entailed, which is unrealistic and unsuccessful for online image re-ranking. Instinctively, only a small subset of the concepts is appropriate to a particular query. Numerous concepts unrelated to the query not only raise the computational cost however also depreciate the accurateness of re-ranking. Though, how to involuntarily locate such related concepts and exploit them for online web image re-ranking was not well discovered in earlier studies.

III. PROPOSED SYSTEM

The illustration of our proposed system is shown in fig. 2. It has offline and online parts. At the offline stage, the reference classes (which represent different concepts) corresponding to query keywords are routinely discovered and their training images are repeatedly gathered in numerous steps. For a query keyword (e.g., “apple”), a set of much related keyword expansions (such as “red apple” and “apple macbook”) are automatically picked utilizing both textual and visual information. This set of keyword expansions describes the reference classes for the query keyword. For automatically attaining the training examples of a reference class, the keyword expansion (e.g., “red apple”) is exploited to recover images by the search engine based on textual information yet again. Images recovered by the keyword expansion (“red apple”) are much less different than those recovered by the novel keyword (“apple”). After routinely eliminating outliers, the recovered top images are employed as the training instances of the reference class.

Some reference classes (such as “apple laptop” and “apple macbook”) have related semantic meanings and their training sets are visually related. So as to develop the efficiency of online image re-ranking, redundant reference classes are eliminated. To better evaluate the resemblance of semantic signatures, the semantic correlation between reference classes is evaluated with a web-based kernel function. For every query keyword, its reference classes forms the foundation of its semantic space. Neural Network classifier on visual and textual features is trained from the training sets of its reference classes and stored offline. Under a query keyword, the semantic signature of an image is mined by calculating the resemblances between the image and the reference classes of the query keyword exploiting the trained Neural Network classifier. If there are K varieties of visual/textual features, such as color, texture, and shape, one could merge them jointly to train a single classifier, which mines one semantic signature for an image. It is also probable to train a separate classifier for every type of features. Then, the NN based on dissimilar types of features extract K semantic signatures, which are joined at the later stage of image matching. According to the word-image index file, an image may be related with multiple query keywords, which have dissimilar semantic spaces. As a result, it may have diverse semantic signatures. The query keyword input by the user makes a decision about which semantic signature to prefer.

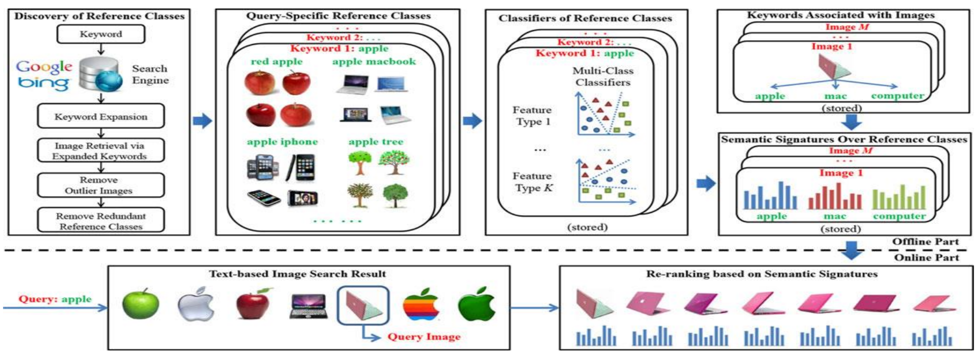


Fig. 2. Diagram of our new image re-ranking framework

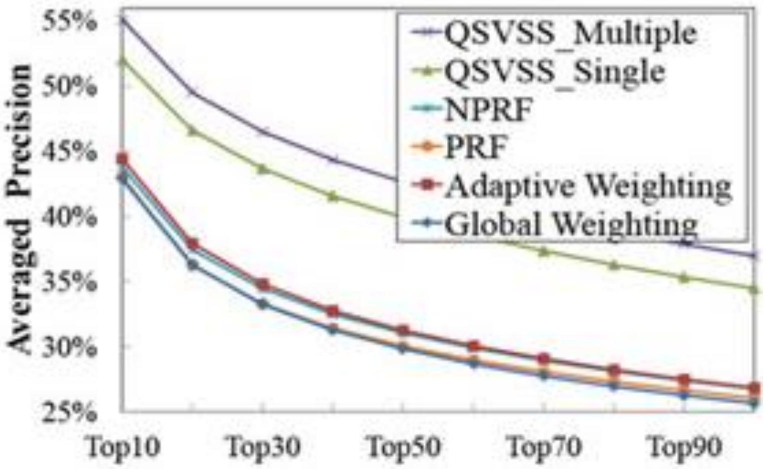
Research Publish Journals

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 3, Issue 1, pp: (352-357), Month: January - March 2015, Available at: www.researchpublish.com

