Pattern Recognition Letters 102 (2018) 23–29

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Multi-bit quantization based on neighboring structure preservation

Liang-Liang Su[a,b], Jun Tang[a,b], Pu Yan[a,b], Dong Liang[a,b,*], Wen-Xia Bao[a,b]

[a] Key Laboratory of Intelligent Computing and Signal Processing, Ministry of Education, Anhui University, Hefei 230601, China
[b] School of Electronics and Information Engineering, Anhui University, Hefei 230601, China

ARTICLE INFO

Article history:
Received 2 December 2016
Available online 2 December 2017

Keywords:
Approximate nearest neighbor search
Neighboring structure
Hashing
Quantization coding

ABSTRACT

Hashing based approximate nearest neighbor search has become a research hotspot in computer vision. Most existing hashing methods concentrate on projection learning, and few efforts are dedicated to quantization coding. In this paper, we present a multi-bit quantization strategy to improve the quantization quality of projection values by adaptively learning quantization thresholds and quantizing each projection dimension with multiple bits. Our method exploits both the similarity and the local structure of samples in the original feature space and the pair-wise samples coding consistency. Extensive experiments on two canonical image datasets have shown that our method consistently outperforms the state-of-the-art quantization methods in terms of query performance.

1. Introduction

As an alternative to the nearest neighbor search technique, approximate nearest neighbor (ANN) search has received increasing attention from the researchers and shown to be enough and useful for many practical problems. Among the methods for ANN, hashing and vector quantization (VQ) are two popular solutions [1]. In general, the VQ method partitions the input space into multiple regions and maps each sample to the nearest regional center, then all the samples lying in the same region are represented by the same center. Therefore, the VQ method is a discrete but not binary encoding method, where each center is called a codeword, and the set of codewords forms a codebook. In order to enhance the representation of the codebook, product quantization [2] and its variants [1,3–6] have been developed to increase the number and improve the quality of the codewords. Furthermore, [4] and [6] represent each sample with hash codes by binarizing the real valued codewords. As a result, the VQ method requires extra space to store large codebook and lookup table of the distance between the codewords, and its query stage is time consuming if asymmetric distance computation is adopted to compute the distance between samples. By contrast, the hashing method can achieve fast query and less storage requirements.

The goal of hashing is to learn compact binary codes for high dimensional samples, which can preserve the neighboring structure between the samples in the input space. The significant gains can be obtained on both speed and storage from the compact data representation. That is, storage requirements are reduced sharply and ANN search can be performed in constant or sub-linear time. Generally, the hashing method contains two distinct stages: projection and quantization coding. Projection is the process of finding the optimal projection dimension, and the projection values should be close together if their corresponding samples are similar. Typically, the projection is generated by random selection or learning. The random projection involves Locality Sensitive Hashing (LSH) [7] and LSH-related methods [8], which are data-independent and generate hash functions based on random projections. Though these methods theoretically guarantee that original metrics are asymptotically preserved in Hamming space, the LSH family usually requires long codes and multiple hashing tables to achieve good performance. In contrast to the random projection, the data-dependent learning-based projections attract more attention as they can yield the discriminative compact binary codes. Representative projection learning methods include Principal Component Analysis Hashing (PCAH) [9], Iterative Quantization (ITQ) [10], Spectral Hashing (SH) [11], kernel-based Supervised Hashing [12], Fast Supervised Hashing [13] and Latent Structure Preserving Hashing[14]. Extensive experimental studies in the literatures have verified that data-dependent hashing methods significantly outperform data-independent hashing ones.

Compared with projection, relatively fewer efforts are devoted to quantization coding. Generally, most existing hashing methods adopt Single Bit Quantization (SBQ) strategy to quantize each projection dimension with one bit. In recent years, multi-bit quantization (MBQ) strategy which assigns multiple bits to each projection dimension has been developed to overcome the weakness of SBQ, such as Double Bits Quantization (DBQ) [15], Hierarchical Quantization (HQ) [16] and Manhattan Hashing (MH) [17]. More-

* Corresponding author at: School of Electronics and Information Engineering, Anhui University, Hefei 230601, China.
E-mail address: dliang@ahu.edu.cn (D. Liang).

