Signal Processing 104 (2014) 15–26

Contents lists available at ScienceDirect

Signal Processing

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

Iterative gradient projection algorithm for two-dimensional compressive sensing sparse image reconstruction

Gao Chen, Defang Li, Jiashu Zhang *

Sichuan Province Key Lab of Signal and Information Processing, Southwest Jiaotong University, Chengdu 610031, People's Republic of China

ARTICLE INFO

Article history:
Received 30 August 2013
Received in revised form
3 March 2014
Accepted 27 March 2014
Available online 3 April 2014

Keywords:
Compressive sensing
Random projections
Sparsity
Reconstruction

ABSTRACT

The basic theories and techniques in compressive sensing (CS) are established on the sampling and reconstruction of one-dimensional (1D) signals. When it is applied to two-dimensional (2D) images, the images are first stacked in a large vector. However, this vectorization not only destroys the spatial structure of the 2D image, but also increases computational complexity and memory requirements. As a result, some researchers proposed the concept of 2D CS. The major challenge of 2D CS is to design a reconstruction algorithm that can directly reconstruct the 2D image data from the 2D random projection. In this paper, a 2D CS sparse image reconstruction algorithm based on iterative gradient projection is proposed. In the proposed algorithm, the sparse solution is searched iteratively in the 2D solution space and then updated by gradient descent of the total variation (TV) and bivariate shrinkage in the dual-tree discrete wavelet transform (DDWT) domain. Numerous experiments are performed on several natural images. Compared with several state-of-the-art reconstruction algorithms, the proposed algorithm is more efficient and robust, not only yielding higher peak-signal-to-noise ratio but also reconstructing images of better subjective visual quality.

1. Introduction

As a fundamental problem in the field of image processing, image compression has been extensively studied in the past two decades [1–4]. The emerging compressive sensing or compressive sampling (CS) theory [5] has pointed us a promising way of developing novel efficient data compression techniques. The CS is a novel technique for signal acquiring, processing and compression. It states that as long as the signal is sparse or compressible, then we can roughly recover the original signal through only a few measurements, namely, we can achieve data sampling and compression at the same time. Due to its advantages of sampling below Nyquist rate and little loss in reconstruction quality, research on CS based image/video coding has become a hot topic in recent years. Han et al. [6] proposed an image representation method for visual sensor networks based on compressive sensing; Wu et al. [7] integrated a piecewise stationary auto regressive model into the recovery process for CS-coded images. Deng et al. [8,9] proposed a robust image coding method with a number of descriptions based on CS for high packet loss transmission, where the discrete wavelet transform (DWT) was applied for sparse representation. However, as the image data was huge, their reconstructions were unnecessarily expensive. This expensiveness could also lead to computational instability. From the results of [10,11], it seemes that the run time necessary to reconstruct 256×256 and 512×512 images ranges from a few minutes to one hour.

Image sparse representation is a key factor that affects the quality of the reconstructed image, the higher sparsity

* Corresponding author. Tel.: +86 28 87601756.
E-mail address: jszhang@home.swjtu.edu.cn (J. Zhang).

http://dx.doi.org/10.1016/j.sigpro.2014.03.039
0165-1684/

of an image is, the higher recovery quality it will have. Wavelet transform [12–14,44] is a good sparse representation for piecewise constant images. Many image sparse representation algorithms based on wavelet transform were proposed in the past decade [8,9,15,16]. TV [17] (total variation) regularization is another candidate for sparseness modeling of strong edges in piecewise constant images. It was first proposed for image denoising [17] and since then has been successfully used for image super-resolution [18–20]. Recently, some TV solvers have become available for CS [21–24]. However since natural images, which include a variety of features, such as rich textures, fine details and strong edges, are typically non-stationary, only using a certain regularization is difficult to effectively represent all of the image features and will result in a poor CS recovery performance. In order to have more flexibility to achieve sparser representation, some scholars put forward learning-based image representation and extended to image super resolution [25–28]. Zhang [25] proposed a new framework for image compressive sensing recovery by using adaptively learned sparsifying basis via ℓ_0 minimization, the intrinsic sparsity of natural images was enforced substantially by sparsely representing overlapped image patches using the adaptively learned sparsifying basis in the form of ℓ_0 norm. Based on the fact that small patches in natural images tend to redundantly repeat themselves, Zhang [26] proposed a single image super resolution approach by learning multiscale self-similarities from a low-resolution image itself. By applying the nonlocal means method to learn the similarity within the same scale, the proposed approach could preserve sharper edges and suppress aliasing artifacts. It seems that these learning methods adequately consider the different geometrical characteristics of images, and hence they can get more effective representation than Wavelets. However, finding a sparse representation of an image is still a remaining challenge.

