J. Vis. Commun. Image R. 24 (2013) 181–186

Contents lists available at SciVerse ScienceDirect

J. Vis. Commun. Image R.

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

Dictionary learning and similarity regularization based image noise reduction

Shuyuan Yang [a,*], Linfang Zhao [a], Min Wang [b], Yueyuan Zhang [a], Licheng Jiao [a]

[a] Key Lab of Intelligent Perception and Image Understanding of Ministry of Education, Department of Electrical Engineering, Xidian University, Xi'an 710071, China
[b] National Key Lab of Radar Signal Processing, Department of Electrical Engineering, Xidian University, Xi'an 710071, China

ARTICLE INFO

Article history:
Available online 3 August 2012

Keywords:
Image denoising
Sparse learning
Image patches
Dictionary learning
Self-similarity
Non-local means
KSVD
Non-parametric Bayesian

ABSTRACT

Redundant dictionary learning based image noise reduction methods explore the sparse prior of patches and have proved to lead to state-of-the-art results; however, they do not explore the non-local similarity of image patches. In this paper we exploit both the structural similarities and sparse prior of image patches and propose a new dictionary learning and similarity regularization based image noise reduction method. By formulating the image noise reduction as a multiple variables optimization problem, we alternately optimize the variables to obtain the denoised image. Some experiments are taken on comparing the performance of our proposed method with its counterparts on some benchmark natural images, and the superiorities of our proposed method to its counterparts can be observed in both the visual results and some numerical guidelines.

1. Introduction

Noise reduction of images aims to remove unknown noise from a measured corrupted image, which is a challenging problem in image processing and computer vision. Considering the classic image noise reduction problem: an image $\mathbf{X}$ is measured in the presence of an additive zero-mean white and homogeneous Gaussian noise n, and we want to recover the original image $\mathbf{X}$ from the noisy image $\mathbf{Y} = \mathbf{X} + n$. With the growing realization of efficiency of redundancy in denoising images, numerous methods have been proposed, including some redundant multiscale transforms [1–6] and the very recent developed spatial redundant dictionary based sparse representation methods [7–9]. In the past decade, using spatial redundant representation and sparsity for images denoising has drawn much attention of researchers. Its basic idea is that the sparse representation of images will help to automatically select the primary components in images while reducing the noise components, as long as the dictionary can well describe the characteristics of images. In more recent works [10–14], small image patches prove to be able to represent the statistical properties of the whole image, from which a dictionary can be learned to recover the noisy patches via sparse representation. That is, they assume that all the image patches have the identical statistical distribution, which is the foundation of the dictionary learning.

Although sparse representation and dictionary learning based methods have proved to lead to state-of-the-art denoising results, they only explore the sparse prior of images. However, there is also similarity about the structure of images. In fact there are often many repetitive image structures (or self-similarity) in an image, especially for natural images. When an image is divided into patches, for each local patch, we can find similar patches in the whole image according to Gaussian neighborhood (in practice, in a sufficiently large area). Such non-local structural similarity redundancy is very helpful to improve the quality of reconstructed images.

The self-similarity of image patches has been well known and successfully increases the efficiency of many image processing tasks [15,16]. In paper [17], a simultaneous sparse coding is proposed to combine the self-similarities of natural images. However, it explores the similarities in the sparse coding of patches and thus is of high computational complexity. In this paper, we explore the self-similarities in a simpler way, and introduce another non-local self-similarity regularizer into the dictionary learning based denoising method to improve the quality of reconstructed images. By formulating the image noise reduction as a multiple variables optimization problem, we alternately optimize the variables to obtain the denoised image. The contribution of the proposed method has the following characteristics: (1) We use both the sparse prior and similarity prior in estimating the original image by combining the dictionary learning with the regularization method; (2) An optimization algorithm is proposed to alternately tune the multiple variables to obtain the denoised image. On the one hand, the nonparametric Bayesian dictionary learning method has the ability to automatically decide the noise level and does not require parameter tuning for images with difference noise statistics. The prior assumption of noise is considered and a truncated beta-Bernoulli process is used to establish a hierarchical Bayesian model. So the noise variance can be accurately estimated from the example patches. On the other hand, the self-similarity prior can help to further improve the image quality in patch aggregation. Therefore the state-of-the-art image reduction

* Corresponding author.
E-mail address: syyang@xidian.edu.cn (S. Yang).

1047-3203/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jvcir.2012.07.011

result can be obtained. Some experiments are taken on comparing the performance of our proposed method with its counterparts on some benchmark natural images, and the superiorities of our proposed method to its counterparts can be observed in both the visual results and some numerical guidelines.

