Image sparse representation with local ARMA and nonlocal self-similarity regularizations for super-resolution

Weirong Liu a,*, Chaopeng Zhang a, Jie Liu b, Chaorong Liu c

a College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China
b National Demonstration Center for Experimental Electrical and Control Engineering Education, Lanzhou University of Technology, Lanzhou 730050, China
c Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou University of Technology, Lanzhou 730050, China

A R T I C L E   I N F O

Keywords:
Super-resolution
Spare representation
ARMA model
Nonlocal self-similarity

A B S T R A C T

Since the single image super-resolution (SR) is an extremely ill posed problem, we introduce a novel autoregressive moving average (ARMA) model-based regularization term into the spare representation-based framework to deal with it in this paper. In our framework, we have a dual regularization. Firstly, we use the ARMA models trained from external samples to establish a regularization term. ARMA model-based regularization serves as a local constraint. Secondly, we introduce the nonlocal (NL) self-similarity as another regularization term. Both the local and the NL regularizations are unified into the sparse representation-based framework. Finally, extensive experiments verify the effectiveness of the proposed method.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Image SR is a methodology to overcome the inherent resolution limitation of low resolution (LR) imaging systems [1]. Image SR techniques always recover a high resolution (HR) image from one or a series of down-sampled and blurred LR images. The development of SR techniques is very helpful in medical imaging, satellite imaging and so on. In general, methods of SR can be roughly categorized into three classes: reconstruction-based approaches [2–4], interpolation-based approaches [5,6] and learning-based approaches [7–11]. Firstly, we briefly summarize the sparse representation-based SR methods as follows.

The sparse representation-based methods focus on how to effectively learn the mapping relationship between the LR and HR image patches. The learned mapping is employed to reconstruct a HR image. Recently, numerous sparse representation-based methods [7–11] have been proposed. For example, Yang et al. [7] learned a HR and LR dictionary pair from an external training set via sparse coding. They recovered a HR image using the learned dictionary pair under the assumption that the HR patch and its corresponding LR counterpart can share the same sparse code. Although their approach performs well, it is not very robust to the change of the training set. Yang et al. provided another approach that utilized the self-samples learning in [9]. They generated image patch pairs directly from an image pyramid with a single frame LR input image. These image patch pairs were clustered to train a dictionary by enforcing group sparsity constraints, and then they constructed HR images via group structured sparse coding. Their method can well deal with the change of the training set. The unitary samples and the flexibility of dictionary may limit the application of this method. Moreover, their method leads to a very high computational complexity. For fast super-resolution, Timofte et al. [10] utilized the ridge regression to learn exemplar neighborhoods in order to precompute the projections which map LR patches to the HR domain. The regressors were incorporated into the sparse dictionary. Their method was named as ANR. Soon afterwards, they [11] proposed the ANR with simple functions. It is an improved variant of ANR. Their algorithm substantially reduces the computational complexity and achieves better performance than ANR. To achieve adaptive sparse coding, Dong et al. [8] introduced an adaptive sparse domain selection (ASDS) scheme. In addition, they proposed two adaptive regularization terms: the autoregressive (AR) model regularization and the NL self-similarity regularization. ASDS is superior to many existing SR algorithms in association with the above two adaptive regularization terms. The question is that AR may lead to overfitting. Since the forecasting accuracy of AR model is not as accurate as ARMA model, we focus on ARMA model to promote the forecasting accuracy. Moreover, we consider the local image structure regularization. The learned ARMA models from external training set are

* Corresponding author.
E-mail address: liu_weirong@163.com (W. Liu).

http://dx.doi.org/10.1016/j.optcom.2017.07.031
Received 16 February 2017; Received in revised form 6 July 2017; Accepted 9 July 2017
Available online xxxx
0030-4018/© 2017 Elsevier B.V. All rights reserved.
employed to predict the local image structures of HR images, so that we can utilize the learned ARMA models to construct a local constraint.

