Improved Sparse Signal Reconstruction based on Approximate Hyperbolic Tangent Function with Smoothed ℓ0 Norm

Yuanwei Zhang [a], Jianyong Yu, Haohao Bai, Jian Kang, Xiaoguang Gong,

Guangbao Hu, Ping Luo

School of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

[a]zhangyuanwei6878@163.com

Abstract

In this paper, reconstruction algorithm for compressive sensing is based on smoothed ℓ0 norm. The smoothed function sequence is introduced into the SL0 algorithm to approximate the ℓ0 norm, and the minimization problem of ℓ0 norm is transformed into the optimization problem of smoothed function.This paper has proposed a reconstruction algorithm with a modified Newton direction (ONSL0 algorithm) which is based on an optimized projection direction in ℓ0 norm.The numerical simulation results show that the ONSL0 algorithm is better than both the SL0 algorithm and the OSL0 algorithm at the peak signal-to-noise ratio and relative error.

Keywords

Compressive sensing, Approximate hyperbolic tangent function OSL0 algorithm ONSL0 algorithm.

1. Introduction

Over the past decades, the sparse representation of signals have been in-depth study. The concept of signal sparsity and ℓ1 norm based on recovery techniques can be traced back to the work of Logan in 1965 [1], the work of Santosa and Symes in 1986 [2], Donoho and Stark's work in 1989 [3]. It is generally accepted that the cornerstone of the compressive sensing theory (CST) is based on Candès et al. [4], Donoho [5], and Candès and Tao [6] in 2006. CST can achieve simultaneous sampling and compression of the signal, and even dissatisfying the Nyquist sampling theorem, it is possible to achieve a better approximation of the original complete signal with only a few sampled data. The central problem in CST is how to reconstruct sparse signals with finite measured values. The essence of the above problem is to find the most sparse solution of the undetermined system of linear equations $y = Dx$, where,given the signal $y \in \mathbb{R}^m$, the dictionary $D \in \mathbb{R}^{m \times n}$ ($\mathbb{R}^{m \times n}$ is a collection of n atoms with $n > m$), the goal is to represent y as a linear combination of the atoms of D in a parsimonious way. To this end, the following problem has to be solved

$$\min_x \|x\|_0 \ s.t. \ y = Dx$$

Where $\| \ \|_0$ is the so-called ℓ0 (pseudo) norm, defined as the number of non-zero entries.ℓ0 norm function is very discontinuous and non-differentiable, which makes the above problems difficult to solve [7] [8]. In order to solve the above-mentioned shortcomings, it is necessary to introduce an alternative sparse promoting function. The most famous estimating function is the ℓ1 norm function, which is the nearest convex norm relative to the ℓ0 norm function [7]. By using the ℓ1 norm function, the sparse signal recovery problem can be expressed as

$$\min_x \|x\|_1 \ s.t. \ y = Dx.$$

Various algorithms have been proposed to solve the above-mentioned sparse recovery problems; e.g. [9].

The algorithms include iterative re-weighted least squares (IRLS) [10], iterative hard thresholding (IHT) [11], and smooth ℓ0 (SL0) norm [12] [15] [16] [18], which have used the sparse promoting function (sparse than ℓ1 norm). In particular, SL0 estimates the non-smooth ℓ0 norm by a differentiable function with a smooth parameter (expressed as σ in [12]).

The rest of the paper is organized as follows. In Section II, the basic principle of this algorithm is briefly presented. Section III is about the experimental simulations and analytical results. The paper is finally concluded in Section IV.

