BVG Trust J. Sci. Trans. Environ. Technov. 2017, 10(3) : 146 - 150 Scientific Transactions in Environment and Technovation

Optimized web image search using meta annotation re-ranking technique

https://doi.org/10.56343/STET.116.010.003.009
http://stetjournals.com

A. Ranjitha[1], V. Geetha[1] and M.V. Srinath[2]

[1]Department of Computer Science, S.T.E.T. Women's College, Sundarakkottai, Mannargudi, Thiruvarur, Tamil Nadu, India.

[2]Department of Master of Computer Application, S.T.E.T. Women's College, Mannargudi, Thiruvarur, Tamil Nadu, India.

Abstract

An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Given a textual query in Traditional text Based Image Retrieval (TBIR),relevant images are to be re ranked using visual features after the initial text based image search. In this paper, a new meta annotation based re-ranking framework for large scale TBIR has been proposed. This problem has been computed on basis of Multiple Instance Learning and Generalized Multiple Instance (GMI) learning method. To address the ambiguities on the instance labels in the positive and negative bags GMI settings have been proposed. Also the user log performs the operation of individual user interaction with the system which improves the performance of image retrieval.

Keywords: Multiple Instance, Generalized Multiple Instance, Image Re- ranking , Text Based Image Retrieval

Received : July 2015 Revised and Accepted : January 2017

INTRODUCTION

In the context of ever-growing number of images on the internet (such as in the online photo sharing Website, the online photo forum, and soon), retrieving relevant images from a large collection of database images has become an important research topic. During the period of last decade, the image ranking as an effective way to improve the results of Web based image search has been adopted by current commercial search engines. Given a query keyword a pool of images are re-ranked by the search engines based on the query. By asking the user to select a particular image from the pool, the remaining images are re-ranked based on the user selected image. To avoid the ambiguities in the re-ranked process and to achieve an effective and efficient re-ranking process a Bag Based Re-ranking approach has been introduced.

Specifically, a user is required to input a keyword as a textual query to the retrieval system. Then, the system returns the ranked relevant images whose surrounding text contains the query keyword, and the ranking score is obtained according to some similarity measurements (such as cosine distance) between the query keyword and the textual features of relevant images. However, the retrieval performance can be very poor, particularly when the textual features of the Web images are sparse and noisy in a high-dimensional space. To solve this problem, many image re-ranking methods have been developed to re-rank the initially retrieved images using visual features (Cui, *et al.*, 2008) proposed a method called Web Search Exploiting Image Contents (WEBSEIC), which uses Kernel Density Estimation (KDE) based on visual features to re-rank the retrieved relevant images. After that, an image-based ranking of Web pages is generated, and the final search result is obtained by combining with the original text-based search result. Florian Schroff *et al.* (2011) presented a re-ranking method via the information bottleneck principle based on mutual information. In their work, they first clustered the initially retrieved images together with some irrelevant images by using a so-called sequential information bottle neck clustering method (Hus, *et al.*,2006). Then, a cluster probability is obtained for cluster ranking. Finally, KDE based on visual features is used to re-rank the relevant images within each cluster. Several graph-based re-ranking methods (Hsu, *et al.*, 2007) have been also developed. The basic idea is to construct a graph representing the local similarity of visual features of images for re-ranking. However, the similarity of low-level visual features among the unconstrained Web images may not reflect the high-level semantic concepts of Web images due to the semantic gap. Moreover, this re-ranking paradigm does not consider label information and can only achieve limited improvements.

RELATED WORK

Web-scale image search engines mostly use keywords as queries and rely on surrounding text to search images. It is well known that they suffer from the ambiguity of query keywords (Jen-hao Hsiao, 2008).