The rest of this paper is organized as follows. In Section II, we depict the dictionary learning and non-local similarity regularizer based denoising approach. In section III, some simulation experiments are taken to illustrate the efficiency and superiority of our proposed method to its counterparts. Finally some conclusions are drawn in Section IV.

2. Image noise reduction via dictionary learning and non-local similarity regularizer

Considering the classic image noise reduction problem: an image is measured in the presence of an additive zero-mean white and homogeneous Gaussian noise, with standard deviation σ. Thus the measured image is $\mathbf{Y} = \mathbf{X} + n$. The goal of image noise reduction is to recover the original image $\mathbf{X}$ from the noisy image $\mathbf{Y}$.

2.1. Sparse representation and dictionary learning based image noise reduction

Assume that the measured image $\mathbf{Y} = \mathbf{X} + n$ belongs to the (ε, L, D)-*Sparseland* [10] signals, that is, the original clean image can be represented sparsely over a dictionary $\mathbf{D} : \mathbf{X} = \mathbf{D}\boldsymbol{\alpha}$, where $\mathbf{D} \in \Re^{n\times K}$ $(K > n)$ is a redundant dictionary that is constructed by some atoms, $\|\mathbf{n}\|_p^2 < \varepsilon$, $\boldsymbol{\alpha}$ is the sparse representation coefficient vector. The image can be denoised by solving such an optimization problem,

$$\hat{\boldsymbol{\alpha}} = \begin{cases} \min\limits_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \\ \text{s.t. } \|\mathbf{Y} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \leqslant \varepsilon \end{cases} \tag{1}$$

where $\boldsymbol{\alpha}$ is the sparse representation of the image **Y**, $\|\cdot\|_0$ counts non zero coefficients which is called l_0 norm, and the denoised image is given by $\hat{\mathbf{X}} = \mathbf{D}\hat{\boldsymbol{\alpha}}$.

Given a set of $\sqrt{p} \times \sqrt{p}$ patches extracted from the image $\mathbf{X}$, we consider the sparse representation of them under a small dictionary $\mathbf{D}$, and the denoising is performed patch by patch. Denote the matrix $\boldsymbol{R}_{ij}$ as a $p \times N$ matrix that extracts the (i,j) patch from the image $\mathbf{X}$, $\boldsymbol{\alpha}_{ij}$ is the coefficient for patch $\boldsymbol{R}_{ij}\mathbf{X}$. In order to avoid the artifacts, the overlapping patches are extracted and then the denoised patches are aggregated to recover the original image $\mathbf{X}$. For an $\sqrt{N} \times \sqrt{N}$ image, the summation of the blocks from the image includes $P = (\sqrt{N} - \sqrt{p} + 1)^2$ items when all image patches of size $\sqrt{p} \times \sqrt{p}$ with overlaps. Then the MAP estimator for denoising the (i,j) patch is built by solving the following formula [38],

$$\begin{cases} \min\limits_{X,\{\boldsymbol{\alpha}_{ij}\}} \sum\limits_{ij} \|\boldsymbol{\alpha}_{ij}\|_0 \\ \text{s.t. } \sum\limits_{ij} \|R_{ij}X - D\boldsymbol{\alpha}_{ij}\|_2^2 \leqslant \varepsilon_1; \quad \|X - Y\|_2^2 \leqslant \varepsilon_2 \end{cases} \tag{2}$$

where ε_1, ε_2 are the admissible errors. However, the dictionary is fixed. In order to learn a good dictionary $\mathbf{D}$, a set of $\sqrt{p} \times \sqrt{p}$ patches $\mathbf{Q} = \{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m | \mathbf{q}_i \in \Re^p\}$ sampled from some training images can be used. Assuming that $\mathbf{Q}$ comes from the noisy image and $\mathbf{Q}$ can be represented as a sparse linear combination by the dictionary $\mathbf{D} = [\boldsymbol{d}_1, \ldots, \boldsymbol{d}_K] \in \Re^{p\times K}$, and (2) can be reformulated as,

$$\min_{\mathbf{X},\mathbf{D},\{\boldsymbol{\alpha}_{ij}\}} \sum_{ij} \|\boldsymbol{\alpha}_{ij}\|_0 + \lambda_1 \sum_{ij} \|\boldsymbol{R}_{ij}\mathbf{X} - \mathbf{D}\boldsymbol{\alpha}_{ij}\|_2^2 + \lambda_2 \|\mathbf{X} - \mathbf{Y}\|_2^2 \tag{3}$$

where λ_1, λ_2 are the Lagrange parameters.

