A Novel Bag-of-Object Retrieval Model To Predict Image Relevance

Gouri K. Ghadge[1], G. J. Chhajed[2]

[1]ME-II Student, Vidya Pratishthan's College of Engineering, Baramati, Savitribai Phule Pune University, Pune, Maharashtra, India,

[2]Assistant Professor, Vidya Pratishthan's College Of Engineering, Baramati, Savitribai Phule Pune University, Pune, Maharashtra, India

Abstract: ***Text based image search is now a days a routine work for which generally, image search re-ranking and image result summarization these two effective approaches are used. But these approaches are not suitable for the object queries. In this system, a novel bag-of-object retrieval model is designed to predict image relevance, which is specifically effective for object queries. In this approach, first, an object vocabulary is constructed which contains query-relative objects based on expanded query sets which considers the frequent object patches in the result image collection. Then training model is constructed on the basis of frequent object patches. The efficiency of the proposed model is demonstrated by comparing the results with image search re-ranking and image search result summarization.***

Keywords: Re-ranking, summarization, object retrieval

1. Introduction

Image search re-ranking and image search result summarization are variant aspects to improve result quality and search experience using visual information. Image search re-ranking upgrade the relevant images to the top of the search result list while suppressing the irrelevant ones to the bottom, such that the user can access satisfying images in top positions. Image search result summarization finds a small group of images that are both relevant and representative to the query, providing the user with an overview of the search result list. For both of the approaches, predicting image relevance serves as a key technology and is considered as one of the most challenging tasks.

The cluster assumption and the pseudo relevance feedback (PRF) assumption these two assumptions have been formerly used to estimate the relevance of the images. The cluster assumption suggests that relevant images consistently have close visual appearance while irrelevant ones are regarded as noise thus distinct with each other. PRF assumption regards the images ranked to the top of the text-based search result as pseudo-relevant, which can be employed to train a classifier or multiple classifiers.

In this system, the main focus is on a typical category of queries, named object queries, where the user intends to find images containing the desired objects, including landmarks, products, vehicles, animals and people. While the two assumptions have been demonstrated generally effective in existing re-ranking approaches, but they are not sufficiently effective to deal with such kind of queries. The first problem is that, for images retrieved by object queries, usually some parts of the image are relevant to the object query, while the others are not.

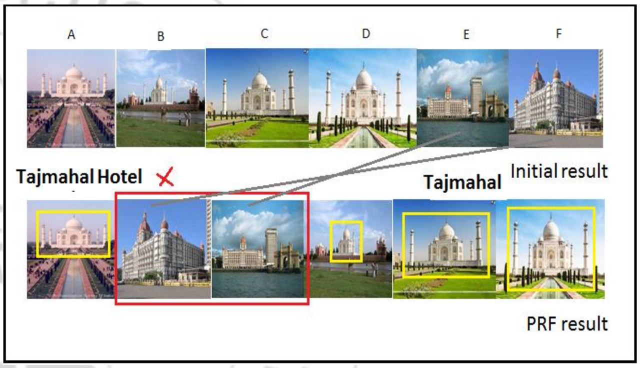

Figure 1: Example of PRF re-ranking process for query Tajmahal

For example, Fig.1 shows an simple PRF re-ranking process for the query Tajmahal where image A is assumed as pseudo-relevant while the other images are ranked with respect to their visual similarities to A. Unfortunately, two irrelevant images, image E and F are boosted to the top because image E and F are the images of Tajmahal hotel, although image A shows the Tajmahal. To overcome this problem a novel Bag-of-Object retrieval model is used in which it represents the query and result images into a language model using object appearance.

To demonstrate the effectiveness of the proposed retrieval model it is tested for re-ranking and summarization purpose.

2. Related Work

This system is used to re-rank images. In this section some previous work on re-ranking techniques are reviewed.

A. Multimedia Search with Pseudo-relevance Feedback [2]

In this method an algorithm is used which mainly focused on video retrieval that gives the decisions of multiple retrieval agents in both text and image methods. While the normalization and combination of evidence shows the successful use of negative pseudo relevance feedback to increase image retrieval performance. While it doesn't sort all problems in video information retrieval, the results are indicating that pseudo-relevance feedback gives great promise for multimedia retrieval with very varied.

B. Object Retrieval Using Visual Query Context [5]

The main goal of this method is retrieving images which contains objects. It is similar to the query object captured in the region of interest of the query image. By considering a large popularity of SIFT image features and bag of visual words image representation, object retrieval has increased significantly. Since existing object retrieval methods perform well in many cases, they may fail to return relevant results if the ROI specified by the user is inaccurate or if the object captured there is too small to be demonstrated using different features and consequently to be matched with similar objects in the image collection.