2. The basic idea of the algorithm in this paper

2.1 Multi-parameter smoothed ℓ0 norm algorithm based on approximate hyperbolic tangent function

The SL0 algorithm needs to select the appropriate smooth continuous function to approximate the ℓ0 norm and solve the minimum solution of the continuous function to obtain the minimum solution of the ℓ0 norm. Now, using the Gaussian function to approximate the ℓ0 norm is a major method. In this paper, we will use the following approximate hyperbolic tangent function (1) to approximate the ℓ0 norm, which can further improve approximation performance of the ℓ0 norm. Among (1), σ ,as a parameter, determines the estimated performance of the function.

$$f_\sigma(s_i)=\frac{\exp(\frac{s_i^{\,2}}{2\sigma^2})-\exp(\frac{-s_i^{\,2}}{2\sigma^2})}{\exp(\frac{s_i^{\,2}}{2\sigma^2})+\exp(\frac{-s_i^{\,2}}{2\sigma^2})} \tag{1}$$

In this paper, the approximate effect between the approximate hyperbolic tangent function and the Gaussian function (2) used in [14] is shown in Fig.1. It can be seen from Fig. 1 that the approximate hyperbolic tangent function is more cliffy than the Gaussian function, so the approximate effect of the Gaussian function on the ℓ0 norm is less than that of the approximate hyperbolic tangent function.

$$\varphi(s_i)=\exp(\frac{-s_i^{\,2}}{2\sigma^2}) \tag{2}$$

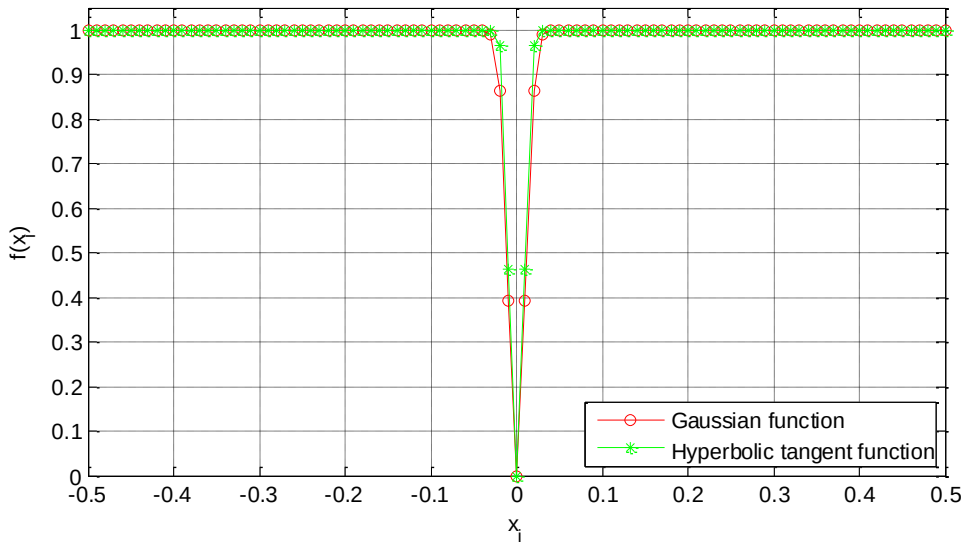

Figure 1. Comparison about the approximation effect between the approximate hyperbolic tangent function and the Gaussian function with sigma = 0.01

For an arbitrary σ the above conclusion is true. Its proof is as follows. Make $u(s)=f_\sigma(s)-(1-\varphi(s))$ .We know $u(s)=0$ from the definition of $u(s)$, when there are $s=0$ or $s=\pm\infty$. $u(s)$ is discussed below $0<|s|<\infty$.

$$u(s)=\frac{\exp(\frac{s^2}{2\sigma^2})-\exp(\frac{-s^2}{2\sigma^2})}{\exp(\frac{s^2}{2\sigma^2})+\exp(\frac{-s^2}{2\sigma^2})}-1+\exp(\frac{-s^2}{2\sigma^2})=\frac{\exp(\frac{s^2}{\sigma^2})-1}{\exp(\frac{s^2}{\sigma^2})+1}-1+\exp(\frac{-s^2}{2\sigma^2})=$$

$$\frac{-2}{\exp(\frac{s^2}{\sigma^2})+1}+\exp(\frac{-s^2}{2\sigma^2})=\frac{1}{\exp(\frac{s^2}{\sigma^2})+1}(-2+\exp(\frac{s^2}{2\sigma^2})+\exp(\frac{-s^2}{2\sigma^2}))=$$

$$\frac{1}{\exp(\frac{s^2}{\sigma^2})+1}(\exp(\frac{s^2}{4\sigma^2})+\exp(\frac{-s^2}{4\sigma^2}))^2>0$$