Since the proposed method is based on the sparse representation, it is necessary to review the sparse prior. The main task of single image SR is to reconstruct a visually fine HR image from a blurred and down-sampled LR image in noisy environment [1]. Let \( y \) denote an observed LR image and \( x \) denote its corresponding HR counterpart. The relationship between LR image \( y \) and HR image \( x \) can be formulated as

\[
y = DBx + v
\]

where \( D \) is a down-sampling operator, \( B \) is a blurring operator and \( v \) is an additive noise. Eq. (1) is an inverse problem with an ill-posed characteristic. Plenty of regularization techniques have been proposed to regularize Eq. (1). Among the popular regularization terms, total variation (TV) is a notable technique. TV utilizes the total variation of signal as a regularization function [12], however, it tends to over-smooth the reconstructed images. Hence, the prior knowledge of image is helpful in SR reconstruction. The sparsity prior is one of the most notable priors, which usually contributes to state-of-art results [7]. In the sparsity prior, the patches of image \( x \) can be represented as a sparse linear combination via a dictionary \( \mathbf{\Theta} \). Therefore, we have the representation \( x = \Phi \alpha \), where \( \alpha \) is the sparsity coefficients. It is easy to incorporate the sparsity prior into image SR, so that the reconstruction question can be rewritten as the sparse coding question

\[
\hat{\alpha} = \arg \min_{\alpha} \{ \| y - D \Phi \alpha \|_2^2 + \lambda \| \alpha \|_1 \}
\]

where the \( \ell_1 \)-norm counts the number of nonzero coefficients in vector \( \alpha \) and \( \lambda \) is a constant that balances sparsity of the solution and the fidelity term.

Although the sparsity prior is very useful in SR reconstruction, it is not enough to recover a better HR image. Since ARMA modeling methods have been successfully used to predict electricity prices [13,14] and stock market [15] etc., we can constrain the local image structure in a small neighborhood with ARMA model. As a complement of local constraint, the NL self-similarity is also introduced as a NL constraint. Therefore, the main contribution of this paper is that we construct an ARMA model-based local regularization and unify local regularization with NL regularization into the sparse representation-based framework.

We utilize the K-means clustering technique to categorize the selected example image patches into \( K \) clusters, and principal component analysis (PCA) technique to learn a sub-dictionary for each cluster. In our implementation, we also train an ARMA model for each cluster in the training stage. Subsequently, we adaptively select ARMA models and construct adaptive local and NL [16] regularization terms to regularize the ill-posed problem of image SR in sparse coding stage.

The remainder of this paper is organized as follows. Section 2 provides the related background. Section 3 introduces the proposed ARMA model and NL-based sparse representation algorithm. Section 4 describes the summarization of the proposed algorithm. Section 5 presents the numerical and simulated results. Section 6 gives the conclusion of this paper.

2. Related background

In the sparse coding problem, Eq. (2) is a non-convex minimization question with a \( \ell_1 \)-norm. It is common that the \( \ell_0 \)-minimization is replaced by the \( \ell_1 \)-minimization, such as [8]. Thus, Eq. (2) can be rewritten as

\[
\hat{\alpha} = \arg \min_{\alpha} \{ \| y - D \Phi \alpha \|_2^2 + \lambda \| \alpha \|_1 \}.
\]

Consequently, the estimated HR image can be recovered via \( \ell_1 \)-minimization.

