J. Vis. Commun. Image R. 25 (2014) 1187–1198

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Image recognition via two-dimensional random projection and nearest constrained subspace

Liang Liao [a,b], Yanning Zhang [a,*], Stephen John Maybank [c], Zhoufeng Liu [b], Xin Liu [d]

[a] School of Computer Science, Northwestern Polytechnic University, Xi'an, Shaanxi 710129, PR China
[b] School of Electronics and Information, Zhongyuan University of Technology, Zhengzhou, Henan 450007, PR China
[c] Department of Computer Science and Information Systems, Birkbeck College, University of London, Bloomsbury, London WC1E 7HX, UK
[d] Information Center of Yellow River Conservancy Commission, Zhengzhou, Henan 450003, PR China

ARTICLE INFO

Article history:
Received 14 September 2012
Accepted 23 March 2014
Available online 3 April 2014

Keywords:
Supervised image classification
Two-dimensional random projection
Compressive sampling
ℓ_1-normminimization
ℓ_0-norm sparse presentation
Constrained subspace
Affine hull
Intrinsic dimension estimation

ABSTRACT

We consider the problem of image recognition via two-dimensional random projection and nearest constrained subspace. First, image features are extracted by a two-dimensional random projection. The two-dimensional random projection for feature extraction is an extension of the 1D compressive sampling technique to 2D and is computationally more efficient than its 1D counterpart and 2D reconstruction is guaranteed. Second, we design a new classifier called NCSC (Nearest Constrained Subspace Classifier) and apply it to image recognition with the 2D features. The proposed classifier is a generalized version of NN (Nearest Neighbor) and NFL (Nearest Feature Line), and it has a close relationship to NS (Nearest Subspace). For large datasets, a fast NCSC, called NCSC-II, is proposed. Experiments on several publicly available image sets show that when well-tuned, NCSC/NCSC-II outperforms its rivals including NN, NFL, NS and the orthonormal ℓ_2-norm classifier. NCSC/NCSC-II with the 2D random features also shows good classification performance in noisy environment.

1. Introduction

For most practical pattern recognition scenarios, feature extraction and classification methods are equally important. Feature extraction should retain most if not all of the useful information in the data while keeping the dimension of the features as low as possible. A careful choice of features is required to achieve low complexity in the classifier and a high accuracy in classification.

1.1. Compressive sampling

Recent developments of compressive sampling (CS) theory give us clues for new methods of feature extraction. Namely, if the sparsity of the data is appropriately harnessed, then the data can be highly compressed by an underdetermined random projection (defined by a full rank random matrix whose row number is less than its column number), to achieve a sampling rate even lower than the classical Nyquist rate without any information loss. The original data can be exactly recovered from the highly compressed measurements by the ℓ_1-norm minimization techniques [1–9].

More specifically, let $\mathbf{x} \in \mathbb{R}^D$ be a κ-sparse ($\kappa < D$) vector, i.e., $\mathbf{x}$ has at most κ nonzero entries, and let $\mathbf{\Phi} \in \mathbb{R}^{d \times D}$ ($d < D$) be a matrix, whose entries are Gaussian distributed (or more generally, Restricted Isometry Property compatible). Then $\mathbf{x}$ can be compressed as follows.

$$\hat{\mathbf{x}} = \mathbf{\Phi} x \quad (1)$$

where $\hat{\mathbf{x}} \in \mathbb{R}^d$ is the vector of CS measurements.

Given $\hat{\mathbf{x}}$ and $\mathbf{\Phi}$, there are an infinite number of vectors $\mathbf{x}$ that satisfy Eq. (1). However, it has been proved that if $d \geqslant O(\kappa \log(\frac{D}{\kappa}))$, then with overwhelming high probability

$$p \geqslant 1 - \exp O(-d) \quad (2)$$

$\mathbf{x}$ can be exactly recovered from $\hat{\mathbf{x}}$ by minimizing the ℓ_0-norm of $\mathbf{x}$ as follows [1,2].

$$\mathbf{x}^* = \underset{\mathbf{x} \in \mathbb{R}^D}{\operatorname{argmin}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \hat{\mathbf{x}} = \mathbf{\Phi}\mathbf{x} \quad (3)$$

where $\mathbf{x}^*$ is the recovered version of $\mathbf{x}$.