Thus, there is $0\le u(s)$ and $f_\sigma(s)$ has a better convergence than $\varphi(s)$ when they approximate $\ell 0$ norm. From (1) and (2), it can be clearly seen that the smaller the σ is, the more the local extreme values of objective function would be, and it would increase the difficulty of obtaining the global optimal value of the objective function. The σ determines the smoothness of the objective function.

The greater the σ is, the smoother the objective function, and it will have less accuracy to estimate the $\ell 0$ norm. On the contrary, the smaller the σ is, the more excellent the approximate accuracy of the $\ell 0$ norm would be.

It is possible to construct a set of attenuated sequences $\{\sigma_1, \sigma_2, ? \cdots, \sigma_j\}$ to optimize each corresponding objective function until σ_j is sufficiently small. This will eliminate the effect of local extremes and make it possible to obtain the global optimal value of the smoothed function. Figure 2 and Figure 3 respectively show the comparison of the approximation effect between the Gaussian function and the approximate hyperbolic tangent function with σ=0.01,0.06,0.11,0.16,0.21. It can be seen from Fig. 2 and Fig. 3 that under σ=0.01the Gaussian function and the approximate hyperbolic tangent function are all the steepest in the five sigma values.

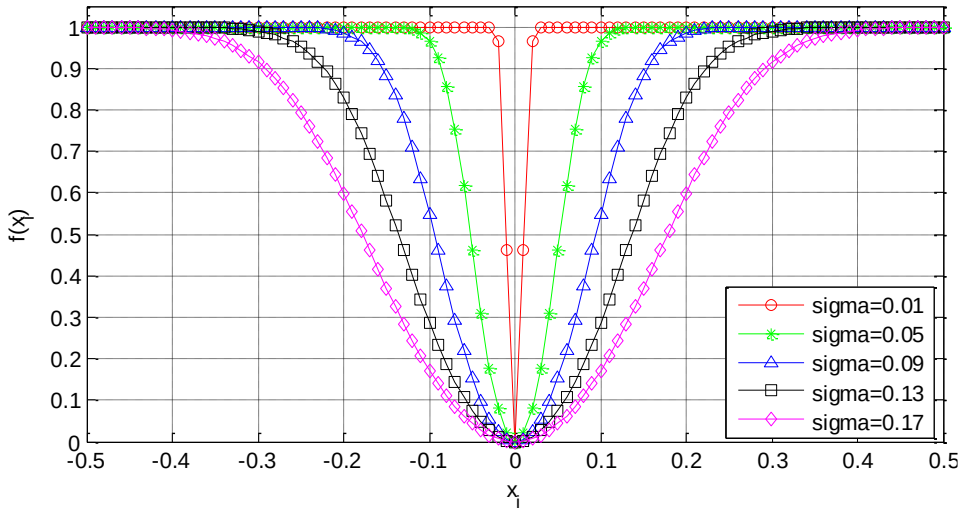

Figure 2. Comparison of the approximate hyperbolic tangent function at different sigma values

