Neurocomputing 173 (2016) 172–180

Contents lists available at ScienceDirect

Neurocomputing

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

Visual search reranking with RElevant Local Discriminant Analysis

Peiguang Jing [a], Zhong Ji [a,b,c,*], Yunlong Yu [a], Zhongfei Zhang [c]

[a] School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China
[b] State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China
[c] Department of Computer Science, State University of New York, Binghamton, NY 13902, USA

ARTICLE INFO

Article history:
Received 5 June 2014
Received in revised form
26 October 2014
Accepted 25 December 2014
Available online 3 September 2015

Keywords:
Image retrieval
Visual search reranking
Dimensionality reduction
Linear discriminant analysis
Locality preserving projections

ABSTRACT

Visual search reranking is a promising technique to refine the text-based image search results with visual information. Dimensionality reduction is one of the key preprocessing steps in it to overcome the "curse of dimensionality" brought by the high-dimensional visual features. However, there are few dimensionality reduction algorithms employing the relevance degree information for visual search reranking. This paper proposes a novel dimensionality reduction algorithm called *RElevant Local Discriminant Analysis* (RELDA) for visual search reranking. As a semi-supervised combination of improved *Linear Discriminant Analysis* (LDA) and *Locality Preserving Projections* (LPP), the proposed RELDA algorithm preserves the local manifold structure of the whole data as well as controls the relevance between labeled examples. Moreover, RELDA algorithm has an analytic form of the globally optimal solution and can be computed based on eigen-decomposition. Extensive experiments on two popular real-world visual search reranking datasets demonstrate the superiority of the proposed RELDA algorithm.

1. Introduction

With the exponential growth of online image data, *Content Based Image Retrieval* (CBIR) has been an active and challenging research area over the past decades. However, due to the well-known "semantic gap", popular image search engines still employ the metadata associated with media contents as features and depend on text-based search techniques. These approaches cannot accurately and completely reflect the visual content of images, which leads to the phenomenon that noisy results always appear on the top of the search list. In the era of "Big Data", visual search reranking is an emerging technique to address this problem, which makes use of visual information to reorder the text-based search results to obtain a better search result.

Lots of efforts have been made in the research of visual search reranking techniques, mainly including classification-based [1,2], clustering-based [3,4], graph-based [5–7], and learning-to-rank-based [8,9] strategies. In these methods, visual features play a critical role. However, the high dimensionality of visual features usually ranges from hundreds to thousands, which degrades the performance of many machine learning algorithms. This is the so-called the "curse of dimensionality". Dimensionality reduction is an important preprocessing step in most of visual search reranking methods. Its goal is to project the data examples into a low-dimensional space while preserving the desired intrinsic information.

Many dimensionality reduction methods have been proposed [10–13]. Manifold learning is one of the most effective methods, which assumes that data examples possibly reside on a nonlinear sub-manifold [14,15]. Popular algorithms include *Laplacian Eigenmap* (LE) [16], *Locally Linear Embedding* (LLE) [17], *Locality Sensitive Discriminant Analysis* (LSDA) [18], *Locality Preserving Projections* (LPP) [19], *Local Fisher Discriminant Analysis*(LFDA) [20], *Maximal Linear Embedding* (MLE) [21], *SEmi-supervised Local Fisher discriminant analysis* (SELF) [22], and *Semi-supervised Discriminant Analysis* (SDA) [23]. Algorithms of LE, LLE and LPP aim at finding the geometric structure of the underlying manifold. However, they fail to discover the discriminant structure in the data because of the unsupervised characteristics. Semi-supervised manifold leaning methods are more promising since they make use of both labeled examples and unlabeled examples simultaneously, where labeled examples are used to find the discriminant structure and unlabeled examples are used to find the local geometric structure. For instance, as a kind of semi-supervised LDA [24], SDA employs both labeled data and unlabeled data to build a graph which incorporates the neighborhood information of the dataset. Based on the notion of graph Laplacian, a smoothness penalty on the graph is incorporated into the objective function of SDA. Recently, SELF was proposed to linearly combine LFDA and PCA [20,25]. Its idea is to smoothly bridge LFDA and PCA so as to employ and

* Corresponding author at: School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China.
E-mail address: jizhong1@gmail.com (Z. Ji).

http://dx.doi.org/10.1016/j.neucom.2014.12.118
0925-2312/© 2015 Published by Elsevier B.V.

control both the global structure of unlabeled samples and the discriminative information brought by the labeled samples.

Meanwhile, many dimensionality reduction algorithms have also emerged for CBIR and ranking tasks [26], e.g., LGD [27], Rank-CCA [28], *Maximum Margin Projection* (MMP) [29], Biased Discriminative Euclidean Embedding (BDEE) [30], *Sparse Transfer Learning* (STL) [31], and Ranking Graph Embedding [32]. Specifically, LGD [27] is a semi-supervised dimensionality reduction algorithm applied to visual search reranking. It localizes user's intention by transferring both the local geometry and the discriminative information from labeled images to unlabeled ones. MMP [29] is designed to discover the local manifold structure by maximizing the margin between positive and negative examples at each local neighborhood. Tian et al. proposed STL method [31], which effectively and efficiently encodes users labeling information in the interactive video search reranking application. These algorithms have successfully obtained experimental results. However, most of them neglect the great differences between the classification task and the ranking task.