*Corresponding Author :
email: *ranjitha.317@gmail.com*

For example, using "apple" as query, images are retrieved which are belonged to different categories, such as "red apple", "apple logo", and "apple laptop". Online image re-ranking has been shown to be an effective way to improve the image search results (Jing and Baluja, 2008). Partition the relevant images into clusters by using visual and textual features. Multimedia and Image (MI) learning method explains this problem. Automatic bag annotation method performs much better re ranking than existing image re ranking method (Xiao gang Wang *et al.*, 2003). A real-time textual query based personal photo retrieval system has been performed by leveraging millions of web images and their associated rich textual descriptions. After a user provides a textual query (*e.g.*, "water"), our system exploits the inverted file to automatically find the positive web images that are related to the textual query "water" as well as the negative web images that are irrelevant to the textual query. Based on these methods it will automatically retrieve relevant and irrelevant web images (Jing and Baluja, 2008). Database containing a large number of images and with high precision is still an arduous manual task. To overcome the download restriction, it will use web search instead of an image search. This search can generate thousands of images. Since the objective in this work is to harvest a large number of images of a particular class automatically, and to achieve this with high precision .

RF increases the retrieval performance to the fact that it enables the system to learn as to what is relevant or irrelevant to the user across successive retrieval-feedback cycles (Kelley,1960). Bag annotation method can achieve better retrieval performance (Kim,2008; Hus *et al.*, 2006). This process is similar to the system proposed by Li *et al.* (2009). To facilitate annotation process the bag ranking score can be used. One leading framework for image object mining is the bag of words approach. The main motive is to encode an image as a collection of visual words of the quantized features. Here they use pseudo positive images produced in response to the original query. It gradually refines the query language model. Most of the work have emphasized on extracting information from data stored in database.

MI learning methods have been proposed to solve learning problems with ambiguity on training samples. In the traditional supervised learning problems, there is clear knowledge on the labels of training samples. In contrast, in MI learning problems, a label only accompanies each training "bag," which consists of several instances (i.e., training samples). Specifically, in the traditional setting of MI learning problems, each positive bag has at least one positive instance, while a negative bag has no positive instances. MI learning methods (Li, *et al.*, 2009) learn models from the training data with such ambiguous label information and predict the label of test bags or instances. Relevance feedback schemes (Liu, 2011) using linear/quadratic estimators have been applied in content-based image retrieval to significantly improve retrieval performance. One major difficulty in relevance feedback is to estimate the support of target images in high dimensional feature space with a relatively small number of training samples. In this paper, we develop a novel scheme based on one class SVM, which fits a tight hyper-sphere in the nonlinearly transformed feature space to include most of the target images based on the positive examples. The use of kernel provides us an elegant way to deal with nonlinearity in the distribution of the target images, while the regularization term in Service Module (SVM) provides good generalization ability.

To validate the efficacy of the proposed approach, we test it on both synthesized data and real-world images. Promising results are achieved in both cases. The ranking models of existing image/video search engines are generally based on associated text while the visual content is actually neglected. Imperfect search results frequently appear due to the mismatch between the textual features and the actual visual content. Visual re-ranking (Li *et al.*, 2009), in which visual information is applied to refine text based search results, has been proven to be effective. However, the improvement brought by visual re-ranking is limited, and the main reason is that the errors in the text-based results will propagate to the refinement stage. In this paper, we propose a Content-Aware Ranking model based on "learning to rank" framework, in which textual and visual information are simultaneously leveraged in the ranking learning process. We formulate the Content-Aware Ranking based on large margin structured output learning by modelling the visual information into a regularization term.

The direct optimization of the learning problem is nearly infeasible since the number of constraints is huge. The efficient cutting plane algorithm is adopted to learn the model by iteratively adding the most violated constraints. Extensive experimental results on a large-scale dataset collected from a commercial Web image search engine, as well as the TRECVID 2007 video search dataset, demonstrate the proposed ranking model significantly outperforms the state-of-the art ranking and re-ranking methods. One of the fundamental problems in Content-Based Image Retrieval (CBIR) has been the gap between low-level visual features and high-level semantic concepts. To narrow down this gap, relevant feedback is introduced into image retrieval.

With the user-provided information, a classifier can be learned to distinguish between positive and negative examples. However, in real-world applications, the number of user feedbacks is usually too small compared to the dimensionality of the image space. In order to cope with the high dimensionality, we propose a novel semi supervised method for dimensionality reduction called Maximum Margin Projection (MMP) (MattiaBroilo, 2008). MMP aims at maximizing the margin between positive and negative examples at each local neighbourhood. Different from traditional dimensionality reduction algorithms such as Principal Component Analysis (PCA) and Linear Discriminate Analysis (LDA), which effectively see only the global Euclidean structure, MMP is designed for discovering the local manifold structure. Therefore, MMP is likely to be more suitable for image retrieval, where nearest neighbour search is usually involved. After projecting the images into a lower dimensional subspace, the relevant images get closer to the query image; thus, the retrieval performance can be enhanced.

