Compressive Sensing for Noisy Video Reconstruction

Huihuang Zhao1,2,a, John Montalbo2,b, Shuxia Li2,c, Yaqi Sun1,d, Zhijun Qiao2,e

1Department of Computer Science, Hengyang Normal University, Hengyang, Hunan 421008, China
2Department of Mathematics, University of Texas-Pan American, Edinburg, TX 78539, USA
abetterlife008@163.com, bjmontalbo@broncs.utpa.edu, csl1@utpa.edu,
 dsunyaqiuyu@163.com, eqiao@utpa.edu

Abstract: In this paper, a compressing and reconstruction method for a noise video based on Compressed Sensing (CS) theory is proposed. At first, the CS theory is presented. Then the noise video is estimated from noisy measurement by solving the convex minimization problem. The video recovery algorithms based on gradient-based method is used to compressing and reconstructing the noise signal. And a compressive sensing algorithm with gradient-based method is proposed. At last, the performance of the proposed approach is shown and compared with some conventional algorithms. Our method can obtain best results in terms of peak signal noise ratio (PSNR) than those achieved by common methods with only a little runtime.

Index Terms—Noisy Video, Compressed Sensing, signal recovery algorithms, Restricted Isometry Property, gradient-based

1. INTRODUCTION

The well-known Nyquist/Shannon sampling theorem that the sampling rate must be at least twice the maximum frequency of the signal is a golden rule used in visual and audio electronics, medical imaging devices and so on. Compressed Sensing (CS) is a sampling paradigm that provides the signal compression at a rate significantly below the Nyquist/Shannon rate. Based on the CS theory, a sparse or compressible signal can be represented by the fewer number of bases than the one required by Nyquist/Shannon theorem, when it is mapped to the space with bases incoherent to the sparse data space [1, 2]. The major algorithmic challenge in compressive sampling is to approximate a signal given a vector of samples. The literature describes a huge number of approaches to solving this problem. Such as Orthogonal Matching Pursuit (OMP) [3], Greedy Basis Pursuit (GBP)[4], Iteratively Reweighted Least Square (IRLS)[5], CoSaMP[6], Suspace Pursuit(SP)[7], and so on.

The contents of most references are about imagery and raw data compressing and reconstruction based on CS theory. CS has been successfully applied to MRI [8], with consistent benefits in a clinical setting [9]. In [10], an iterative image reconstruction method in X-ray CT is proposed based on compressive sensing (CS). Reference [11] proposes a new method of fast encoding for Synthetic Aperture Radar (SAR) raw data by using the CS theory to complete SAR raw data compressing and reconstruction. But, there are so many problems in compressive sampling to approximate a noise signal [12, 13]. And noise video compressive sampling is also a challenge.
research problem. In this paper, a compressive sensing method for noisy video reconstruction based on gradient-based are studied.

This paper is organized as follows. In Section 2, the CS theory is presented. The processing of noisy video compressive sensing is introduced in Section 3. In Section 4, Noisy video reconstruction base on compressive sensing with gradient-based method is presented, and fast gradient-based compressed sensing reconstruction algorithm (FGB-CS) recovery algorithm are presented in detail. Experimental results are obtained with the different methods in Section 5. Finally, the conclusions are summarized at the end of this paper.

2. Compressed Sensing theory

CS is based on the assumption of the sparse property of signal and incoherency between the bases of sparse domain and the bases of measurement vectors. CS has three major steps: the construction of k-sparse representation, the compression, and the reconstruction. The first step is the construction of k-sparse representation, where k is the number of the non-zero entries of sparse signal. Most natural signal can be made sparse by applying orthogonal transforms such as Wavelet Transform, Fast Fourier Transform, and Discrete Cosine Transform. This step is represented as [2].

\[ s = \Psi^T x \]  

where \( x \) is an N-dimensional non-sparse signal; \( s \) is a weighted N-dimensional vector (sparse signal with \( k \) nonzero elements), and \( \Psi \) is an \( N \times N \) orthogonal basis matrix. The second step is compression. In this step, the random measurement matrix is applied to the sparse signal according to the following equation

\[ y = \Phi s = \Phi \Psi^T x \]  

where \( \Phi \) is an \( M \times N \) random measurement matrix (\( M < N \)).

