Neurocomputing 141 (2014) 202–210

Contents lists available at ScienceDirect

Neurocomputing

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

A comparative study of dimensionality reduction methods for large-scale image retrieval

Li Zhuo*, Bo Cheng, Jing Zhang

Signal & Information Processing Laboratory, Beijing University of Technology, 100 Ping leyuan, Chaoyang District, Beijing 100124, China

ARTICLE INFO

Article history:
Received 5 December 2013
Received in revised form
5 March 2014
Accepted 21 March 2014
Communicated by M. Wang
Available online 21 April 2014

Keywords:
Large-scale image retrieval
Dimensionality reduction
OPTIMIZED SIFT
HSV histogram
Vocabulary tree

ABSTRACT

“Curse of Dimensionality” is one of the important problems that Content-Based Image Retrieval (CBIR) confronts. Dimensionality reduction is an effective method to overcome it. In this paper, six commonly-used dimensionality reduction methods are compared and analyzed to examine their respective performance in image retrieval. The six methods include Principal Component Analysis (PCA), Fisher Linear Discriminant Analysis (FLDA), Local Fisher Discriminant Analysis (LFDA), Isometric Mapping (ISOMAP), Locally Linear Embedding (LLE), and Locality Preserving Projections (LPP). For comparison, Scale Invariant Feature Transform (SIFT) and color histogram in Hue, Saturation, Value (HSV) color space are firstly extracted as image features, meanwhile SIFT feature extraction procedure is optimized to reduce the number of SIFT features. Then, PCA, FLDA, LFDA, ISOMAP, LLE, and LPP are respectively applied to reduce the dimensions of feature vectors, which can be used to generate vocabulary trees. Finally, we can process large-scale image retrieval based on the inverted index built by vocabulary trees. In the experiments, the performance of various dimensionality reduction methods are analyzed comprehensively by comparing the retrieval performance, advantages and disadvantages, computational complexity and time-consuming of image retrieval. Through a series of experiments, we can conclude that dimensionality reduction method of LLE and LPP can effectively reduce computational complexity of image retrieval, while maintaining high retrieval performance.

1. Introduction

With the rapid development of the Internet, the popularity of various digital equipment, increasing mass storage devices, and the number of digital images have growth exponentially. It is important on how to effectively manage and rationally utilize such large-scale image information, which has become a challenging and urgent problem in the field of multimedia information retrieval [1–3].

Since the 1990 s, Content-Based Image Retrieval (CBIR) has appeared and flourished gradually, and become a mainstream image retrieval method nowadays [4–6]. The image features are used to represent the content of images. The similarity between images is measured and the images that have similar features with the query image will be regarded as the search results. With specific image feature extraction algorithms, extracted image features are invariant and objective, and require no manual operation, which enable CBIR as a retrieval method to reflect image content automatically and objectively.

The key technologies of CBIR include feature extraction, similarity measurement, and relevance feedback, etc. The most critical parts are feature extraction and similarity measurement. For CBIR, image features are used to represent the image content, therefore the extracted image features will directly have effects on performance of image retrieval. The features extracted from images can be defined into two levels: low-level features which refer to one aspect of physical characteristics; while high-level features are semantic characteristics which can reflect semantic level and understanding of image content. Therefore, how to extract image features rapidly and efficiently is a key problem of CBIR.

The existing CBIR system usually adopts visual features to represent image content, such as color, texture, shape, contour and spatial relationships, etc. High dimension features are usually extracted to describe image content accurately. Especially for large-scale image retrieval system, the number of features is large-scale as well. These high-dimensional image features may lead to “Curse of Dimensionality” which will cause a degraded retrieval performance. The primary reason of “Curse of Dimensionality” is that the distribution of original data is extremely sparse. Large deviation will appear in similarity measure of image features, which will result in inefficient performance of processing algorithm of features. Correspondingly, dimensionality reduction is an effective method to overcome this problem. Through this

* Corresponding author. Tel.: +86 10 6739 2859.
E-mail address: zhuoli@bjut.edu.cn (L. Zhuo).

http://dx.doi.org/10.1016/j.neucom.2014.03.014
0925-2312/

method, image features are preprocessed by projecting original data from a high-dimensional space to a lower dimensional space. In brief, dimensionality reduction method is an important step for CBIR.

The existing dimensionality reduction methods can be roughly divided into two categories: the unsupervised and the supervised. Unsupervised dimensionality reduction can minimize the loss of data information. The commonly used unsupervised dimensionality reduction methods include Principal Component Analysis (PCA) [7], Multidimensional Scaling (MDS) [8] and kernel PCA (KPCA) [9]. PCA aims to find the optimal projection matrix in terms of least mean square. MDS makes the Euclidean distance between data consistent with the distance of original data after dimension reduction. KPCA is an improved variety of PCA. By giving appropriate kernel function, PCA reconstruction in kernel space becomes relatively simple to handle nonlinear data. On the other hand, supervised methods can maximize inter-class discriminant information. The two commonly used supervised methods include Fisher Linear Discriminant Analysis (FLDA) [10], and Local Fisher Discriminant Analysis (LFDA) [11].