PROPOSED DESIGN

After the user log in, initially user log displays the information about the previous user recently searched images. From that a particular query selected by the user or a new query given by the user which retrieves images from the database. In the existing system, classification of images can be displayed by means of semantic signature. In our approach visual and textual features can be compared with the user selected image

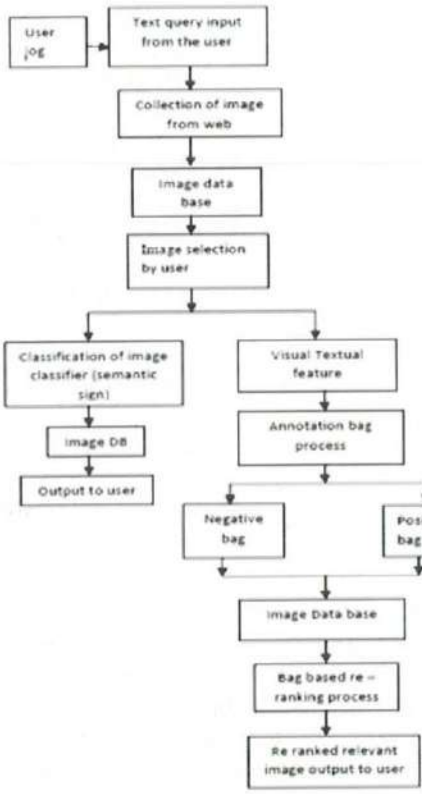

Fig. 1. System architecture

by means of shape, colour and texture. In the annotation bag process K-means algorithm can be used to split the positive and negative bags which contain relevant and irrelevant images respectively.

Here positive bag is only considered for re-ranking operation, this can send to the bag based re-ranking process which perform GMI learning method. This will provide an effective and efficient re-ranked image output to the user. Generally it performs the following five operations

i. construction of web image repository,

ii. discovery of reference classes,

iii. bag annotation process,

iv. image Re-Ranking, and

v. performance comparison.

Construction of web image repository

Initially user log is performed which provide an authentication to each user. Depends on the user given query, collection of the images can be stored in the database

Discovery of Reference classes

Here the user given query relevant words are searched and it provides an easy access for relevant images. It will contain maximum of three relevant words which is relevant to the query word. This process contains two operations such as keyword expansion and select reference classes. In the keyword expansion, the relevant previous word and next word images queries are listed which is already stored in the database that could be searched by the corresponding users. In the select reference classes corresponding images could be displayed to the user.

For a keyword q, we automatically define its reference classes through finding a set of keyword expansions E(q) most relevant to q, to achieve this a set of images

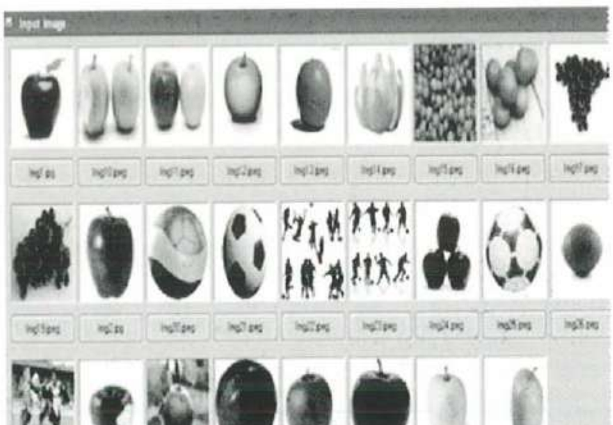

Fig. 2. Input image

S(q) are retrieved by the search engine using q as query based on textual information. keyword expansions are found from the words extracted from the images in S(q).

Meta Annotation Re-ranking

In meta annotation process, visual and textual features such as colour, shape and textures are considered and compared with user selected image. Here each image colour could be counted and that can be counted and stored in database. These countings can be compared with remaining other images. For this we can use instance ranking score. This can calculate positive and negative bag images. S(BI) performs this task. Depending on the ranking percentage the chart could be plotted which represent relevant and irrelevant images percentage. By using this user can find relevant image percentage score. The image keyword relationship compares the total image versus image keyword as shown in Fig. 3.

