Signal Processing: *Image Communication* 41 (2016) 101–114

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Object retrieval with image graph traversal-based re-ranking

Siyuan Qi *, Yupin Luo

Tsinghua University, Tsinghua National Laboratory for Information Science and Technology (TNList), Department of Automation, Beijing 100084, China

ARTICLE INFO

Article history:
Received 8 April 2015
Received in revised form
19 November 2015
Accepted 21 December 2015
Available online 29 December 2015

Keywords:
Image graph traversal
Image attribute
Object retrieval
Re-ranking

ABSTRACT

The topic of this paper is the retrieval of a particular object. A graph traversal-based re-ranking framework for the baseline bag-of-words (BOW) approach is proposed. For an image, we consider not only its similarity with the query image, but also the relationship between other dataset images. We integrate these information as image attributes via an extended image graph and propose a graph traversal algorithm to efficiently obtain their values. By comprehensively considering these attributes, we propose an attribute similarity measure for re-ranking, which brings much performance improvement. We further use our method for the multiple-query retrieval with a simple extension of the virtual query. The experimental results show that our method significantly improve the baseline approach and achieves competitive performance compared with the other state-of-the-art methods. Additionally, our re-ranking method requires only a little extra memory space and time costs.

1. Introduction

For a specific object image, the task of object retrieval is ranking images that contain the object at the top of the result list. This topic has been studied extensively over recent years, with the BOW being the best-known framework [1], and other effective feature encoding methods are also widely used, such as VLAD (Vector of Locally Aggregated Descriptors) [2–4] and FV (Fisher Vector) [5]. In this paper, we focus on improving the performance of the BOW framework. In a standard BOW baseline approach, features are extracted from images and classified into visual words. Then, images are represented as vectors of visual words. During retrieval, images are ranked in terms of vector similarities with the query and a verification step further ranks the top images by their spatial consistency [6]. Many recent studies have addressed object retrieval based on the baseline approach. These approaches emphasize different key issues, including feature detection and quantization [7–11], component weighting [12–14], similarity metrics [15,16], and spatial constraints [17–19].

Another category of effective methods is re-ranking, which uses feedback from the baseline method to further improve the retrieval results. Our method also belongs to this category, and we further consider the relationship between dataset images by integrating these relationships into similarity evaluation and ranking. For an object image, there are often other images in the dataset that contain a same object but in different viewpoints or environments. Image graph proposed in [20] is a good tool of using this knowledge. When a query is executed, the feedback from retrieval results of the baseline approach and the image graph allows us to further build an extended image graph. Based on this, we introduce the concept of image attributes, which reflects the relationship between a dataset image and the query from different perspectives, and we can quickly obtain the attribute values via our proposed graph traversal algorithm. With comprehensive consideration of the properties of the attributes, a combination of re-ranking algorithm and attribute similarity

* Corresponding author.
E-mail address: qsy10@mails.tsinghua.edu.cn (S. Qi).

http://dx.doi.org/10.1016/j.image.2015.12.004
0923-5965/

measure is proposed to re-rank the dataset images in the result list.

We further expand our method to deal with the multiple-query retrieval. A virtual query is set as the start vertex of the traversal algorithm. Hence, without further modification, we can solve multiple-query search under our single-query retrieval framework.

The main contributions of this paper are three-fold: firstly, we propose the image attributes with an efficient graph traversal algorithm. Therefore, we can obtain relations between the query and dataset images in many perspectives without costly computation of low-level feature in re-ranking. Secondly, we propose an attribute similarity evaluation and a re-ranking procedure by comprehensively considering the attributes, which can effectively explore relevant images of different environments and viewpoints. Thirdly, the whole framework of our method is also compatible for multiple-query retrieval with a simple extension, and it achieves competitive results in evaluated datasets as well.

The remainder of this paper is organized as follows. Some related work is discussed in Section 2. In Section 3, our approach is introduced in detail. Experimental results and analysis are presented in Section 4. Finally, Section 5 presents our conclusion and future work.

2. Related work

The re-ranking methods always use feedback from the baseline system to further improve the retrieval results. A famous re-ranking method is the query expansion (QE) [21], which aims to make up the lost features owing to environment change via some relevant images. This method back-projects features of the spatially verified images confirmed by the baseline method to issue a new query, which is then used for a second round of search using the baseline method. This method has been shown to bring significant improvements of retrieval performance [9,22–24]. However, as QE method relies heavily on the spatial consistent results confirmed by baseline approach, it is difficult to gain much improvement with very few feedback. Moreover, the expansion of the query increases the required time to compute feature correspondences and spatial consistency in re-ranking. In [25], the authors proposed a discriminative query expansion (DQE) which uses machine learning methods to solve the retrieval problem. The spatially consistent and the most dissimilar results are selected as positive and negative samples, respectively, to train a SVM classifier that is further used for ranking in terms of the distance away from the decision boundary. This approach also suffers from a lack of spatial consistent images. Shen et al. re-queries the top-ranked images in the initial result list obtained by the baseline method, and merge the retrieval results for the final ranking result [19]. This method only considers the images that are confirmed relevant by the baseline method, which is difficult to accurately rank images that have large variances in appearance. These methods all focus on the low-level features or only consider the spatially consistent images obtained by the baseline method, which limits further performance improvements.

Dataset-side feature augmentation is a related method of another category, which could be regarded as the QE for each dataset image. It also aims at mitigating the effects of feature loss caused by environment change. Turcot et al. build an image graph to represent the relevant relationships between dataset images using the baseline method [20]. For each image in the graph, the BOW representation of each image is further augmented with the visual words of its neighbors. This method increases retrieval recall, but also requires extra space for the expanded features. The image graph is further improved by filtering features that are estimated to be invisible in the expanded image by [25]. Though image graph is used in these methods, their concerns are still on expanding low-level features. The relationships between dataset images contained in image graph, which are more reliable in exploring relevant image, are relatively ignored. The image graph is also used for data mining in [26], but object retrieval is not discussed in their work. A reciprocal structure describing image relations, which is built based on vector similarity, is proposed in [27]. However, as no reliable spatial information is considered, the performance of their work is limited. We realize the importance and reliability of relationship contained in the image graph, and hence apply it to improve retrieval performance in our work.

Multiple-query search deals with the problems by using multiple images of an object for retrieval. Arandjelović et al. propose a framework for multiple-query retrieval that overcomes the difficulties of single-query search [28]. In their work, several simple rules are introduced to merge the results of each single query. [29] addressed this issue by combining the DQE with certain effective learning methods. However, the classifier training rounds are time-consuming for real-time retrieval.

3. Our approach

Our approach explores relevant images via reliable relations between dataset images and the query, and also the relations between datasets images themselves. The whole framework of our approach consists of the off-line part and the on-line part. In off-line dataset processing, we compute the image graph using the baseline approach. The extension of the image graph, the graph traversal for attributes, the attribute similarity measure and the re-raking procedure are processed for final results in the real-time retrieval.

3.1. Baseline approach

Because our re-ranking approach is based on the initial results of the baseline approach, and the image graph is also conducted via it, we first briefly introduce the baseline approach.

The baseline approach used in this paper completely follows the guidelines of [6]. For each image, features are extracted with the Hessian-Affine detector [30] and represented as SIFT descriptors [31]. Then the descriptors