In addition to the dimensionality reduction methods mentioned above, manifold learning [12,13] is a novel method occurred in recent years, which learns and discovers smooth low-dimensional manifold embedded in high-dimensional space of finite discrete samples. It reveals low-dimensional structure in high-dimensional data, and then reconstructs and performs non-linear dimensionality reduction. The essence of manifold learning is a kind of nonlinear dimensionality reduction that can find intrinsic geometric constructions or rules while sample space is smooth high-dimensional, and obtain corresponding low-dimensional embedding. This method is better than traditional methods of dimensionality reduction in terms of reflecting the essence of data, and is more conducive in the understanding of data and further processing. The representative methods of the manifold learning are Isometric Mapping (ISOMAP) [14,15], Locally Linear Embedding (LLE) [16], and Locality Preserving Projections (LPP) [17] as well.

In this paper, six dimensionality reduction methods are compared and analyzed comprehensively to demonstrate their own performances in image retrieval. For effective comparison, Scale Invariant Feature Transform (SIFT) features [18] and color histogram in Hue, Saturation, Value (HSV) [19] color space are firstly extracted as image features, and SIFT feature extraction procedure is optimized to reduce the number of SIFT features. Then, PCA, FLDA, LFDA, ISOMAP, LLE, and LPP are respectively applied to reduce the dimensions of the image features, which can be used to generate vocabulary trees. Finally, through building an inverted index of vocabulary trees, large-scale image retrieval is implemented. In the experiments, the performance of various dimensionality reduction methods are analyzed by comparing the retrieval performance, advantages and disadvantages [20], computational complexity and time-consuming of image retrieval. It can be concluded that LLE and LPP methods can effectively reduce the computational complexity of image retrieval, while maintaining the high retrieval performance.

The remainder of this paper is organized as follows: Section 2 introduces the large-scale image retrieval framework based on vocabulary tree adopted in this paper, and a series of comparisons between the six dimensionality reduction methods. Based on this framework, the performance of various dimensionality reduction methods on image retrieval is compared. Section 3 presents the experimental results. Conclusions are drawn in Section 4.

2. Large-scale image retrieval framework based on vocabulary tree

In this paper, a large-scale image retrieval framework based on vocabulary tree [21] is firstly set up for a performance comparison of the six dimensionality reduction methods. The retrieval process is shown in (Fig. 1. First, image features are extracted from image database. Then, the dimension of these features is reduced and then used to construct two vocabulary trees by means of hierarchical K-means clustering scheme. By counting inverted index [22] based on vocabulary tree, the features and their indexes are stored in the image database.

When operating a query, the features of query image are extracted and indexed using the vocabulary tree. The similarity of images is matched and ranked by comparing the index of the query image with the indexes that are stored in the database. Finally, the top-k images are returned to the users.

As Fig. 1 shows, the image retrieval framework includes four key segments: feature extraction, dimensionality reduction, the construction of image index based on visual vocabulary tree and similarity matching, which will be further introduced below.

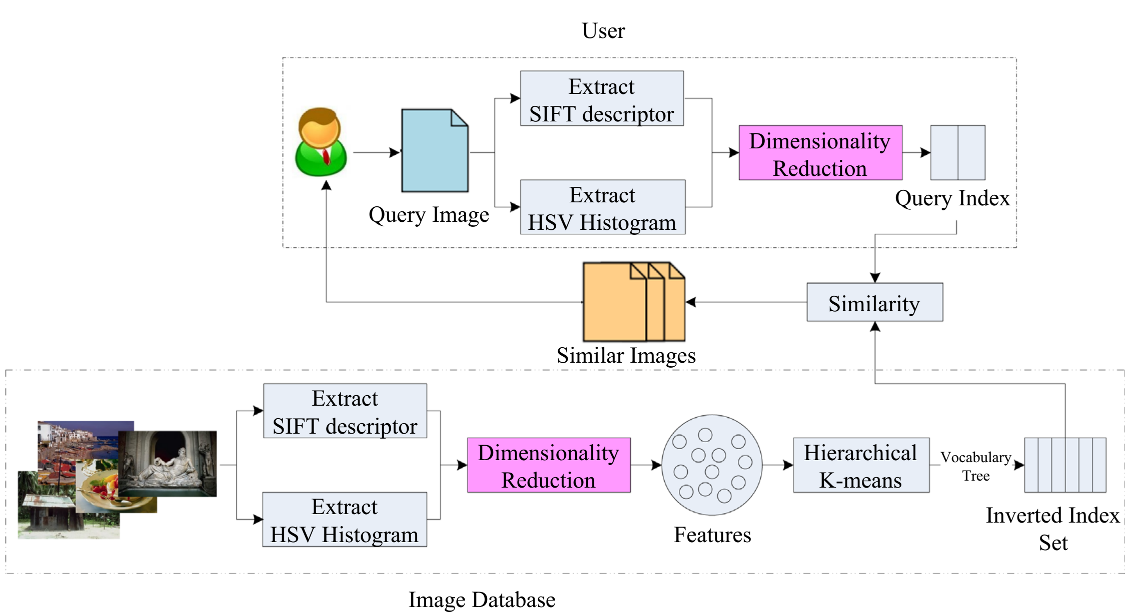

Fig. 1. The framework of large-scale image retrieval based on vocabulary tree.