In Fig.3 the positive image ranking score can be displayed for every keyword. It shows the ranking percentage of comparison result with other images Here every image ranking percentage can be varied according to the relevant images in the positive bag. The higher percentage of ranking score has the more relevant images which moves to the positive bag whereas below 70 percentage moves to the negative bag which contain irrelevant images. This provides more efficient re-ranking of images compared to the existing method.

Image Re-Ranking

In this process input is selected by the user and the corresponding output can be displayed by re-ranking of images. Here the positive bag and negative bag percentage could be calculated by means of Generalized Multiple Instance(GMI) Learning process. If the output ratio exceeds 69 per cent then it moves to positive bag otherwise it moves to negative bag . This negative bag images are irrelevant that cannot be considered here. Fig.5, shows the re-ranking percentage for every image. Depending on the number of relevant images the percentage could be calculated for every image in the database.

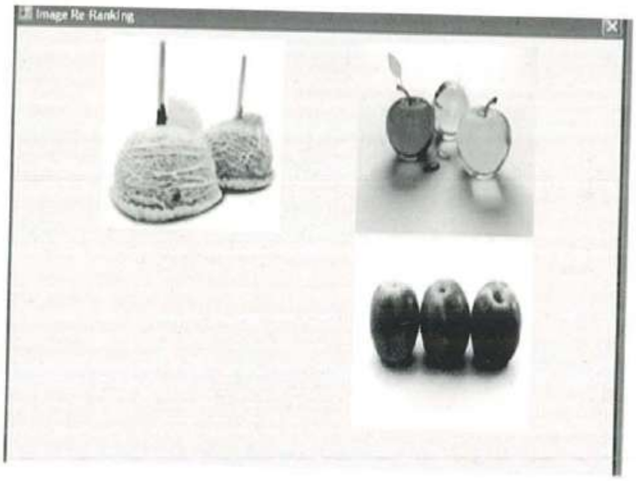

Fig. 3. Meta annotation ranking process

EXPERIMENT AND RESULTS

In this experiment, manual bag annotation process has been simulated by using the ground-truth labels of the images. For each concept, we use one positive bag and one negative bag as the training data. Since the negative bags do not contain any positive instances in most cases, only the positive bags need to be manually annotated. Specifically, if a positive bag contains at least proportion of truly positive instances, it is annotated as a truly positive bag. Based on the initial bag ranking results according to the bag ranking score in (2), we observe that only 72 concepts have at least one truly positive bag after checking with the ground-truth labels. Therefore, we report MAPs over the 72 concepts only in this experiment.

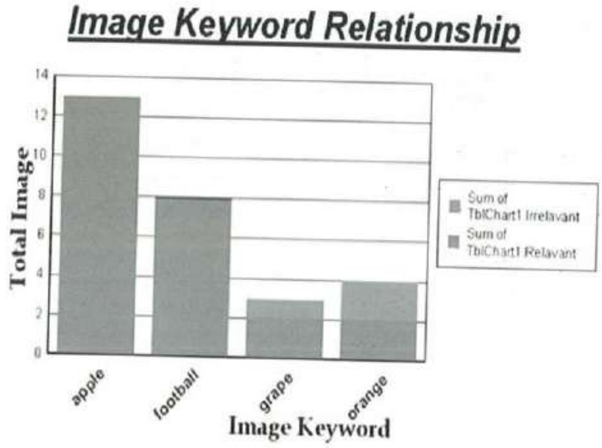

Fig. 4. Image keywords re-ranking relationship

It is worth mentioning that the users are generally reluctant to conduct manual annotations. Thus annotation based technique as an additional extension has not been maintained as the main focus of this paper. More details (e.g., how to develop a novel and effective annotation user interface to facilitate the bag annotation process in the real applications and as to how to fairly compare our approach with conventional RF methods) will be investigated in the future.