Let \( M \) be the number of measurements (the row dimension of \( y \)) sufficient for high probability of successful reconstruction, and \( M \) is determined by

\[ M \geq C \mu^2 (\Phi, \Psi)k \log N. \]  

For some positive constant \( C \), \( \mu(\Phi, \Psi) \) is the coherence between \( \Phi \) and \( \Psi \), and defined by

\[ \mu(\Phi, \Psi) = \sqrt{N} \max_{i,j} |\langle \phi_i, \psi_j \rangle| \]  

If the elements in \( \phi \) and \( \psi \) are correlated, the coherence is large. Otherwise, it is small. From linear algebra, it is known that \( \mu(\Phi, \Psi) \in [1, \sqrt{N}] \). In the measurement process, the noise may occur. The noise is added into the compressed measurement vector as follows

\[ y = \Phi s + \text{noise} \]  

where noise is an \( M \)-dimensional noise vector.
As expected, signal $x$ in (5) may be estimated from noisy measurement $y$ by solving the convex minimization problem, called second-order cone program (SOCP), as follows.

$$\begin{align*}
\text{minimize} & \quad \|x\|_1 \\
\text{subject to} & \quad \|\Phi \Psi^T x - y\|_2 \leq \varepsilon
\end{align*}$$

(6)

where $\varepsilon$ is a bound of the amount of noise in the data.

3 The processing of noisy video compressive sensing

The framework of our approach is shown in Figure 1. For a video, it mainly contains the following steps:

(i) Read each frame, make it sparse by orthogonal transformation, consider its compression and recovery as a convex optimization problem, and then apply for the gradient-based method. Here, we assume the problem to be convex with the Lipschitz gradient.

(ii) Aiming at improving the efficiency, we replace an iteration parameter by the Lipschitz constant and a fast compressive sensing algorithm with gradient-based method is proposed.

(iii) Reconstruct each frame based on compressive compressing.

The processing of noisy video compressive sensing approach is shown in Figure 1.

Fig 1 The processing of noisy video compressive sensing

4 Noisy video reconstruction compressive sensing

4.1 Convex optimization problems based on Gradient-based methods

At first, we consider the unconstrained minimization problem of a convex function $g(x)$

$$\min \{ g(x) : x \in \mathbb{R}^n \}. \quad (7)$$
One of the simplest methods for solving eq. (8) is the gradient algorithm which generates a sequence $x_k$ via

$$x_0 \in \mathbb{R}^n, x_k = x_{k-1} - t_k \nabla g(x_{k-1}), \quad (8)$$

where $t_k > 0$ is a suitable step size. It is very well known that the gradient iteration Eq. 8 can be viewed as a proximal regularization [13] of the linearized function $g$ at $x_{k-1}$, and written equivalently as

$$x_k = \arg \min_x \{ g(x_{k-1}) + \langle x - x_{k-1}, \nabla g(x_{k-1}) \rangle + \frac{1}{2t_k} \| x - x_{k-1} \|_2^2 \}. \quad (9)$$

Adopting this same basic gradient idea to the nonsmooth $l_1$ regularized problem

$$\min \{ g(x) + \lambda \| x \|_1; x \in \mathbb{R}^n \} \quad (10)$$

leads to the iterative scheme [14]

$$x_k = \arg \min_x \{ g(x_{k-1}) + \langle x - x_{k-1}, \nabla g(x_{k-1}) \rangle + \frac{1}{2t_k} \| x - x_{k-1} \|_2^2 + \lambda \| x \|_1 \}. \quad (11)$$

After ignoring constant terms, this can be rewritten as

$$x_k = \arg \min_x \{ \frac{1}{2t_k} \| x - (x_{k-1} - t_k \nabla g(x_{k-1})) \|_2^2 + \lambda \| x \|_1 \}. \quad (12)$$

### 4.2 Noise signal optimization with Lipschitz gradient

The noise is added into the compressed measurement vector as follows

$$y = \Phi s + \text{noise} \quad (13)$$

where $\text{noise}$ is an $M$-dimensional noise vector. We think about an objective function $F(x) = g(x) + n(x)$, which is a composite type convex function, the Lipschitz gradient is named [15]

$$\| \nabla g(x) - \nabla g(y) \|_2 \leq L(g) \| x - y \|_2 \text{ for every } x, y, \quad (14)$$

where $L(g) > 0$ is a (Lipschitz) constant. where $\| \cdot \|$ denotes the standard Euclidean norm and $L(g) > 0$ is the Lipschitz constant of $\nabla g$. And we approximate the function $F(x)$ at point $x_{k-1}$ by quadratic function

$$Q_k(x, x_{k-1}) = g(x) + \langle x - x_{k-1}, \nabla g(x_{k-1}) \rangle + \frac{L}{2} \| x - x_{k-1} \|_2^2 + n(x) \quad (15)$$
which admits a unique minimizer

\[ PL(x_{k-1}) = \arg \min_x \{ Q(x, x_{k-1}), x \in R^n \}. \]  