C. Learning from Search Engine and Human Supervision for Web Image Search [6]

In this method it combines two learning strategies for deriving the re-ranking model, first is learning from search engine and another is learning from human supervision. The first strategy learns the re-ranking model in a pseudo supervised way by interpreting parts of the initial text based search result as pseudo-relevant. The second strategy gives manual relevance labeling of the text-based search results generated for a limited number of representative queries.

D. Visual Categorization with Bags of Key points [3]

This is a novel method generic visual categorization. The problem of identifying the object content of natural images is generalizing across variations integrates to the object class. This bag of key points method is based on vector quantization of invariant descriptors of image patches. In this comparison of two alternative implementations is done using some classifiers, Nave Bayes and SVM. The main advantages of the method are that it is simple, efficient and intrinsically invariant. It gives results for simultaneously classifying seven semantic visual categories. These results show that the method is robust to background clutter and gives good categorization accuracy.

E. Supervised Re-ranking for Web Image Search [4]

In this method the learning-to-re-rank paradigm is used, which derives the re-ranking function in a supervised fashion from the human labeled training data. While supervised learning approach does not suffer from scalability issues while a unified re-ranking model is learned that can be applied to all queries. A query-independent re-ranking model will be trained for all queries using query- dependent re-ranking features. The query-dependent re-ranking feature extraction is challenging while the text query and the visual documents have different representation.

F. Bayesian video search re-ranking [7]

Content-based video search re-ranking can be considered as a process that uses visual content to recover the true ranking list which was generated based on textual information. The Bayesian framework solves the problem, of maximizing the ranking score consistency among visually similar video shots while minimizing the ranking distance, which gives the disagreement between the objective ranking list and the initial text based. It is Different from existing point wise ranking distance measures, which calculate the distance in terms of the individual scores, two new methods are used to measure the ranking distance based on the disagreement in terms of pair wise orders.

Table 1: Summary of literature survey

Sr No.	Method	Comments
1	Multimedia search with pseudo relevance feedback	It gives successful use of negative pseudo-relevance feedback to improve image retrieval performance
2	Object retrieval using visual query context	An object retrieval method that gives the information about the visual context of the query object
3	Learning from Search Engine and Human Supervision for Web Image Search	It combines two learning strategies for deriving the re-ranking model, learning from search engine and learning from human supervision.
4	Visual Categorization with Bags of Key points	This bag of key points method is based on vector quantization
5	Supervised re-ranking for web image search	Re-ranking features are proposed by representing the textual query using visual context and pseudo relevant images from the initial search result

3. Problem Statement

In the existing systems, in response to object query, the images retrieved have either the part of object or the object present in image is of very small size. Another problem is, since many times the response is result of labeled based search , retrieved images may be related to the object query where expected objects are missing. To overcome this problem a user is provided with prediction of relevance using a novel Bag-of-Object retrieval model .It represents the query and result images into a language model using object appearance. In order to focus on the valuable object categories and suppress the background noise, a common object discovery (COD) algorithm is used to pick up the query-relevance ROIs and construct a query-relevant object vocabulary.

4. System Architecture

Fig. 2 shows proposed system architecture which indicates relevant images related to user query. The training part shows actual implementation methods for the relevance prediction.

Paper ID: NOV164602 http://dx.doi.org/10.21275/v5i6.NOV164602

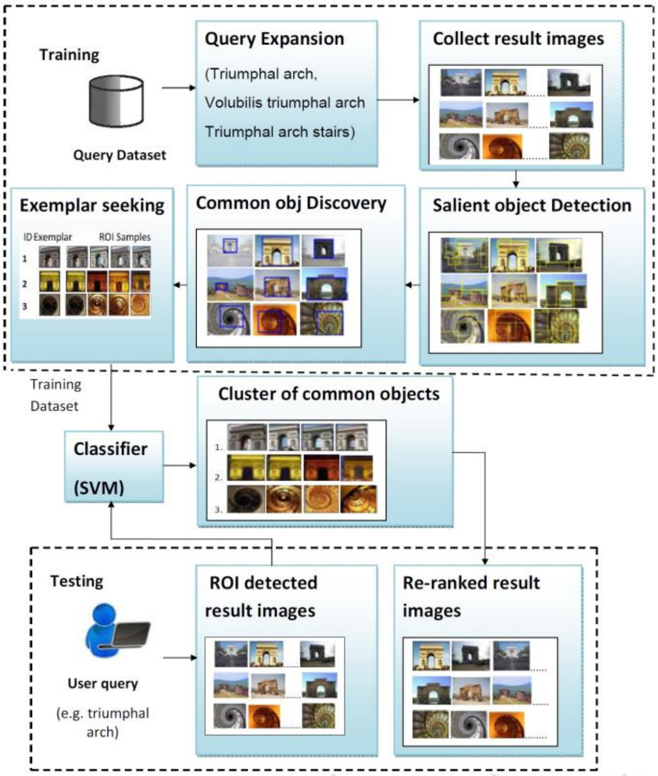

Figure 2: Proposed System Architecture

i. Query Expansion

Given a query q the goal of this step is to discover a set of key words which are semantically related to the query.

