International Journal of Research

Available at https://edupediapublications.org/journals

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 13
September 2016

Image Reranking Using Semantic Signatures

Ms.Ch.Anusha, **Mr.K.SIVAREDDY**

ABSTRACT:

Image re-ranking, as an effective way to improve the results of web-based image search, has been adopted by current commercial search engines such as Bing and Google. Given a query keyword, a pool of images is first retrieved based on textual information. By asking the user to select a query image from the pool, the remaining images are re-ranked based on their visual similarities with the query image. A major challenge is that the similarities of visual features do not well correlate with images' semantic meanings which interpret users' search intention. Recently people proposed to match images in a semantic space which used attributes or reference classes closely related to the semantic meanings of images as basis. However, learning a universal visual semantic space to characterize highly diverse images from the web is difficult and inefficient. In this paper, we propose a novel image re-ranking framework, which automatically offline learns different semantic spaces for different query keywords. The visual features of images are projected into their related semantic spaces to get semantic signatures. At the online stage, images are re-ranked by comparing their semantic signatures obtained from the semantic space specified by the query keyword. The proposed query-specific semantic signatures significantly improve both the accuracy and efficiency of image re-ranking. The original visual features of thousands of dimensions can be projected to the semantic signatures as short as 25 dimensions. Experimental results show that 25-40 percent relative improvement has been achieved on re-ranking precisions compared with the state-of-the-art methods.

INTRODUCTION

WEB-SCALE image search engines mostly use keywords as queries and rely on surrounding text to search images. They suffer from the ambiguity of query keywords, because it is hard for users to accurately describe the visual content of target images only using keywords. For example, using "apple" as a query keyword, the retrieved images belong to different

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 13
September 2016

categories (also called concepts in this paper), such as "red apple," "apple logo,"
and "apple laptop." In order to solve the ambiguity, content- based image retrieval [1], [2] with relevance feedback [3], [4], [5] is widely used. It requires users to select multiple relevant and irrelevant image examples, from which visual similarity metrics are learned through online training. Images are re-ranked based on the learned visual similarities. However, for web-scale commercial systems, users' feedback has to be limited to the minimum without online training.

Our Approach

In this paper, a novel framework is proposed for web image re-ranking. Instead of manually defining a universal concept dictionary, it learns different semantic spaces for different query keywords individually and automatically. The semantic space related to the images to be re-ranked can be significantly narrowed down by the query keyword provided by the user. For example, if the query keyword is
"apple," the concepts of "mountain" and "Paris" are irrelevant and should be excluded. Instead, the concepts of "computer" and "fruit" will be used as dimensions to learn the semantic space related to "apple." The query-specific semantic spaces can more accurately model the images to be re-ranked, since they have excluded other potentially unlimited number of irrelevant concepts, which serve only as noise and deteriorate the re-ranking performance on both accuracy and computational cost.

RELATED WORK

The key component of image re-ranking is to compute visual similarities reflecting semantic relevance of images. Many visual features [36], [37], [38], [39], [40] have been developed in recent years. However, for different query images, the effective low-level visual features are different. Therefore, Cui et al. [6], [7] classified query images into eight predefined intention c categories and gave different feature weighting schemes to different types of query images. But it was difficult for the eight weighting schemes to cover the large diversity of all the web images. It was also likely for a query image to be classified to a wrong category. In order to reduce the semantic gap, query-specific semantic signature was first proposed in [41]. Kuo et al. [42] recently augmented each image with relevant semantic features through propagation over a visual graph and a textual graph which were correlated. Another way of learning visual similarities without adding users' burden is pseudo relevance feedback [43], [44], [45]. It takes the top N images most visually similar to the query image as expanded positive examples to learn a similarity metric. Since the top N images are not necessarily semantically-consistent with the query image, the learned similarity metric may not reliably reflect the semantic relevance and may even deteriorate re-ranking performance. In object retrieval, in order to purify the expanded positive examples, the spatial configurations of local visual features are verified [46], [47], [48]. But it is not applicable to general web image search, where relevant images may not contain the same object.