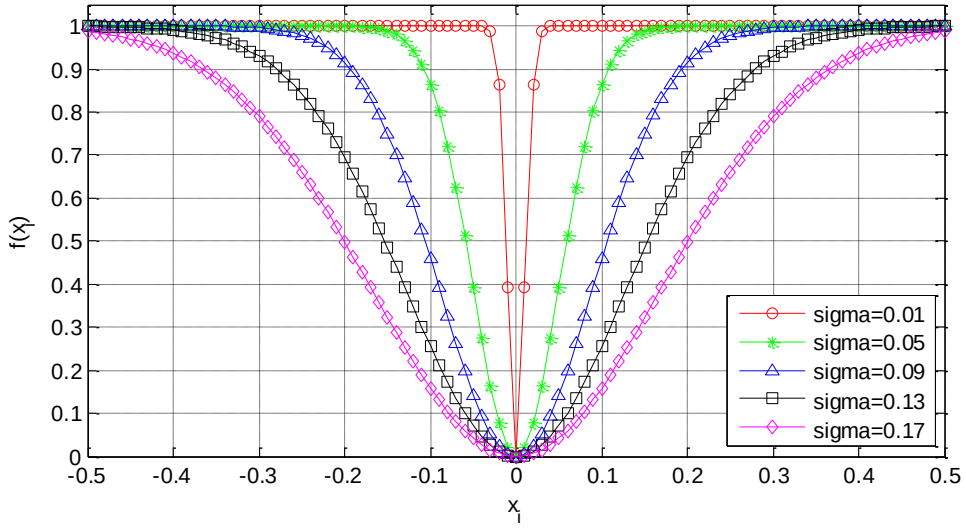

Figure 3. Comparison of the Gaussian function at different sigma values

The result obtained by finding the limit of the approximate hyperbolic tangent function is shown in the following equation (3).

$$\lim_{\sigma\to 0} f_\sigma(s_i) = \begin{cases} 0, & s_i = 0 \\ 1, & s_i \neq 0 \end{cases} \tag{3}$$

Therefore, if there is a formula (4), then the limit value of $f_\sigma(s_i)$ at $\sigma \to 0$ is equal to the number of elements that are not zero and in the vector $\mathbf{s}$.This is consistent with the definition of $\ell 0$ norm,so we can gain (5) .

$$F_\sigma(\mathbf{s}) = \sum_{i=1}^{N} f_\sigma(s_i) \tag{4}$$

$$\|\mathbf{s}\|_0 = \lim_{\sigma\to 0} F_\sigma(\mathbf{s}) \tag{5}$$

Using the Newton direction instead of the steepest descent direction can avoid the "jagged" phenomenon, which can speed up the convergence rate of the SL0 algorithm. The formula for correcting the Newton direction is shown in (6).

$$d = -\nabla^2 F_\sigma(s)^{-1} \nabla F_\sigma(s) \tag{6}$$

Substituting equations (1) and (4) into equation (6) can yield equation (7) [13].

$$\nabla F_\sigma(s) = \left[\frac{\partial f_\sigma(s_1)}{\partial s_1}, \cdots\cdots, \frac{\partial f_\sigma(s_N)}{\partial s_N} \right]^T =$$

$$\left[\frac{\frac{4s_1}{\sigma^2}}{\left[\exp(\frac{s_1^2}{2\sigma^2}) + \exp(\frac{-s_1^2}{2\sigma^2}) \right]^2}, \cdots\cdots, \frac{\frac{4s_N}{\sigma^2}}{\left[\exp(\frac{s_N^2}{2\sigma^2}) + \exp(\frac{-s_N^2}{2\sigma^2}) \right]^2} \right]^T \tag{7}$$

Hessian matrix is shown in the following equation (8), and its eigenvalues may be non-positive and not necessarily positive resulting in that d is sometimes not descending. Therefore, the Hessian matrix needs to be corrected and then the corrected Newton direction needs to be calculated.

$$\nabla^2(F_\sigma(s))=\begin{bmatrix} \frac{\partial^2 f_\sigma(s_1)}{\partial s_1^2} & 0 & \cdots & 0 \\ 0 & \frac{\partial^2 f_\sigma(s_2)}{\partial s_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial^2 f_\sigma(s_N)}{\partial s_N^2} \end{bmatrix} \tag{8}$$

The following equation (9) [17]is obtained by performing the second-order partial derivative operation on (1).

