Equiangular Kernel Dictionary Learning with Applications to Dynamic Texture Analysis

Yuhui Quan[1,2], Chenglong Bao[2], and Hui Ji[2]
[1]School of Computer Science & Engineering, South China Univ. of Tech., Guangzhou 510006, China
[2]Department of Mathematics, National University of Singapore, Singapore 117542
{csyhquan@scut.edu.cn, matbc@nus.edu.sg, matjh@nus.edu.sg}

Abstract

Most existing dictionary learning algorithms consider a linear sparse model, which often cannot effectively characterize the nonlinear properties present in many types of visual data, e.g. *dynamic texture (DT). Such nonlinear properties can be exploited by the so-called kernel sparse coding. This paper proposed an equiangular kernel dictionary learning method with optimal mutual coherence to exploit the nonlinear sparsity of high-dimensional visual data. Two main issues are addressed in the proposed method: (1) coding stability for redundant dictionary of infinite-dimensional space; and (2) computational efficiency for computing kernel matrix of training samples of high-dimensional data. The proposed kernel sparse coding method is applied to dynamic texture analysis with both local DT pattern extraction and global DT pattern characterization. The experimental results showed its performance gain over existing methods.*

1. Introduction

In recent years, sparse dictionary learning has become one important tool in computer vision. Most existing methods for sparse dictionary learning consider a sparse linear model, which assumes that most local or global patterns of data, can be represented by the linear combinations of a small number of atoms from a dictionary. In other words, the underlying assumption of these linear model based dictionary learning methods is that the data for processing is dominated by the stationary patterns generated by some linear process. Clearly, such approaches will be less efficient when processing the data whose main structures are driven by nonlinear stochastic systems.

Indeed, there are many types of visual data, especially those high-dimensional ones, showing strong nonlinear behaviors in terms of visual features. One such representative data is dynamic texture (DT). DTs are referred to as the sequences of moving textures with certain stationary temporal changes in pixel intensities. The spatio-temporal behaviors of DTs are nonlinear in general. For instance, various distinguishable shapes may be observed from flickering fires with changes in wind, which implies multiple modalities of spatio-temporal appearance. Likewise, turbulent water exhibits chaotic behaviors with non-smooth motion, where pixel intensities do not change smoothly. Moreover, camera motion is likely to further aggregate the nonlinearities of correlations among the frames of a DT sequence.

In order to exploit the nonlinear properties existing in high-dimensional visual data, the so-called *kernel sparse coding* has been proposed in the literature which considers a nonlinear sparse model; see *e.g.* [12, 12, 44, 37, 11, 18, 25].The basic idea of kernel sparse coding is linearizing the nonlinear patterns existing in data in some implicit space and then studying the resultant linear structures by sparse coding under an implicit kernel dictionary.

1.1. Motivations

By considering a linear sparse model in implicit infinite-dimensional feature space, this paper aims at developing a nonlinear sparse model for high-dimensional visual data and investigating its applications in DT analysis and recognition. The motivations of applying kernel sparse coding to processing high-dimensional visual data are three-fold. Firstly, sparse coding is able to automatically discover the multiple modalities of patterns and distinguish the mixture of linear subspaces [1]. Secondly, sparse coding with dictionary learning adapts to the data and thus can better encode the stationary behaviors of data than the handcrafted features [39]. Thirdly, the nonlinearity of local structures could be partially linearized in a proper feature space induced by certain kernels, which improves both the accuracy and discriminability of sparse coding; see *e.g.* Fig. 2.

A direct call of existing kernel sparse coding approaches is not suitable for the tasks of processing high-dimensional visual data, such as DT analysis and recognition. Most existing kernel sparse coding methods do not consider the is-

1063-6919/16 $31.00 © 2016 IEEE
DOI 10.1109/CVPR.2016.40