APPROACH OVERVIEW

The diagram of our approach is shown in Fig. 3. It has offline and online parts. At the offline stage, the reference classes (which represent different concepts) related to query keywords are automatically discovered and their training images are automatically collected in several steps. For a query keyword (e.g., "apple"), a set of most relevant keyword expansions (such as "red apple" and "apple macbook") are automatically selected utilizing both textual and visual information. This set of keyword expansions defines the reference classes for the query keyword. In order to automatically obtain the training examples of a reference class, the keyword expansion (e.g., "red apple") is used to retrieve images by the search engine based on textual information again. Images retrieved by the keyword expansion ("red apple") are much less diverse than those retrieved by the original keyword ("apple"). After automatically removing outliers, the retrieved top images are used as the training examples of the reference class. Some reference classes (such as "apple laptop" and "apple macbook") have similar semantic meanings and their training sets are visually similar. In order to improve the efficiency of online image re-ranking, redundant reference classes are removed. To better measure the similarity of semantic signatures, the semantic correlation between reference classes is estimated with a web-based kernel function. For each query keyword, its reference classes forms the basis of its semantic space. A multi-class 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 extracted by computing the similarities between the image and the reference classes of the query keyword using the trained multiclass classifier. If there are K types of visual/textual features, such as color, texture, and shape, one could combine them together to train a single classifier, which extracts one semantic signature for an image. It is also possible to train a separate classifier for each type of features. Then, the K classifiers based on different types of features extract K semantic signatures, which are combined at the later stage of image matching. Our experiments show that the latter strategy can increase the re-ranking accuracy at the cost of storage and online matching efficiency because of the increased size of semantic signatures.

SEMANTIC SIGNATURES

Given M reference classes for keyword q and their training images, a multi-class classifier on the visual features of images is trained and it outputs an M-dimensional vector p, indicating the probabilities of a new image I belonging to different reference classes. p is used as the semantic signature of I. The distance between two images Ia and Ib are measured as the L1-distance between their semantic signatures pa and pb,

CONCLUSION AND FUTURE WORK

We propose a novel framework, which learns query-specific semantic spaces to significantly improve the effectiveness and efficiency of online image re-ranking. The visual features of images are projected into their related semantic spaces automatically learned through keyword expansions offline. The extracted semantic signatures can be 70 times shorter than the original visual features, while achieve 25-40 percent relative improvement on reranking precisions over state-of-the-art methods.

In the future work, our framework can be improved along several directions. Finding the keyword expansions used to define reference classes can incorporate other metadata and log data besides the textual and visual features.For example, the co-occurrence information of keywords in user queries is useful and can be obtained in log data. In order to update the reference classes over time in an efficient way, how to adopt incremental learning [72] under our framework needs to be further investigated.

Although the semantic signatures are already small, it is possible to make them more compact and to further enhance their matching efficiency using other technologies such as hashing [76]

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," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2009.

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

[12] A. Kovashka, D. Parikh, and K. Grauman, "WhittleSearch: Image Search with Relative Attribute Feedback," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2012.

[13] R. Fergus, P. Perona, and A. Zisserman, "A Visual Category Filter for Google Images," Proc. Eighth European Conf. Computer Vision (ECCV), 2004.

[14] R. Fergus, L. Fei-Fei, P. Perona, and A. Zisserman, "Learning Object Categories from Google's Image Search," Proc. 10th IEEE Int'l Conf. Computer Vision (ICCV), 2005.

[15] W. Hsu, L. Kennedy, and S.F. Chang, "Video Search Reranking via Information Bottleneck Principle," Proc. 14th Ann. ACM Int'l Conf. Multimedia, 2006.

[16] F. Schroff, A. Criminisi, and A. Zisserman, "Harvesting Image Databases from the Web," Proc. IEEE 11th Int'l Conf. Computer Vision (ICCV), 2007.

[17] W. Hsu, L. Kennedy, and S.F. Chang, "Video Search Reranking through Random Walk over Document-Level Context Graph," Proc. ACM 15th Int'l Conf. Multimedia, 2007.

[18] M. Fritz and B. Schiele, "Decomposition, Discovery and Detection of Visual Categories Using Topic Models," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2008.

[19] T. Berg and D. Forsyth, "Animals on the Web," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2008.

[20] D. Grangier and S. Bengio, "A Discriminative Kernel-Based Model to Rank Images from Text Queries," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30, no. 8, pp. 1371-1384, Aug. 2008.

[21] Y. Jing and S. Baluja, "Visual Rank: Applying Page Rank to Large-Scale Image Search," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30, no. 11, pp. 1877-1890, Nov. 2008.