$$\frac{\partial^2 f_\sigma(s_i)}{\partial s_i^{\,2}}=\frac{\frac{4}{\sigma^2}\left[(1+\frac{2s_i^{\,2}}{\sigma^2})\exp(\frac{-s_i^{\,2}}{2\sigma^2})+(1-\frac{2s_i^{\,2}}{\sigma^2})\exp(\frac{s_i^{\,2}}{2\sigma^2})\right]}{\left[\exp(\frac{s_i^{\,2}}{2\sigma^2})+\exp(\frac{-s_i^{\,2}}{2\sigma^2})\right]^3} \tag{9}$$

Now we build the matrix G on the basis of the original Hessian matrix to make G a positive definite matrix. The corrected method is $G=\nabla^2 F_\sigma(s)+\varepsilon_k I$,where ε_k is a group of proper positive numbers, I is an identity matrix, and the diagonal elements in matrix G are all positive. When G's eigenvalues are all positive, G is positive definite matrix. If you can find out the right ε_k which can guarantee that G's diagonal elements are all positive, then G is a positive definite matrix.

After looking for many ε_k ,we gain

$$\varepsilon_k=\frac{\frac{4}{\sigma^2}\left[\frac{3s_i^{\,2}}{\sigma^2}\exp(\frac{-s_i^{\,2}}{2\sigma^2})-\frac{s_i^{\,2}}{\sigma^2}\exp(\frac{s_i^{\,2}}{2\sigma^2})\right]}{\left[\exp(\frac{s_i^{\,2}}{2\sigma^2})+\exp(\frac{-s_i^{\,2}}{2\sigma^2})\right]^3}.$$

At this moment there is

$$\frac{\partial^2 f_\sigma(s_i)}{\partial s_i^{\,2}}=\frac{4(1+\frac{s_i^{\,2}}{\sigma^2})}{\sigma^2\left[\exp(\frac{s_i^{\,2}}{2\sigma^2})+\exp(\frac{-s_i^{\,2}}{2\sigma^2})\right]^2}>0.$$

The last Newton direction has become

$$d=\left[-\frac{\sigma^2 s_1}{\sigma^2+s_1^{\,2}},\cdots\cdots,-\frac{\sigma^2 s_N}{\sigma^2+s_N^{\,2}}\right]^{\mathrm{T}}.$$

2.2 Multi-parameter Smoothed ℓ0 norm algorithm based on the approximate hyperbolic tangent function

Unoptimized algorithm (SL0) is as follows.

(1) Initialize every parameter of the SL0;

①Let $\mathbf{s}=\mathrm{A}^{\mathrm{T}}\left(\mathrm{AA}^{\mathrm{T}}\right)^{-1}\mathbf{y}$.

②Select an appropriate descending sequence for σ , i.e. $\{\sigma_1, \sigma_2, ?\ \cdots, \sigma_j\}$, $\sigma_j=\beta\sigma_{j-1}$.

(2) External loop: j=1，2，3，……，J.

①Let $\sigma=\sigma_j$.

② Let $s=\hat{s}_{j-1}$.

③ Let r_0=0.

④ Internal loop:

(a) Calculate the corrected Newton direction $d=\left[-\frac{\sigma^2 s_1}{\sigma^2+s_1^2},\cdots\cdots,-\frac{\sigma^2 s_N}{\sigma^2+s_N^2}\right]^{\mathrm{T}}$.

(b) Update the reconstructed signal $\mathbf{s}\leftarrow\mathbf{s}$+d.

(c) According to the gradient projection principle, gain $\mathbf{s}\leftarrow\mathbf{s}-\mathrm{A}^{\mathrm{T}}\left(\mathrm{AA}^{\mathrm{T}}\right)^{-1}\left(\mathrm{A}\mathbf{s}-\mathbf{y}\right)$.

(d) Calculate the margin r=$\mathbf{y}-\mathrm{A}\mathbf{s}$.

(e) If there is $\|r-r_0\|_2<e$, the internal loop with current σ is ended; otherwise, r_0=r.

⑤ $\hat{s}_j$=s.

(3) Obtain a reconstructed signal $\hat{s}=\hat{s}_j$.