(16)

Simple algebra shows that (ignoring constant terms in \( x_{k-1} \))

\[ PL(x_{k-1}) = \arg \min_x \left\{ \frac{L}{2} \| x - (x_{k-1} - \frac{1}{L} \nabla g(x_{k-1})) \|_2^2 + n(x) \right\}. \]  

(17)

Clearly, the basic step in Eq.9 is replaced by

\[ x_k = PL(x_{k-1}) \]  

(18)

with \( L \) set to \( 1/t_k \). Note that as long as the constant \( L \) in Eq.14 is taken to be no less than Lipschitz Constant \( L(g) \), it follows that [16]

\[ g(x) + n(x) \leq g(x_{k-1}) + \langle \nabla g(x_{k-1}), x - x_{k-1} \rangle + \frac{L}{2} \| x - x_{k-1} \|_2^2 + n(x). \]  

(19)

During our work, we replaced \( 1/t_k \) by a constant \( L \) which will be related to the Lipschitz constant \( L(g) \). We can find that the right-hand side of Eq.19 is precisely equal to \( Q_L(x, y) \) in Eq.15. In other words, \( Q_L(x, y) \) is an easier-to-deal-with convex upper bound of the objective function \( F(x) \) and by minimizing the upper bound, \( Q_L(x, y) \) with \( x_k \) given by Eq.18 offers a tight upper bound of \( F(x) \), provided that \( L \geq L(f) \), and

\[ F(x_k) - F(x) \leq \frac{2 L(f) \| x_k - x \|_2^2}{(k+1)^2}. \]  

(20)

So the convergence is \( O(1/k^2) \).

4.3 Compressed sensing for noisy video reconstruction

The major algorithmic challenge in compressive sampling is to approximate a noise signal given a vector of samples. The literature describes few numbers of approaches to solving this problem. In our method, Eq.(6) is often found more natural to study the closely related problem

\[ \min \| \Phi^T x - y \|_2^2 + \lambda \| x \|_1. \]  

(21)

We begin with considering the problem of Eq.21 assumed to be smooth and convex with smooth Lipschitz gradient. For any \( L > 0 \), compressed sensing of noisy signals formulated by Eq.21 becomes,
\[ x_k = \arg \min_x \left\{ \frac{L}{2} \| x - x_{k-1} \|^2 + \lambda \| x \|_1 \right\} \]  

(22)

where \( x_k = PL(x_{k-1}) \), so

\[ x_k = \arg \min_x \left\{ \frac{L}{2} \| x - (x_{k-1} - \frac{1}{L} \nabla f(x_{k-1})) \|^2 + \lambda \| x \|_1 \right\} \]  

(23)

or equivalently

\[ x_k = \arg \min_x \left\{ \frac{L}{2} \| x - d_k \|^2 + \lambda \| x \|_1 \right\} \]  

(24)

where \( d_k = x_{k-1} - \frac{1}{L} \nabla f(x_{k-1}) \). According to Eq.22, this \( d_k \) can be rewritten as

\[ d_k = x_{k-1} - \frac{1}{L} (\Phi \Psi^T)^T (\Phi \Psi^T x_{k-1} - y) \]  

(25)

Because both the 1-norm and square of the 2-norm are separable, i.e. each of them is mere the sum of \( n \) nonnegative terms and each of these terms involves only a single (scalar) variable, the iterate \( x_k \) in Eq.24 can be computed exactly by a straightforward shrinkage step (assuming \( d_k \) in Eq.24 has been calculated) as

\[ x_k = \Gamma_{\lambda L}(d_k) \]  

(26)

where \( \Gamma_{\lambda} \) is an shrinkage operator which maps \( R^n \) to \( R^n \) with the \( i \)-th entry of the output vector given by

\[ \Gamma_{\lambda}(d) = (|d_i| - \alpha)_+ \text{ sgn}(d_i) \]  

(27)

where \( (u)_+ = \max(u, 0) \).

4.3 FGBCS algorithm

In our study, a gradient-based compressive sensing algorithm (FGBCS) for noisy video reconstruction method is proposed, in which FGBCS algorithm is used to noise signal recovery, the detail algorithm is shown as follows.

\[ \text{FGBCS}(\Phi, \Psi, s, \lambda, K) \]

Input:

(i) \( L = L(g) - \alpha \) Lipschitz cons tant of \( \nabla g(x) \) in Eq.14.;

(ii) a video and its length is \( P \). Each frame is \( s \). \( \Psi \in R^{N \times N} \) is a signal sparse transform matrix,

(iii) a measurement matrix \( \Phi \in R^{M \times N} \), \( x = \Phi s \); the iteration counter \( K \) and noise parameter \( \lambda \).