ii. Salient object detection

A group of salient ROIs are extracted from each image using the salient object detection method. At the very beginning, a group of salient ROIs are extracted from each image using the salient object detection approach in [8]. Each image is then represented as a bag of ROIs, Then the true object ROIs can be further located using common object discovery.

iii. Common object detection

To focus on the valuable object categories and suppress the background noise Common Object Discovery (COD) algorithm [9] is used. This is the fast and scalable alternating optimization technique to detect regions of interest (ROIs) in cluttered Web images without labels.

iv. Exemplar seeking

In this procedure the detected foreground ROIs are clustered and cluster centers are adopted as exemplars to represent the foreground objects.

v. Nearest Neighbor Hard Voting

According to the cluster assumption, if the object has more "sponsors" in the result image collection , it is more likely to be relevant. Here, we gather the sponsors of each object by assigning each image to its closest object category.

4.1 Algorithms

1. Query Expansion

Input: Query keywords ,Image context;
Output: Expanded Query keywords ;
Steps
1) Extract image context
2) Pre-process context
3) Calculate TF*IDF with respect to query terms
4) Sort context keywords by TF*IDF score

2. Salient object detection

Input: RGB Image;
Output: Object detected without background;
Steps:
1)Collect pixels from foreground model
2)Scan all pixels of image
3)Segment pixels with respect to foreground model , non-foreground model and un-categorized
4)Select segments with higher matching probability with foreground model
5)Replace segment with pixel values in new image

3. Common object detection

Input: Set of objects, Similarity threshold;
Output: Set of common objects;
Steps:
1) Extract SIFT features of each object
2) Calculate similarity using Euclidean distance
3) Match similarity threshold
4) Collect common objects with respect to similarity

5. Results and Discussion

1. Query is given to the search engine and query expansion is done.

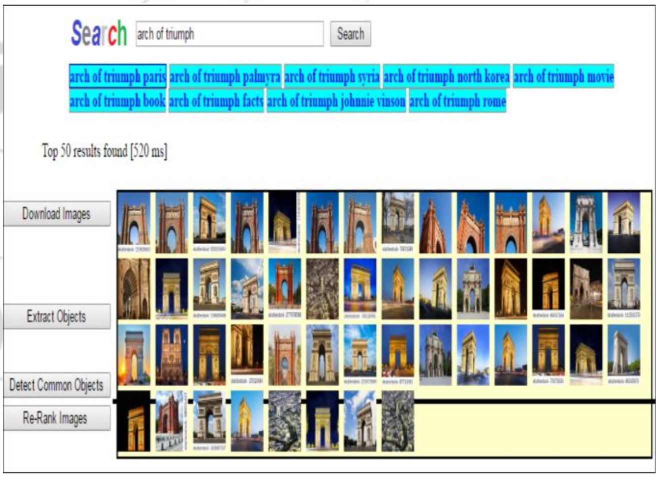

Figure 3: Query expansion and image results

2. Objects are Detected from Result Images

Figure 4: Objects are detected from images

International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

3 .Common objects are discovered and clusters are formed

Figure 5: Common objects are detected

4. On the basis of clusters ranking of images is done

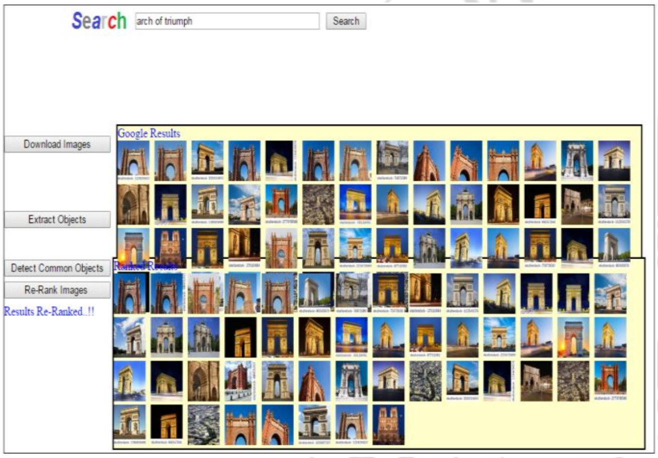

Figure 6: Images are ranked using nearest neighbor hard voting

Table 2: Result table

Query	No. of img results	No. of objects	No. of retrieved objects	No .of common objects	No. of common objects retrieved	No. of clusters
Triumphal arch	40	33	29	8	7	6
apple	37	25	22	5	4	4
Chelsea logo	39	15	12	4	3	3

6. Precision and Recall

Table 3: Precision and Recall

Query	Precision	Recall
Triumphal arch	0.75	0.87
Apple	0.80	0.88
Chelsea logo	0.75	0.80

Table 3.Shows precision and recall values for different queries. By considering these values graph is constructed.