[22] L. Yang and A. Hanjalic, "Supervised Reranking for Web Image Search," Proc. ACM Int'l Conf. Multimedia, 2010.

[23] L. Chen, D. Xu, I. Tsang, and J. Luo, "Tag-Based Web Photo Retrieval Improved by Batch Mode Re-Tagging," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2010.

[24] X. Tian, L. Yang, J. Wang, X. Wu, and X. Hua, "Bayesian Visual Reranking," IEEE Trans. Multimedia, vol. 13, no. 4, pp. 639-652, Aug. 2011.

[25] B. Geng, L. Yang, C. Xu, and X. Hua, "Content-Aware Ranking for Visual Search," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2010.

[26] 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.

[27] W. Liu, Y. Jiang, J. Luo, and F. Chang, "Noise Resistant Graph Ranking for Improved web Image Search," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2011.

[28] N. Morioka and J. Wang, "Robust Visual Reranking via Sparsity and Ranking Constraints," Proc. ACM Int'l Conf. Multimedia, 2011.

[29] V. Jain and M. Varma, "Learning to Re-Rank: Query-Dependent Image Re-Ranking Using Click Data," Proc. 20th Int'l Conf. World Wide Web (WWW), 2011.

[30] J. Huang, X. Yang, X. Fang, and R. Zhang, "Integrating Visual Saliency and Consistency for Re-Ranking Image Search Results," IEEE Trans. Multimedia, vol. 13, no. 4, pp. 653-661, Aug. 2011.

[31] 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.

[32] J. Cai, Z. Zha, W. Zhou, and Q. Tian, "Attribute-Assisted Reranking for Web Image Retrieval," Proc. 20th ACM Int'l Conf. Multimedia,

2012.

[33] M. Wang, L. Yang, and X. Hua, "MSRA-MM: Bridging Research and Industrial Societies for Multimedia Information Retrieval," technical report, Microsoft Research Asia, 2009.

[34] T. Deselaers, T. Gass, P. Dreuw, and H. Hey, "Jointly Optimising Relevance and Diversity in Image Retrieval," Proc. ACM Int'l Conf. Image and Video Retrieval, 2009.

[35] J.J. Foo, J. Zobel, and R. Sinha, "Clustering Near-Duplicate Images in Large Collections," Proc. Int'l Workshop Multimedia Information Retrieval, 2007.

[36] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Int'l J. Computer Vision, vol. 60, pp. 91-110, 2004.

[37] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2005.

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

[39] J. Philbin, M. Isard, J. Sivic, and A. Zisserman, "Descriptor Learning for Efficient Retrieval," Proc. 11th European Conf. Computer Vision (ECCV), 2010.

[40] Y. Qiao, W. Wang, N. Minematsu, J. Liu, M. Takeda, and X. Tang, "A Theory of Phase Singularities for Image Representation and Its Applications to Object Tracking and Image Matching," IEEE Trans. Image Processing, vol. 18, no. 10, pp. 2153-2166, Oct. 2009.

[41] 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.

[42] Y. Kuo, W. Cheng, H. Lin, and W. Hsu, "Unsupervised Semantic Feature Discovery for Image Object Retrieval and Tag Refinement," IEEE Trans. Multimedia, vol. 14, no. 4, pp. 1079-1090, Aug. 2012.

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 13
September 2016

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

[44] R. Yan, A.G. Hauptmann, and R. Jin, "Negative Pseudo-Relevance Feedback in Content-Based Video Retrieva

Author's profile:

Mr. K.SIVA REDDY received M.Tech(CSE) Degree from Acharya Nagarjuna University & M.B.A.(HR)Degree from Shathinikethan P.G. College, affiliated to Osmania University, Hyderabad. He is currently working as Assistant Professor in the Department of Computer Science and Engineering in Nalgonda Institute of Technology and Science, Nalgonda, Telangana, India. His interests includes Computer Organization, Computer Networking, Software Engineering, Software Project Management, Software Testing Methodologies, Cloud Computing and Management Science.

Ms. Ch.Anusha received B.Tech Degree from Nalgonda Institute of Technology and Science in Nalgonda. She is currently pursuing M.Tech Degree in Computer Science and Engineering specialization in Nalgonda Institute of Technology and Science, Nalgonda, Telangana, India.