The quality of the reconstructed HR image is highly dependent on the learned dictionary, which plays a key role in the sparse representation-based image SR. These existing dictionaries are mainly divided into two categories: the analytical dictionary and the learning dictionary. The analytically designed dictionary, such as DCT, wavelet, curvelet, and contourlet can be easily acquired. However, they have limited abilities to represent the different types of data. The learning-based dictionary usually focus on learning a universal overcomplete dictionary for various image structures, such as the coupled dictionaries trained from LR and HR image patch pairs proposed by Yang et al. [7]. However, sparse decomposition over a highly redundant dictionary is potentially unstable and tends to cause visual artifacts [17]. Recently, researches have focused on adaptive sparse coding using the dictionary learned in LR and HR spaces. For instance, Dong et al. [8] proposed an ASDS scheme to sparsely code image patches with the learned PCA sub-dictionaries. Suppose that \( \{ \Phi_k \}^K_{k=1} \in \mathbb{R}^{n \times n} \) includes \( K \) orthonormal PCA sub-dictionaries where the \( K \) sub-dictionaries come from the \( K \) different sample clusters respectively. Let \( x \in \mathbb{R}^N \) be a vector of an image and \( R \in \mathbb{R}^{N \times N} \) be a patch extraction operator. \( x = R \cdot x \in \mathbb{R}^n \), symbolizes the \( i \)-th patch of \( x \). If a sub-dictionary \( \Phi_k \) is adaptively selected to code \( x \) sparsely, the representation of \( x \) can be written as \( x = \hat{\Phi}_k \alpha \), where \( \hat{\Phi}_k \) is an estimation of \( \Phi_k \) and \( \alpha \) is the sparse coefficients of \( x \).

For the convenience of expression, we define the “o” operator

\[
\hat{x} = \Phi \circ \alpha = \left( \sum_{i=1}^{N} R_i^2 \right)^{-1/2} \sum_{i=1}^{N} R_i^2 \Phi_i \alpha_i
\]

where \( \Phi \) is the concatenation of all sub-dictionaries \( \{ \Phi_k \}^K_{k=1} \) and \( \alpha \) is the concatenation of all \( \alpha_i \). With Eq. (4) the sparse coding problem can be reformulated as

\[
\hat{\alpha} = \arg \min_{\alpha} \{ \| y - D \Phi \circ \alpha \|_2^2 + \lambda \| \alpha \|_1 \}.
\]

3. The proposed method

3.1. Sub-dictionaries training

We present our PCA sub-dictionaries learning method in this section. In order to learn \( K \) compact sub-dictionaries \( \{ \Phi_k \}^K_{k=1} \), we collect and crop some HR images to get image patches. We randomly select \( o \) patches \( S = \{ s_1, s_2, \ldots, s_o \} \) where \( s_j \in \mathbb{R}^{N \times N} \) is the cropped image patch from the collected images. Note that we divide the patches into two sets according to whether their intensity variances are greater than \( \Delta \) or not, where \( \Delta \) is a presupposed threshold. In addition, to enhance the robustness, we employ a high-pass filter to extract the high frequency components that benefit for sparse coding before sub-dictionary learning. Denote \( S_A = \{ s_1^A, s_2^A, \ldots, s_o^A \} \) as the filtered result of \( S \). By introducing K-means clustering algorithm, dataset \( S_A \) can be divided into \( K \) clusters \( \{ C_1, C_2, \ldots, C_K \} \). Denote \( \mu_k \) as the centroid of the cluster \( C_k \). The clustering process can be mathematically written as:

\[
k = \arg \min_{k \in \{1,2,\ldots,K\}} \left\| x^k - \mu_k \right\|_2.
\]

We can get \( K \) subsets via Eq. (6). Denote \( S_I \in \mathbb{R}^{n \times m_I} \) as the \( k \)-th subset of \( S \) where \( m_I \) is the number of samples in \( S_I \).

We introduce the PCA technique to learn sub-dictionaries. The PCA is widely used for signal de-correlation and dimensionality reduction in pattern recognition. We employ PCA to construct the sub-dictionary \( \Phi_k \) from the corresponding subset \( S_I \). Denote \( \Theta_k \) as the co-variance matrix of the dataset \( S_I \). By applying the PCA to \( \Theta_k \), an orthogonal transformation matrix \( P_k \) can be obtained. We only extract the first \( m \) most important eigenvectors in \( P_k \) to construct a sub-dictionary \( \Phi_k \). Let \( A_k \) denote a diagonal matrix. \( A_k \) consists of the eigenvalues of \( \Theta_k \). We can learn each sub-dictionary via

\[
\Theta_k = P_k A_k P_k^T, \quad \Phi_k = [p_1, p_2, \ldots, p_L], \quad p_L \in \mathbb{R}^n
\]

where \( p_L \) is the \( k \)-th most important eigenvector.
دریافت فوری متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات