VISUAL QUERY COMPRESSION WITH LOCALITY PRESERVING PROJECTION ON GRASSMANN MANIFOLD

ZHAOBIN ZHANG, LI LI, ZHU LI, AND HOUQIANG LI

UNIVERSITY OF MISSOURI, KANSAS CITY

UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA

ABSTRACT

For a variety of visual search and visual key points based navigation applications, compression of visual key point features like SIFT is an important part of the overall system that can directly affect the efficiency and latency. In this work, we examine a new approach in visual key points compression, that utilizes subspaces that optimized for preserving key point feature matching properties than the reconstruction performance, and allows for a set of optimal subspaces on Grassmann manifold that can better adapt to the local manifold geometry. The simulation demonstrates that such scheme has very low overhead in signaling subspaces, and has very much improved performance on the repeatability of the keypoint matching subject to bit rate constraints.

INTRODUCTION

Increasing application of mobile devices promotes the emerging industry of query-by-capture applications on mobile devices. Visual descriptors are usually extracted and compressed on the mobile device while matching is performed on the server using the transmitted feature data as the query. One of the most challenging work is how to minimize transmission data in order to reduce network latency. Therefore, visual compression is a key process for robust identification of mobile visual content, especially in a very large repository. existing methods either lays emphasis on statistic information or reducing dimensionality using a single transform which is not optimal enough to preserve their identification information. They failed to take local subspace relations into consideration such as PCA. In addition, PCA suffers from a number of shortcomings, such as its implicit assumption of Gaussian and is unable to reveal nonlinear relationships.

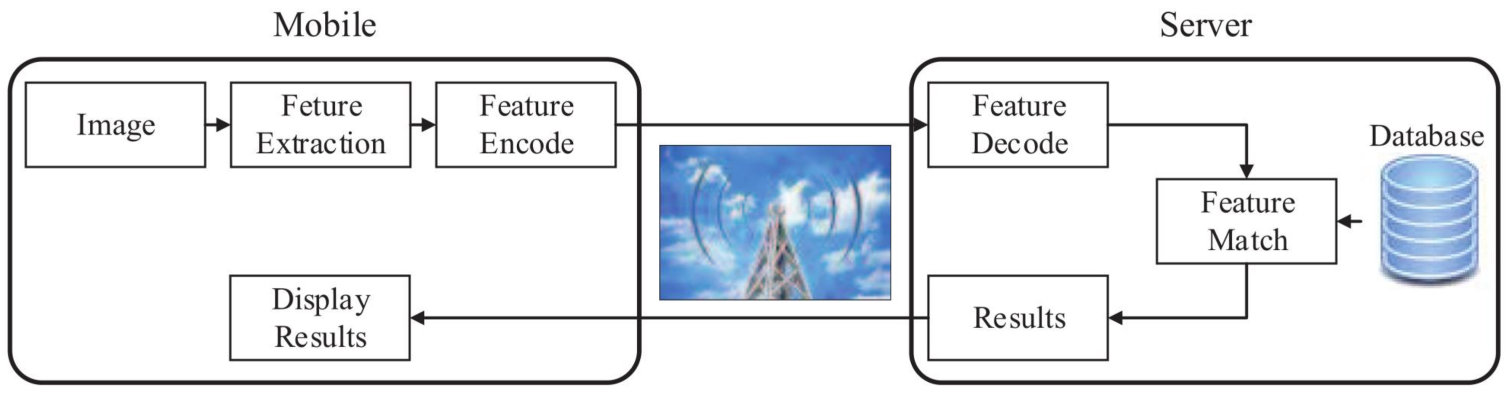

Fig. 1. Mobile server model

PROPOSED ALGORITHMS

1. Affinity matrix construction

Given a set of query descriptors $X = \{x_1, x_2, \dots, x_N\}$ where $x_i \in \mathbb{R}^n$, the nearest neighbor relationships are represented by their affinity matrix W which is computed as

$$W_{ij} = \begin{cases} 0, & if\ \|x_i - x_j\| > \theta \\ e^{-\|x_i - x_j\|^2/\sigma}, & otherwise \end{cases}$$

where σ and θ are the kernel parameters and the cut off threshold respectively and they are controlling the shape of Gaussian distribution and the sparsity of affinity matrix respectively. We optimize θ by searching for the value which preserves the most matching pairs whilst σ is empirically set to 0.25. The optimized θ values are listed as follows,

Table 1. Optimized θ value of each node.