2.2. Self-similarity regularizer

Just as we have mentioned above, there are often many repetitive image structures (or self-similarity) in an image, especially for natural images. Such non-local redundancy is very helpful to improve the quality of reconstructed images. In this section, we will introduce another non-local regularizer based on self-similarity of images into (3) to better preserve edge sharpness. For each local patch $\mathbf{x}$ in the image $\mathbf{X}$ especially natural image, we can find similar patches in the whole image according to Gaussian neighborhood (in practice, in a sufficiently large area around G). The value of the ith pixel in $\mathbf{x}$, $\mathbf{x}^i$, can then be regarded as a mean of the values of all points whose Gaussian neighborhood looks like the neighborhood of $\mathbf{x}^i$, that is,

$$\mathbf{x}^i = \sum_{(\mathbf{x})' \in G} (\mathbf{x}^i)' w' \tag{4}$$

where $(\mathbf{x}^i)'$ is the value of the ith pixel in the patch $\mathbf{x}'$ that belongs to the neighborhood G, and w' is the corresponding connected weights, which is calculated by

$$w' = \frac{\exp(-\|\mathbf{x}' - \mathbf{x}\|/h)}{\sum_{\mathbf{x}' \in G} \exp(-\|\mathbf{x}' - \mathbf{x}\|/h)} \tag{5}$$

The formula (5) amounts to say that each recovered value $\mathbf{x}^i$ in $\mathbf{x}$ is a mean of the values of all points $(\mathbf{x}^i)'$ whose Gaussian neighborhood looks like the neighborhood of $\mathbf{x}$. It makes a systematic use of all possible self-predictions the image can provide, in the spirit of self-similarity of images.

Assume that the number of pixels in the non-local area G is L, i.e., $|G| = L$, and denote the neighbored vector of $\mathbf{x}^i$ as $\mathbf{Nx}^i \in \Re^{1\times L}$ and the corresponding weight vector as $\mathbf{w}^i \in \Re^{L\times 1}$, the formula (4) can be rewritten as

$$\mathbf{x}^i = \mathbf{Nx}^i \times \mathbf{w}^i. \tag{6}$$

Then the ith patch $(\mathbf{x})_i$ can be described as,

$$(\mathbf{x})_i = [\mathbf{Nx}^1, \mathbf{Nx}^2, \ldots, \mathbf{Nx}^n]^T \cdot \begin{bmatrix} \mathbf{w}^1 \\ \mathbf{w}^2 \\ \vdots \\ \mathbf{w}^n \end{bmatrix} = (\mathbf{NX})_i \cdot (\mathbf{W})_i \tag{7}$$

where $(\mathbf{NX})_i$ is the neighborhood matrix of patch $(\mathbf{x})_i$, and $(\mathbf{W})_i$ is the weighted matrix of patch $(\mathbf{x})_i$. Both the arrangement of $(\mathbf{NX})_i$ and the calculation of $(\mathbf{W})_i$ are determined by the other patch $(\mathbf{x})_j$ in the non-local area G, that is, $(\mathbf{x})_i = t((\mathbf{x})_j)$; $(\mathbf{x})_j \in G$, where t is denoted by a non-local means operation in G. The recovered image can then be written as

$$\mathbf{X} = [(\mathbf{x})_i, \ldots, (\mathbf{x})_P] = \mathbf{D}[\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_P] = \mathbf{D}\boldsymbol{\alpha} = T(\mathbf{D}\boldsymbol{\alpha}) = T(\mathbf{X}) \tag{8}$$

where T represents a non-local means operation on $\mathbf{X}$. This constraint (8) can be served as a new regularizer term in the *denoising* problem (3). Adding the formula (8) on (3), we get

$$\min_{\mathbf{X},\mathbf{D},\{\boldsymbol{\alpha}_{ij}\}} \sum_{ij} \|\boldsymbol{\alpha}_{ij}\|_0 + \lambda_1 \sum_{ij} \|\boldsymbol{R}_{ij}\mathbf{X} - \mathbf{D}\boldsymbol{\alpha}_{ij}\|_2^2 + \lambda_2 \|\mathbf{X} - \mathbf{Y}\|_2^2 + \lambda_3 \|\mathbf{X} - T(\mathbf{X})\|_2^2 \tag{9}$$

It exploits the *non-local similarity* constraints of images in the dictionary learning, and reduces the denoising problem to a multiple variables optimization problem. In the following we will depict the procedure of the proposed algorithm.

2.3. Procedure of the optimization algorithm

From the formula (9) we can see that there are three set of variables to be optimized, and we can observe the training data set Q is