Accurate Image Retrieval with Unsupervised 2-Stage k-NN Re-Ranking

Dawei Li, Lehigh University, Bethlehem, PA, USA

Mooi Choo Chuah, Lehigh University, Bethlehem, PA, USA

ABSTRACT

Many state-of-the-art image retrieval systems include a re-ranking step to refine the suggested initial ranking list so as to improve the retrieval accuracy. In this paper, we present a novel 2-stage k-NN re-ranking algorithm. In stage one, we generate an expanded list of candidate database images for re-ranking so that lower ranked ground truth images will be included and re-ranked. In stage two, we re-rank the list of candidate images using a confidence score which is calculated based on, rRBO, a new proposed ranking list similarity measure. In addition, we propose the rLoCATe image feature, which captures robust color and texture information on salient image patches, and shows superior performance in the image retrieval task. We evaluate the proposed re-ranking algorithm on various initial ranking lists created using both SIFT and rLoCATe on two popular benchmark datasets along with a large-scale one million distraction dataset. The results show that our proposed algorithm is not sensitive for different parameter configurations and it outperforms existing k-NN re-ranking methods.

KEYWORDS

k-NN, Multimedia Retrieval, Ranking Consistency, Reciprocal k-NN, Re-Ranking

INTRODUCTION

Image retrieval systems allow a user to search an image of interest by returning a list of similar images (e.g., including the same physical object) which are ranked by the similarity to the query. A typical image retrieval system involves two required components: image representation and image similarity calculation, and one optional but quite effective component: result re-ranking. Let us consider how an image retrieval system using the popular Bag-of-Visual-Word (BOVW) framework (Sivic, 2003) works. First, local image features such as SIFT (Scale-invariant feature transform) (Lowe, 2004) are extracted from each image and quantized using a pre-trained codebook into a histogram of visual words as the image representation, and the histogram is further weighted using the *tf-idf* (term frequency–inverse document frequency) scheme. Extensive studies for improving the robustness of the BOVW image

representation have made significant advancements, such as Hamming Embedding (Jegou, 2008) and Multiple Assignment (Jegou, 2009). Second, an inverted index is constructed over the database images for real-time image similarity calculation, so that only images with common visual words to those in a query image are compared and included in the returned ranked list. Finally, a re-ranking method is applied to the initial list using some metrics and additional image information, e.g., feature geometric verification (Philbin, 2007).

In this paper, we focus on the problem of the final re-ranking step assuming an initial ranking list is given. Traditionally used raw feature geometric verification method which relies on applying the RANSAC (Random sample consensus) algorithm over local features' coordinates (Fischler, 1981) is expensive in terms of both computation and storage overhead. For better efficiency and effectiveness, a new class of k-Nearest Neighbor (k-NN) re-ranking methods (Chen, 2014; Pedronette, 2014; Qin, 2011; Shen, 2012; Li, 2015a) were proposed and have been demonstrated to outperform the geometric verification method. The k-NN based re-ranking methods are motivated by the fact that an image retrieval system will return similar ranked lists for visually similar images, and thus they refine an initial ranked list by comparing the k-NN of a query image to those of the candidate database images (e.g., the initially highly ranked images). The only additional information required for performing k-NN re-ranking is the nearest neighbors of each candidate database image which can be pre-calculated during the database index construction stage for less real-time computational cost.

Inspired by several previous works (Chen, 2014; Pedronette, 2014; Qin, 2011; Shen, 2012), we propose a novel unsupervised 2-stage k-NN based re-ranking algorithm. In stage one of our method, we generate an expanded list of candidate database images for re-ranking. Such expanded image list is constructed by adding top neighbors' reciprocal nearest neighbors to the k-NN list of the query image. The benefit of this expanded list is that those low ranked ground truth images will be included for re-ranking, and it also avoids the issue of having to re-rank too many images to achieve a high recall that other methods suffer from. In stage two, we re-rank the candidate images in the expanded list by calculating a confidence score for each image. The confidence score incorporates both ranking consistency (Webber, 2010) (i.e., whether two ranked lists share common items and common item sequences) and reciprocal k-NN (Qin, 2011; Pedronette, 2014) (i.e., whether items in two ranked list are reciprocal k-NN) information.

We evaluate our proposed method over two popular benchmark datasets (Holiday and UKbench) as well as a large-scale dataset which is constructed by adding 1 million distraction images to the Holiday dataset. The initial ranking lists used in our experiment are generated using two types of local image features, SIFT (Lowe, 2004) and *rLoCATe* which is an enhanced version of a recently proposed local color and texture image feature (Iakovidou, 2014). Our results show improved re-ranking performance over the state-of-the-art methods by achieving mAP (mean average precision) of 92.64% for UKbench Dataset and 78.04% for Holiday Dataset. Our results also show that our re-ranking scheme is not too sensitive to the parameter settings associated with our scheme. Figure 1 presents three sample re-ranking results using our proposed method. For each query (left), initial ranked list (the first row) and re-ranked list (the second row) are demonstrated. They show that our method returned more ground truth images in the top-ranked list.

The rest of the paper is organized as follows. We first review the related work. Then, we describe our proposed 2-stage k-NN re-ranking method. In the following section, we introduce the new *rLoCATe* feature that will be used to generate the initial ranking list. Next, we present the experimental result. Finally, we conclude the paper.

RELATED WORK

Re-ranking is a critical step for improving retrieval accuracy. A traditional method for re-ranking is to verify the spatial consistency among detected local image features between the query image and top-K results (Philbin, 2007; Zhang, 2011; Chum, 2009; Perd'och 2009). However, the drawback of

Related Content

Interactive Digital Television
Margherita Pagani (2009). *Encyclopedia of Multimedia Technology and Networking, Second Edition (pp. 726-734).*
www.igi-global.com/chapter/interactive-digital-television/17472?camid=4v1a

Collaborative Work and Learning with Large Amount of Graphical Content in a 3D Virtual World Using Texture Generation Model Built on Stream Processors

Andrey Smorkalov, Mikhail Fominykh and Mikhail Morozov (2014). *International Journal of Multimedia Data Engineering and Management (pp. 18-40).*

www.igi-global.com/article/collaborative-work-and-learning-with-large-amount-of-graphical-content-in-a-3d-virtual-world-using-texture-generation-model-built-on-stream-processors/113305?camid=4v1a

Weighted Association Rule Mining for Video Semantic Detection

Lin Lin and Mei-Ling Shyu (2012). *Methods and Innovations for Multimedia Database Content Management (pp. 12-27).*

www.igi-global.com/chapter/weighted-association-rule-mining-video/66685?camid=4v1a

Exploring Multimedia Performance in Educational Research

Carl Bagley (2008). *Handbook of Research on Digital Information Technologies: Innovations, Methods, and Ethical Issues (pp. 14-26).*

www.igi-global.com/chapter/exploring-multimedia-performance-educational-research/19832?camid=4v1a