We aim at proposing a novel dimensionality reduction method directly designed for the image ranking task. We observe that the relevance degree information essentially differs from the traditional class label information. In addition, we find that the local geometric structure information is one of the key factors to improve the performance of visual search reranking. Thus, ranking information should be employed to control the relevance between labeled examples. Meanwhile, the local geometric structure should be applied to preserve the local manifold structure of the whole data. Based on the above two considerations, we propose a novel semi-supervised dimensionality reduction algorithm named RElevant Local Discriminant Analysis (RELDA) to address the visual search reranking problem. Inspired by the idea of SELF, we improve LDA [24] and combine it with LPP [19] to build the final objective function. The optimal projection can be calculated by Eigen-decomposition. Extensive experiments on two popular datasets show that our proposed algorithm outperforms the traditional feature dimensionality reduction algorithms and the state-of-the-art visual search reranking algorithms.

The rest of paper is organized as follows. Section 2 gives a brief review of the visual search reranking, SELF, and LPP. The proposed RELDA algorithm is described in Section 3. In Section 4, the experimental results on MSRA-MM 1.0 and MSRA-MM 2.0 image datasets are presented and discussed. Section 5 concludes the paper.

2. Related work

Since the proposed RELDA algorithm is designed for visual search reranking and inspired by SELF and LPP, this section gives brief descriptions of them.

2.1. Visual search reranking

As a new paradigm in the CBIR domain, a visual search reranking technique is defined as reordering the initial text-based search results by incorporating their visual information. Four strategies should be highlighted in the visual search reranking domain: classification-based, clustering-based, graph-based, and learning-to-rank-based strategies. Classification-based strategy [1,2] formulates visual search reranking as a binary classification problem to identify whether each returned image is relevant or not. *Support vector machine* (SVM) is generally employed, where human labeling [1] and pseudo-relevance feedback [2] are two typical solutions to selecting the training data.

Clustering-based strategy [3,4] considers that the relevant scores of visually similar documents should be closer to each other. Thus, it first groups the images in the initial results into several clusters, and then orders the clusters together with the examples within them with some measure criteria to get the final reranking result. For example, information bottleneck and NCuts clustering algorithms were employed in [3] and [4] respectively to refine the initial performance.

Graph-based strategy [5–7,33] is proposed recently, in which a graph is constructed with the visual documents as the nodes and the edges between them being weighted by their visual similarity. For example, Hsu et al. [5] formulated the reranking process as a random walk over a context graph. Wang et al. [6] proposed a novel graph-based method by exploring multiple modalities. The approach simultaneously learns the relevance degrees, weights of the modalities, and the distance metric with its scaling for each modality. Based on the Bayesian framework, Tian et al. [7] maximized the ranking score consistency among visually similar data examples while the ranking distance is minimized, which represents the disagreement between the objective ranking list and the initial text-based one.

Recently, learning-to-rank-based strategy [8,9] has shown promising results and receives an increasing attention. Learning-to-rank [34,35] is a newly proposed machine learning technology used to build a ranking model, and is shown a promising performance in the information retrieval domain. Typically, learning-to-rank-based reranking techniques first extract the image visual features from the initial search results, and then builds a ranking function with the labeled training data, and finally reorders the images with the ranking function. One good example is the work proposed by Yang et al. [8], in which learning-to-rank algorithms, i.e., Ranking SVM [34] and ListNet [35], were used to build the ranking function and then the initial search results were reranked. Liu et al. [9] studied the differences of classification and ranking in the reranking task, and presented two novel pairwise reranking models by formulating reranking as an optimization problem. They first converted the individual documents into "document pairs" and then found the optimal document pairs and their relevant relations with the proposed pairwise reranking models, and finally adopted a round robin criterion to recover the final ranked list.

In visual search reranking, dimensionality reduction techniques are usually utilized in two ways. One way is the direct utilization of the conventional algorithms. For example, Yao et al. [33] directly employed *Principal Component Analysis* (PCA) to reduce the dimensionality. The other way is to design specific algorithms for visual search reranking. For example, Tian et al. [27] proposed a local–global discriminative dimensionality reduction algorithm, where sub-manifold is learned by transferring the local geometry and the discriminative information from the labeled images to the global image data. From the query keyword perspective, Wang et al. [1] proposed a novel query-specific dimensionality reduction algorithm by automatically offline learning different visual semantic spaces for different query keywords through keyword expansions. However, few studies focus on the great differences between the classification task and the ranking task.

2.2. SEmi-supervised Local Fisher discriminant analysis (SELF)

SELF [22] is a recently proposed semi-supervised method, which reformulates LDA, LFDA, and PCA in a unified pairwise manner and combines the objective functions of unsupervised PCA with supervised LFDA.

Suppose that all the data examples $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are distributed in a D-dimensional space. Without loss of generality, the transformation vector is denoted by $\mathbf{w}$. There is a single vector