Node	1	2	3	4	5	6	7	8
θ	0.17	0.49	0.40	0.10	0.26	0.31	0.39	0.19
Node	9	10	11	12	13	14	15	16
θ	0.23	0.49	0.22	0.66	0.15	0.20	0.51	0.56

PROPOSED ALGORITHMS

2. Grassmann subspace indexing

A single LPP transform is not good enough in capturing all manifold geometry characteristics in feature space. We focus on exploring multiple transforms such that each image descriptor is optimally projected to lower dimensional space. A binary tree is constructed and Grassmann manifold is incorporated for subspace indexing.

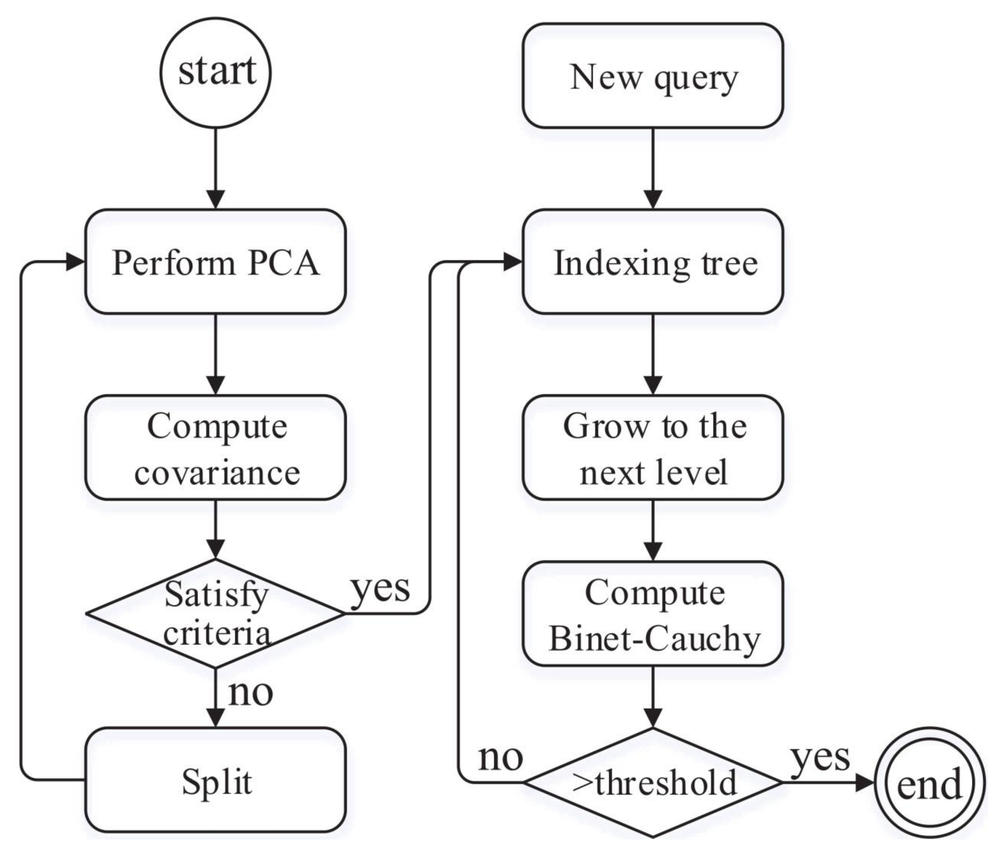

Fig. 3. Subspace indexing model.