CONCLUSION

In this paper a meta annotation based frame work for large scale TBIR has been proposed. By given a textual query, relevant images are to be re-ranked after the initial text based search. To address the ambiguities on the instances of both positive and negative bags, Generalized Multiple Instance (GMI) has been developed to further enhance retrieval performance. Our framework using the automatic bag annotation method can achieve the best performance compared to

P - ISSN 0973 - 9157
E - ISSN 2393 - 9249

other traditional image re-ranking methods. It also has the user log in for individual users which provide time consuming process for effective image retrieval.

FUTURE WORK

Currently, we use the -means clustering method based on visual and textual features to partition the relevant images into bags/clusters in our weak bag annotation process. In the future, it would be worth while to investigate more effective clustering methods to further improve the performance of our framework.

REFERENCES

Cui, J., Wen, F. and Tang, X. 2008. Real time Google and live image search re-ranking. In Proc ACM Multimedia.

Florian Schroff., Antonio Criminisi., and Andrew Zisserman. 2011. Harvesting Image Databases *from the Web. Proc. IEEE Conf. Computer Vision and Pattern Recognition.*

Hsu, W. H., Kennedy, L. S. and Chang, S. F. 2006. Video search rerankingvia information bottleneck principle, *Proc. 14th ACM Int. Conf. Multimedia*, P. 35–44.

Hsu, W. H., Kennedy, L. S. and Chang, S. F. 2007. Video search rerankingthrough random walk over document-level context graph, *Proc.15th ACM Int. Conf. Multimedia*, P. 971–980.

Jen-hao Hsiao, Chu-song chen and Ming-syanchen, 2008. A novel language model based approach for image object mining and re-ranking. IEEE DOI 10.1109/ ICDM.

Jing, Y. and Baluja, S. 2008. Textual query of personal photos facilitated by large scale web data. In: *Proceeding of the 17th International Conference on World Wide Web*, P.307–316.

Jing, Y. and Baluja, S. 2008. Pagerank for product image search. *Proceeding of the 17th International Conference on World Wide Web*, P. 307–316.

Jing, Y. and Baluja, S. 2008. Pagerank for product image search, *Proc.17th Int. Conf. World Wide Web*, P. 307–316.

Kelley, J. E. 1960. The cutting plane method for solving convex programs, *SIAM J. Appl. Math.*, 8(4): 703–712.

Kim, S. J. and Boyd, S. 2008. A minimax theorem with applications to machine learning, signal processing, and finance, *SIAM J. Optim.*, 19(3):1344–1367.

Li, Y.F., Kwok, J. T., Tsang, I. W. and Zhou, Z. H. 2009. A convexmethod for locating regions of interest with multi-instance learning, *Proc. Eur. Conf. Mach. Learn. Principles Pract. Knowl. DiscoveryDatabases*, P.15–30.

Li, Y.F., Tsang, I. W., Kwok, J. T. and Zhou, Z.H. 2009. Tighter and convex maximum margin clustering, *Proc. 22nd Int. Conf. Artif. Intell.Stat.*, P. 344–351.

Liu, Y., Xu, D., Tsang, I. W. and Luo, J. 2011. Textual query of personalphotos facilitated by large-scale web data," *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(5):1022–1036.

Maron, O. and Lonzano-Pérez, T. 1998. A framework for multiple-instancelearning, in *Advances in Neural information Processing Systems.Cambridge, MA: MIT press*, P. 570-576.

Maron, O. and Ratan, A.L. 1998. Multiple- Instance learning for natural scene classification, *Proc, 15th Int. Conf. Mach. Learn*, P. 341-349.

MattiaBroilo. 2008. A Stochastic Approach to Image Retrieval Using Relevance Feedback and Particle Swarm OptimizationIn *Proceedings 10th Workshop on Multimedia Signal Processing-MMSP, Cairns, Australia.*

Rui, Y., Huang, T. S. and Mehrotra, S. 1997. Content-based image retrieval withrelevance feedback in mars. *Proceedings of the IEEE InternationalConference on Image Processing*, P.815–818.

Xiao gang Wang, Ken Liu and Xiao Tang. 2003. Query-Specific Visual Semantic Spaces for Web Image Re- In *Proceedings of the ACM International Conference on Image and Video Retrieval*, P. 238–247.

Yan, R., Hauptmann, A. G. and Jin, R. 2003. Multimedia search withpseudo-relevance feedback. *Proceeding of the ACM International Conference on Image and Video Retrieval*, P. 238-247.