https://doi.org/10.1016/j.patrec.2017.11.018
0167-8655/

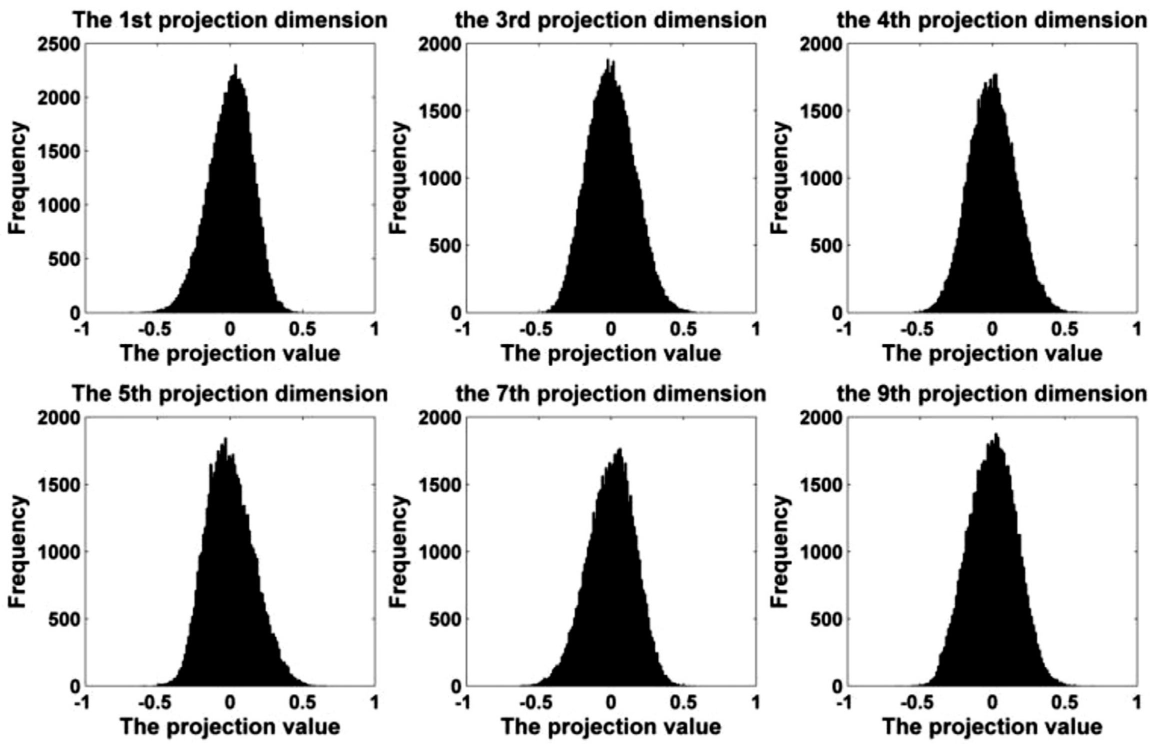

Fig. 1. The distribution of the projection values over six projection dimensions on CIFAR-10 dataset by ITQ.

over, more sophisticated approaches, including Neighborhood Persevering Quantization (NPQ) [18], Hamming Compatible Quantization (HCQ) [19] and Adaptive Quantization (AQ) [20], can allocate fixed or variable number of bits to each projected dimension. These MBQ methods significantly improve the search performance over SBQ and verify the importance of quantization coding in hashing methods. However, the above quantization schemes partially consider the neighboring structure of the data and overlook the local structure of samples. To address these issues, we propose an MBQ approach by comprehensively considering the similarity and local structure of samples and the pair-wise samples coding consistency.

The rest of this paper is organized as follows. Section 2 gives a brief review of quantization works. Section 3 describes our MBQ strategy. Experiments and results are reported in Section 4. Finally, we conclude the paper in Section 5.

2. Related work

As mentioned above, the SBQ approach often selects the average over the sample projection values as the threshold to quantize each projection dimension. Typically, the threshold is set to 0 when the data are centralized. Fig. 1 shows that the threshold 0 lies in the densest region for each projection dimension that are randomly chosen from the top 10 projection dimensions produced by ITQ [10] (In addition, similar results for other projection methods). As a result, many neighboring samples whose projection values are close to the threshold may be allocated to different bits. On the contrary, the samples far away from each other but in the same side of the threshold may be assigned the same bits. Therefore, this strategy is unable to well preserve the neighboring structure of original data, thus degrading the quantization effect.

In order to reduce the error of SBQ, researchers have proposed several promising MBQ methods to quantize each projection dimension with multiple bits, such as DBQ [15], HQ [16], MH [17], NPQ [18] and AQ [20]. Specifically, the first three are unsupervised methods and explore quantization thresholds by using the distribution of the projection values. However, these methods only consider the projection dimensions in the quantization stage and ignore the neighboring structure of original data. Moran et al. proposed the NPQ [18] method to obtain better thresholds by combining a supervised pairwise affinity matrix with a regularization term. To maintain the capability of similarity metric between the input space and Hamming space, Wang et al. [19] proposed the HCQ method to solve the desired thresholds. Yet the HCQ method has a high time complexity during threshold learning, which is difficult to be extended to quantize single projection dimension with more than 2 bits. In addition, varying bit assignment methods [20–22] can assign variable number of bits across projection dimensions under the assumption that the information contained in each projection dimension is not equal. Since there are a lot of hashing methods in hands, Liu et al. [23] presented a general hash bit selection framework to select the most informative binary code from a pool of candidate bits generated by different hashing methods. As a special case, Wang et al. [24] proposed an affinity preserved quantization method to transform samples into binary codes by incorporating the neighboring structure in the pre- and post-projection data space into vector quantization.

In this paper, we focus on improving the quality of quantization in hashing method. By combining the affinity between samples, the local structure and the distribution of samples, we present a novel quantization strategy to optimize threshold learning. Experiments on two benchmark datasets demonstrate that the proposed method outperforms other state-of-the-art quantization methods.

3. Quantization threshold learning

In this section, we describe the details of preserving the neighboring structure between samples in Hamming space from two aspects, i.e., neighboring structure preservation and pair-wise samples coding consistency.