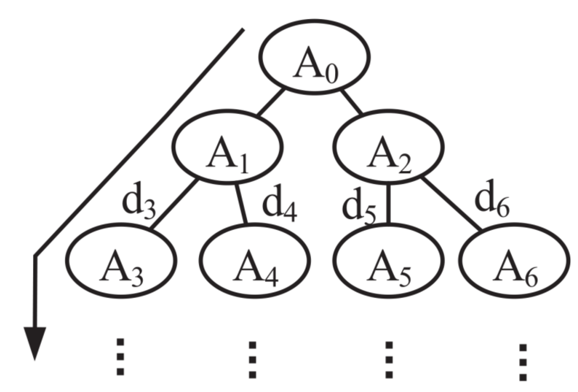

Fig. 4. Binary tree growing scheme.

EXPERIMENTAL RESULTS

Experiments are conducted over SIFT descriptors from CDVS dataset. Repeatability over bitrate is evaluated for proposed method (K-LPP) with comparison to PCA.

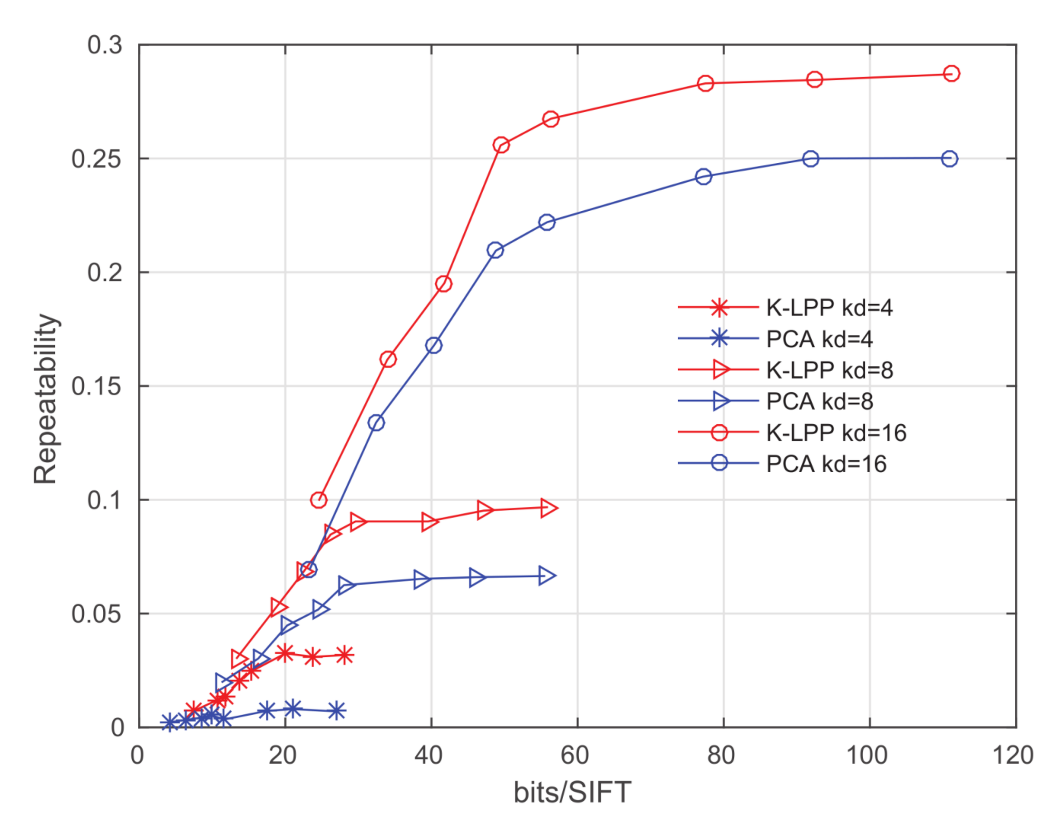

Fig. 5. Average repeatability of K-LPP and PCA.

CONCLUSION

This work proposes using multiple projective transforms instead of single transform. LPP is incorporated to preserve more local identification information and try to explore nonlinear matching relationships. We optimize training performance by dividing global training samples into small patches with Grassmann distance constraints. The proposed method is evaluated in MPEG CDVS dataset. The experimental results show that the proposed method which using multiple transforms with optimization on Grassmann manifold can outperform traditional PCA in repeatability under the same bitrate.