Another challenge of CS-based image coding is to design an optimal sampling scheme based on the image characteristics that target to improve CS encoding efficiency. A large amount of outstanding work on this topic has been developed [29–34]. A novel sampling scheme for compressive sensing framework was proposed in [29] by designing a weighting scheme for the sampling matrix. The structure of the sampling matrix could be tuned to favor the frequency components that were important to human perception by adjusting the weighting coefficients, so that those components could be more precisely recovered in the reconstruction procedure. A reweighted sampling method based on the statistical characteristics of images was proposed in [30], in which the sampled coefficients were determined in encoding side according to the statistics of image signals, which could get much more performance gains than the other reweighted reconstruction methods. A novel block discrete cosine transform (DCT) based sampling scheme with coefficients random permutations for image compressive sensing was proposed in [31]. By introducing the coefficients random permutations in the sampling process, the proposed sampling scheme could make the significant frequency components in texture region reassign evenly in sampled blocks, so that they could be better recovered in the reconstruction process. A saliency-based compressive sampling scheme for image signals was proposed in [32]. By exploiting the saliency information of images and taking human visual attention into consideration, it improved the reconstructed image quality considerably. Comparing to frame-based CS, block-based CS (BCS) of image [33,34] was an efficient alternative for significantly reducing the computational complexity. In BCS the original image was divided into small blocks and each block was sampled independently using the same measurement operator. There is a common feature in the above-mentioned methods: in order to collect a set of linear measurements of an image, they stacked the columns of the image or image blocks into a large column vector. This so-called vector space model for image processing destroys the intrinsic image spatial structure.

To maintain the intrinsic image spatial structure, some researchers have proposed the concept of two-dimensional compressive sensing (2DCS) [35–38]. In 2DCS, 2D random projection [35] was exploited to directly leverage the matrix structure of images. One challenge of 2DCS is to design a 2D signal reconstruction algorithm. A fast algorithm called $2DS\ell_0$ (modified from smoothed ℓ_0) was proposed in [36] and [37], it achieved signal reconstruction in the matrix domain, both benefits of computation and storage were obtained. However, the quality of the reconstructed images was poor by directly using $2DS\ell_0$ to natural images. A novel 2D compressive sensing reconstruction algorithm was proposed in [38]. However, the reconstruction of this method must use ℓ_1 optimization row by row, so the computational complexity was very high. Besides, it should satisfy that one of the random projection matrix must be an invertible matrix. A novel recovery algorithm called 2DOMP (2D orthogonal matching pursuit), which was an extension of 1DOMP (1D orthogonal matching pursuit), was proposed in [39]. In this algorithm, each atom in the dictionary is a matrix. At each iteration, the decoder projected the sample matrix onto 2D atoms to select the best matched atom, and then renewed the weights for all the already selected atoms via the least squares. However, as each atom in the dictionary was a matrix, this algorithm needed a large memory to store the dictionary.

In this paper, we consider the problem of 2D CS sparse image reconstruction. A gradient based algorithm for 2D sparse image reconstruction called 2DPG (two-dimensional projected gradient) is proposed. In the proposed algorithm, we use a new model for natural images sparse representation in which the dual-tree discrete wavelet transform (DDWT) [40] ℓ_0-norm and the approximate total variation (TV) norm are applied simultaneously. In order to directly recover the original 2D image data from the 2D measurements, the sparse solution is searched iteratively in the 2D solution space and then updated by gradient descent of the TV and bivariate shrinkage in the DDWT domain. The main advantages of our proposed algorithm include: (1) images could not be divided into small blocks or stacked into a large vector, so the computational complexity and memory requirements are lower and (2) compared with the state-of-the-art CS based image reconstruction algorithm, the reconstructed images