Proc. of SPIE Vol. 9484  94840C-6
Procedure.
For frame=1 to \( P \)

Initialize, \( y_1 = x_0 \in R^n, t_1 = 1, s_p = \Psi s \);

If \( 1 \leq k \leq K \) Compute

(i) \( x_k = PL(y_k) \) by solving the problem in Eq. 18.

(ii) \( t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2} \) \( y_{k+1} = x_k + \frac{t_k - 1}{t_{k+1}}(x_k - x_{k-1}) \)

End

Output:
A sparse approximation \( x_k \) of the each frame \( s \) then reconstruction result signal \( s' = \Psi^T x_k \)

End

5 Experiment Result

In order to evaluate the quality of the reconstructed results, the mean square error (MSE) and peak signal noise ratio (PSNR) can be utilized. They are defined as [17]

\[
MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (\hat{f}(i,j) - f(i,j))^2
\]

\[
PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right)
\]

Where \( M \) and \( N \) are the image dimensions, \( \hat{f} \) is the denoised image, and \( f \) is the original noiseless image. During formula (28), 255 means the pixel values is 0 to 255 in a optical gray image. Many researchers used PSNR to estimate the result in image processing. In our study, the PSNR is used to compare the experiment result.

The video gsaesmang15.avi (total 48 frames) [18] was used as a test data in our experiment. CoSaMP, GBP, IRLS and OMP are chosen as the comparison methods. The video was degraded by rand noise (\( \sigma = 15 \)). The 1th frame and its noisy image are shown as in figure 2(a) and figure 2(b). The reconstruction results based on different methods with the matrix’s rows \( M = 230 \) can be shown as figure 2(c-g). During FGbCS, we set \( \lambda = 20 \) and \( K = 40 \), and the result can be shown as figure 2(h).
The 10th frame and its noisy image are shown as in figure 3(a) and figure 3(b). The reconstruction results based on different methods with the matrix’s rows M=230 can be shown as figure 3(c-g). During FGbCS, we set \( \lambda = 20 \) and \( K = 40 \), and the result can be shown as figure 3(h).

The 20th frame and its noisy image are shown as in figure 3(a) and figure 3(b). The reconstruction results based on different methods with the matrix’s rows M=230 can be shown as figure 3(c-g). During FGbCS, we set \( \lambda = 20 \) and \( K = 40 \), and the result can be shown as figure 3(h).
In order to compare their reconstruction performance in detail, the more experiments were also done. The reconstruction average time and PSNR with different rows of measurement matrix can be shown as figure 5(a,b).

We can see from figure 5 that,

(i) The noisy video reconstruction accuracy decrease with the increase of measurement matrix rows. And among of those methods, FGB algorithm can obtain best result than those other methods, and the second is IRLS algorithm.

(ii) The noisy video reconstruction accuracy decrease with the increase of measurement matrix rows. The method based on OMP and FGB algorithm can run the fastest than those other methods in noise signal reconstruction, and SP, CoSaMP, GBP and IRLS is in the second, third and fourth. Among of them, the runtimes decrease with the increase of measurement matrix rows. IRLS method can get good reconstruction result, but it spends the most time.

We can see from figure 5 that, the proposed method can obtain best reconstruction result comparing with OMP ,SP, CoSaMP, GBP and IRLS.

6 Conclusion

There are many ices in compressive sampling to approximate a noise signal given a vector of samples. And there many algorithms, such as OMP, SP, CoSaMP, GBP and IRLS are proposed to solve
the problem. But those algorithms are not good for noisy video reconstruction. Among of them, IRLS achieves the good performance on low error, but it spends the most time. This paper presented here has focused on the formulation of the compressed sensing for noise signal reconstruction algorithm. The noisy and video is estimated as a convex minimization problem. And the gradient-based methods is used to solve the problem. In order to improve the run speed, the step size in gradient iteration is replaced by a constant $1/L$ which is related to the Lipschitz constant. We proposed a new compressive sensing for noisy video reconstruction methods base on gradient-based. The experiments have been shown that the proposed method can obtain best reconstruction result comparing with OMP ,SP, CoSaMP, GBP and IRLS.

References