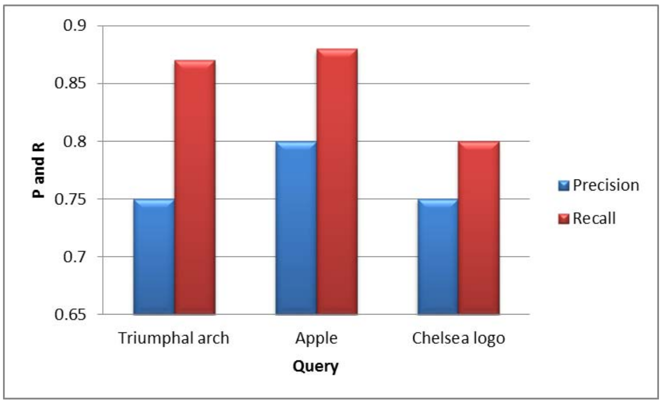

Figure 7: Precision and Recall

7. Conclusion

In image search re-ranking and image search result summarization relevance prediction is one of the biggest challenges. A novel bag-of-object retrieval model is used to give a more accurate prediction of image relevance. In this system, the query related objects are collected and then these are considered in common object discovery. This common object discovery approach is used to pick up the query-relevance ROIs and construct a query-relevant bag-of-object. Since all previous image search re-ranking method builds one generic model to all kinds of queries, this system serves as the first attempt to deal the queries from a specific domain. This system is tested for different queries and it gives different bags of objects on the basis of expanded query. It supports the user to predict image relevance.

References

[1] Yang Yang, Linjun Yang, Gangshan Wu, and Shipeng Li, "Image Relevance Prediction Using Query-Context Bag-of- Object Retrieval Model," IEEE transactions on Multimedia, Vol. 16, No. 6, October 2014.

[2] R. Yan, A. Hauptmann, and R. Jin, "Multimedia search with pseudo relevance feed- back," Image Video Retrieval, pp. 649654, 2003.

[3] G. Csurka, C. Dance, L. Fan, J.Willamowski, and C. Bray, "Visual categorization with bags of key points," in Proc. Workshop Statist. Learn. Comput. Vision ECCV, 2004, vol. 1, p. 22.

[4] L. Yang and A. Hanjalic, "Supervised reranking for web image search," in Proc. ACM Multimedia, 2010, pp. 183192

[5] L. Yang, B. Geng, Y. Cai, A. Hanjalic, and X. Hua, "Object retrieval using visual query context," IEEE Trans. Multimedia, vol. 13, no. 6, pp. 12951307, Dec. 2011.

[6] L. Yang and A. Hanjalic, "Learning from search engine and human supervision web image search," in Proc ACM Multimedia, 2011, pp. 13651368.

[7] X. Tian, L. Yang, J. Wang, Y. Yang, X. Wu, and X. Hua, "Bayesian video search reranking," in Proc. ACM Multimedia ,2008, pp. 131140.

[8] J. Feng, Y.Wei, L. Tao, C. Zhang, and J. Sun, "Salient object detection by composition," in Proc. 2011 IEEE Int. Conf. Computer Vision, 2011, pp. 10281035.

Author Profile

Ms. Gouri K. Ghadge received her B.E. Degree in Information Technology in 2014 from Vidya Pratishthan's College Of Engineering, Baramati and pursuing Master of Computer Engineering in Vidya Pratishthan's College Of Engineering, Baramati, Savitribai Phule Pune University, Pune.

Prof. Mrs. G. J. Chhajed obtained her B.E. Degree in Computer Science and Engineering in 1991-95 from S.G.G.S.I.E.T, Nanded and Postgraduate Degree in Computer Engineering from College of Engineering, Pune (COEP) 2005-2007 both with distinction. She is approved Undergraduate and Postgraduate teacher of Pune University and has 19 years of experience. She has total 32 Publications in the National, International level Journals and proceedings of Conferences. She has membership of IEEE, LMCSI, IET, LMISTE and International Association LMIACSIT.