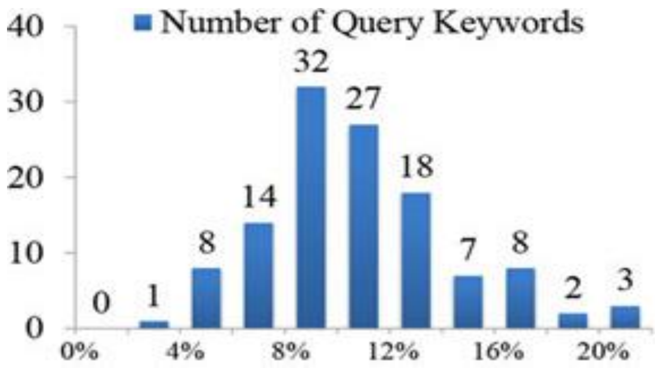
As an instance shown in fig. 2, an image is associated with three keywords "apple," "mac" and "computer." When employing any of the three keywords as query, this image will be recovered and re-ranked. Yet, under dissimilar query keywords, diverse semantic spaces are used. As a result an image could have numerous semantic signatures gained in dissimilar semantic spaces. They all required to be calculated and stored offline. At the online stage, a pool of images is retrieved by the search engine according to the query keyword. As all the images in the pool are related with the query keyword according to the word-image index file, they all have pre-calculated semantic signatures in the similar semantic space indicated by the query keyword. Once the user selects a query image, these semantic signatures are employed to calculate image resemblances for re-ranking. The semantic correlation of reference classes is integrated when calculating the resemblances.

IV. EXPERIMENTAL RESULTS

The images for testing the performance of re-ranking and the training images of reference classes can be gathered at diverse time (because the update of reference classes may be interrupted) and from dissimilar search engines. Given a query keyword, 1,000 images are recovered from the entire web exploiting a search engine. In our data set 1,20,000 testing images for re-ranking were gathered from the Bing Image Search with 120 query keywords. These query keywords cover various themes including animals, plants, food, places, people, events, objects, and scenes, etc. The training images of reference classes were also gathered from the Bing Image Search around the similar instant.



(a)



(b)

Fig. 3. (a) Averaged top m precisions. (b) Histograms of improvements of averaged top 10 precisions by comparing QSVSS Multiple with Adaptive Weighting

We evaluate with two image re-ranking schemes employed in [6], which openly evaluate visual features, and two schemes of pseudo-relevance feedback [18], [19], which online learns visual similarity metrics.

Global weighting: Fixed weights are implemented to combine the distances of diverse visual features.

Adaptive weighting: Cui et al. [6] proposed adaptive weights for query images to fuse the distances of diverse visual features. It is implemented by Bing Image Search.

PRF: The pseudo-relevance feedback scheme proposed in [18]. It utilized top-ranked images as positive instances to train a one-class NN.

NPRF: The pseudo-relevance feedback scheme proposed in [19]. It employed top-ranked images as positive cases and bottom-ranked images as negative instances to train Neural Network. For our scheme, two dissimilar methods of comparing semantic signatures are compared.

Query-specific visual semantic space employing single signatures (QSVSS Single): For an image, a particular semantic signature is evaluated from one NN classifier trained by combining all types of visual features.

Query-specific visual semantic space employing multiple signatures (QSVSS Multiple): For an image, multiple semantic signatures are calculated from multiple NN classifiers, each of which is trained on one category of visual features individually.

V. CONCLUSIONS AND FUTURE WORK

A new framework is proposed, which studies query-specific semantic spaces to considerably develop the efficiency of online image re-ranking. The visual features of images are planned into their associated semantic spaces to obtain semantic signatures offline. The obtained semantic signatures can be 70 times shorter than the novel visual features, while attain 25-40 percent relative progress on re-ranking precisions over state-of-the-art techniques. In the future work, our framework can be advanced along numerous directions. Discovering the keyword expansions employed to describe reference classes can include other metadata and log data as well as the textual and visual characteristics. For an instance, the co-occurrence information of keywords in user queries is helpful and can be attained in log data. Although the semantic signatures are previously little, it is probable to make them more solid and to additionally improve their matching efficiency exploiting further technologies such as hashing.