Because the optimization problem of Eq. (3) is NP-hard, the recovery of $\mathbf{x}$ is equivalently reformulated as the ℓ_1-norm minimization problem as follows.

$$\mathbf{x}^* = \underset{\mathbf{x} \in \mathbb{R}^D}{\operatorname{argmin}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \hat{\mathbf{x}} = \mathbf{\Phi}\mathbf{x} \quad (4)$$

* Corresponding author.
E-mail addresses: liaoliangis@126.com (L. Liao), ynzhang@nwpu.edu.cn (Y. Zhang), sjmaybank@dcs.bbk.ac.uk (S.J. Maybank), lzhoufeng@hotmail.com (Z. Liu), liuxinis@126.com (X. Liu).

http://dx.doi.org/10.1016/j.jvcir.2014.03.007
1047-3203/© 2014 Published by Elsevier Inc.

This problem can be solved by algorithms such as Basis Pursuit [10] or Orthogonal Matching Pursuit [11].

Since the data dimension can be efficiently reduced without significant information loss, the above mentioned projection technique serves as a tool for feature extraction. Mathematical analyses show that compressive recognition, detection and other processings in compression domain $\mathbb{R}^d$ are feasible [12–20].

Note that the above mentioned projection technique is applied to vectors. For image data which are naturally represented by matrices, the 1D representation discards structural information about the image.

Due to this concern, different 2D (matrix) representations are exploited for feature extraction. For example, the 2D representation methods include 2DPCA [21] and its variants [22–24], 2DLDA [25], bilinear subspace learning [26,27], tensor analysis [28–31], and the recent common interest in 2D random projection [32–34].

Among the 2D representations, either supervised or unsupervised, 2DPCA, 2DLDA and bilinear subspace learning, etc., are obtained by deterministic two-dimensional projection. Another category of 2D representation include those obtained by random linear projection[32–34]. Both categories are actually the order-two tensor analyses, which exploit the correlations among image pixels with different dimensions and in this sense is believe to lead to good classification performance for different applications such as image recognition and human gait recognition [28–31].

1.2. NN, NFL and NS

Besides the feature extraction, classifier design is equally important. Classical but still popular subspace-based classifiers include NN (Nearest Neighbor), NFL (Nearest Feature Line, proposed by Li et al. [35]) and NS (Nearest Subspace).

NN, NFL and NS share some common traits and can be summarized in a generalized way — given a query sample $\mathbf{y}$ and n training samples belonging to K classes, NN, NFL and NS all use on the same strategy to determine the class of $\mathbf{y}$ as follows.

$$\begin{cases} r_i(\mathbf{y}) = \min\limits_{\mathbf{x}\in\mathbb{M}_i}\|\mathbf{y}-\mathbf{x}\|_2, \quad \forall i = 1,\ldots,K \\ \text{class}(\mathbf{y}) = \underset{i\in\{1,\ldots,K\}}{\text{argmin}}\, r_i(\mathbf{y}) \end{cases} \tag{5}$$

where $r_i(\mathbf{y})$ is the distance of $\mathbf{y}$ to class i and $\mathbb{M}_i$ is a classifier-specific dataset defined by training set i.

Denoting the ith training set by $\mathbb{X}_i = \left\{\mathbf{x}_i^{(1)},\ldots,\mathbf{x}_i^{(n_i)}\right\}$, in NN, $\mathbb{M}_i$ is the i-th training set itself, namely

$$\mathbb{M}_i = \mathbb{X}_i \tag{6}$$

In NFL, $\mathbb{M}_i$ is a set of feature lines defined by $\mathbb{X}_i$, namely,

$$\mathbb{M}_i = \left\{\alpha\mathbf{x}^{(a)} + (1-\alpha)\mathbf{x}^{(b)} | \alpha \in \mathbb{R},\ \mathbf{x}^{(a)},\mathbf{x}^{(b)} \in \mathbb{X}_i\right\} \tag{7}$$

In NS, $\mathbb{M}_i$ is the linear subspace spanned by $\mathbf{x}_i^{(1)},\ldots,\mathbf{x}_i^{(n_i)}$. Denote the matrix whos columns are the training samples of the ith class by

$$\mathbf{A}_i = \left[\mathbf{x}_i^{(1)},\ldots,\mathbf{x}_i^{(n_i)}\right] \tag{8}$$

then, in NS, $\mathbb{M}_i$ can be written as follows.