REFERENCES

[1] R. Datta, D. Joshi, and J.Z. Wang, “Image Retrieval: Ideas, Influences, and Trends of the New Age,” ACM Computing Surveys, vol. 40, article 5, 2007.

[2] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, “Content-Based Image Retrieval,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, Dec. 2000.

[3] Y. Rui, T.S. Huang, M. Ortega, and S. Mehrotra, “Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval,” IEEE Trans. Circuits and Systems for Video Technology, vol. 8, no. 5, pp. 644-655, Sept. 1998.

[4] X.S. Zhou and T.S. Huang, “Relevance Feedback in Image Retrieval: A Comprehensive Review,” Multimedia Systems, vol. 8, pp. 536-544, 2003.

[5] D. Tao, X. Tang, X. Li, and X. Wu, “Asymmetric Bagging and Random Subspace for Support Vector Machines-Based Relevance Feedback in Image Retrieval,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 28, no. 7, pp. 1088-1099, July 2006.

[6] J. Cui, F. Wen, and X. Tang, “Real Time Google and Live Image Search Re-Ranking,” Proc. 16th ACM Int’l Conf. Multimedia, 2008.

[7] J. Cui, F. Wen, and X. Tang, “Intent Search: Interactive on-Line Image Search Re-Ranking,” Proc. 16th ACM Int’l Conf. Multimedia, 2008.

[8] X. Tang, K. Liu, J. Cui, F. Wen, and X. Wang, “Intent Search: Capturing User Intention for One-Click Internet Image Search,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 34, no. 7, pp. 1342-1353, July 2012.

[9] N. Rasiwasia, P.J. Moreno, and N. Vasconcelos, “Bridging the Gap: Query by Semantic Example,” IEEE Trans. Multimedia, vol. 9, no. 5, pp. 923-938, Aug. 2007.

[10] C. Lampert, H. Nickisch, and S. Harmeling, “Learning to Detect Unseen Object Classes by Between-Class Attribute Transfer,” Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

[11] Q. Yin, X. Tang, and J. Sun, "An Associate-Predict Model for Face Recognition," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011.

[12] .A. Kovashka, D. Parikh, and K. Grauman, "Whittle Search: Image Search with Relative Attribute Feedback," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012.

[13] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60, pp. 91-110, 2004.

[14] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005.

[15] Y. Cao, C. Wang, Z. Li, L. Zhang, and L. Zhang, "Spatial-Bag-of- Features," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010.

[16] X. Wang, K. Liu, and X. Tang, "Query-Specific Visual Semantic Spaces for Web Image Re-Ranking," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010.

[17] Y. Kuo, W. Cheng, H. Lin, and W. Hsu, "Unsupervised Semantic Feature Discovery for Image Object Retrieval and Tag Refinement," IEEE Transactions on Multimedia, Vol. 14, No. 4, pp. 1079-1090, August 2012.

[18] R. Yan, E. Hauptmann, and R. Jin, "Multimedia Search with Pseudo-Relevance Feedback," Proc. Int'l Conf. Image and Video Retrieval, 2003.

[19] R. Yan, A.G. Hauptmann, and R. Jin, "Negative Pseudo-Relevance Feedback in Content-Based Video Retrieval," Prof. ACM International Conference Multimedia, 2003.

[20] L. Yang and A. Hanjalic, "Supervised Reranking for Web Image Search," Proceedings ACM International Conference on Multimedia, 2010.

[21] J. Krapac, M. Allan, J. Verbeek and F. Jurie, "Improving Web Image Search Results Using Query-Relative Classifiers," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010.

[22] J. Cai, Z. Zha, W. Zhou, and Q. Tian, "Attribute-Assisted Reranking for Web Image Retrieval," Proceedings of 20th ACM International Conference on Multimedia, 2012.

[23] V. Sharmanska, N. Quadrianto and C.H. Lampert, "Augmented Attribute Representations," Proceedings of 12th European Conference Computer Vision (ECCV), 2012.

[24] D. Parikh and K. Grauman, "Relative Attributes," Proceedings of International Conference Computer Vision (ICCV), 2011.