$$\mathbb{M}_i = \{\mathbf{A}_i\boldsymbol{\alpha} | \boldsymbol{\alpha} \in \mathbb{R}^{n_i}\} \tag{9}$$

It follows from Eqs. (6)–(9), that in all cases we have $\mathbb{X}_i \subseteq \mathbb{M}_i$. For notation convenience, we respectively denote the training superset $\mathbb{M}_i$ of NN, NFL and NS as $\mathbb{M}_i^{\text{NN}}, \mathbb{M}_i^{\text{NFL}}$ and $\mathbb{M}_i^{\text{NS}}$. It is not difficult to see that $\mathbb{M}_i^{\text{NN}} \subset \mathbb{M}_i^{\text{NFL}} \subset \mathbb{M}_i^{\text{NS}}$. Since $\mathbb{M}_i^{\text{NS}}$ is a linear subspace and $\mathbb{M}_i^{\text{NN}}$ and $\mathbb{M}_i^{\text{NFL}}$ are just the appropriate subsets of it, we call $\mathbb{M}_i^{\text{NN}}$ and $\mathbb{M}_i^{\text{NFL}}$ *the constrained subspaces* for the ith class.

1.3. NM and its relationship to NN, NFL, NS

In NM (Nearest Manifold), it is assumed that the data of a class lie on or near to a manifold, and that the dimension of the manifold is much less than the dimension of the feature space.

NM uses the same strategy of Eq. (5) to classify $\mathbf{y}$ with

$$\mathbb{M}_i = \mathcal{M}_i, \quad i = 1,\ldots,K. \tag{10}$$

where $\mathcal{M}_i$ is the data manifold of the ith class.

If suitable manifolds for all $i = 1,\ldots,K$ can be found, then NM has a high classification accuracy.

Note that $\mathbb{M}_i$ in Eqs. (6), (7) and (9) can be viewed as different approximations to $\mathcal{M}_i$ for the ith class. From this perspective, we contend that NN, NFL and NS are all approximations to NM and propose later a novel classifier, called NCSC (Nearest Constrained Subspace Classifier), and show by experiments that NCSC is a better approximation to NM than NN, NFL and NS.

1.4. Contributions of this study

Based on our previous work [34], we discuss the technique of 2DCS (two-dimensional compressive sampling), which is inspired by 1DCS (traditional compressive sampling) and 2DPCA [21]. The 2D (matrix based) approach is computationally less complex than the 1D (vector based) approach to image data. The reconstruction of the original data is still guaranteed with a high probability. In this sense, 2DCS is more efficient than 1DCS for feature extraction. Our experiments show that when 2DCS features are exploited by some state-of-the-art classifiers, the performance of image recognition is improved.

This interest is somehow shared almost at the same time by Eftekhari et al. [32] and Leng et al. [33]. Although addressing the same problem, the focuses of Eftekhari et al. [32] and Leng et al. [33] and ours are different. Besides the theoretical analysis of 2D random projection and the assumption that 2D signal is sparse, Eftekhari et al. reported a reconstruction algorithm of 2D sparse signal based on smoothed ℓ_0-norm minimization. A reconstruction method of natural images (not explicitly sparse) and the problem of designing a cutting-edge classifier exploiting the 2D random projection features were not addressed in [32]. On the other hand, in [33], 2D random projection and its variations combined with PCA, LDA etc., were studied and compared with other feature extractors but without mention of the problems of 2D reconstruction and classifier design.

In our work, we propose a two-steps (including row processing and column processing) 2DCS reconstruction scheme for natural images via TV minimization. We also propose a classifier called NCSC (Nearest Constrained Subspace Classifier) and its fast version called NCSC-II, in which the subspace associated with the target class is constrainedly spanned by training samples. The constrained subspace is a union of a series of affine hulls.

We prove that NCSC is a generalized version of NN (Nearest Neighbor), NFL (Nearest Feature Line) and has a close relationship with NS (Nearest Subspace). Employing the intrinsic dimension as a freedom degree parameter, the constrained subspace, rather than the unconstrained one, is believed to be a more accurate approximation to the data manifold. The intrinsic dimension of the constrained subspace in NCSC is defined by a ℓ_0-norm sparse representation, and NCSC itself is in fact an approximation to the conceptual NM (Nearest Manifold) classifier, which is believed to