The iterative Newton SL0 algorithm described above sometimes iterates many times before satisfying $\|r-r_0\|_2<e$. This will cause an unnecessary extension on the computation time, and even be in an infinite loop. We can set a maximum iterative number about an internal loop(L_{max}) with reference to the optimized SL0 algorithm(OSL0). We cannot set the parameters of ONSL0 with reference to OSL0 under affecting the reconstructed precision.

The optimized iteration Newton SL0 algorithm (referred to as the ONSL0) is as follows.

(1) Initialize every parameter of the OSL0;

①Let $\mathbf{s}=\mathrm{A}^{\mathrm{T}}\left(\mathrm{AA}^{\mathrm{T}}\right)^{-1}\mathbf{y}$.

②Select an appropriate descending sequence for σ, i.e. $\{\sigma_1,\sigma_2,?\cdots,\sigma_j\}$, $\sigma_j=\beta\sigma_{j-1}$.

(2) External loop: j=1，2，3，……，J.

①Let $\sigma=\sigma_j$.

② Let s=$\hat{s}_{j-1}$.

③ Let r_0=0.

④ Internal loop: L=1，……，L_{max}.

(a) Calculate the corrected Newton direction $d=\left[-\frac{\sigma^2 s_1}{\sigma^2+s_1^2},\cdots\cdots,-\frac{\sigma^2 s_N}{\sigma^2+s_N^2}\right]^{\mathrm{T}}$.

(b) Update the reconstructed signal $\mathbf{s}\leftarrow\mathbf{s}$+d.

(c) According to the gradient projection principle, gain $\mathbf{s}\leftarrow\mathbf{s}-\mathrm{A}^{\mathrm{T}}\left(\mathrm{AA}^{\mathrm{T}}\right)^{-1}\left(\mathrm{A}\mathbf{s}-\mathbf{y}\right)$.

(d) Calculate the margin r=$\mathbf{y}-\mathrm{A}\mathbf{s}$.

(e) If there is $\|r-r_0\|_2<e$, the internal loop with current σ is ended; otherwise, r_0=r.

⑤ $\hat{s}_j$=s.

(3) Obtain a reconstructed signal $\hat{s}=\hat{s}_j$.

3. Experimental Simulation and Results Analysis

3.1 Experimental Environment

This part will give performance analysis and comparison between the ONSL0 and the main reconstruction algorithm: SL0, OSL0. In order to fairly compare the reconstruction performance about different algorithms, all experiments are performed under Windows 10 and MATLAB V8.0 (R2012a) running on an ASUS notebook with an Intel (R) Core(TM), CPU i3 at 2.40GHz, and 4GB of memory.

In this experiment, the object of the reconstruction algorithms are three 512×512 gray-scale images which are named after Debbie, Peppers and Barbara respectively. Debbie contains a lot of flat areas, in which edge areas and texture areas are relatively simple.Peppers contains a lot of edge areas and flat areas, and its texture details are less than Barbara's. Barbara also contains a large number of flat areas, but its edge areas and texture areas are richer than Debbie's.

Measurement matrix Φ is a Gaussian random matrix, which satisfies the restricted isometric properties (RIP) and is easy to generate. Since the Gaussian random matrices are random, this would lead to the fluctuation on the reconstructed image quality and the reconstructed time. Therefore the PSNR value, the relative error and the reconstruction time are all the average values after independently testing 5 times. The compression ratio (assuming its size is M * N) is a result which equals the number of rows of Φ (assuming its size is M * N)divided by the number of columns of Φ;i.e. M/N . M/N is denoted by the sampling rate because the lower the sampling rate is, the smaller the compression ratio is. For sparse base Ψ , using a wavelet transform matrix symlets8 is an optimal selection.

3.2 Results and Analysis

Table 1 tabulates the PSNR value about the three reconstruction algorithms mentioned above.

From Table 1, it is clear that average PSNR value about the three images under ONSL0 is 0.334dB higher than that under OSL0 and is 1.009 dB more than that under SL0.Especially, the psnr value for the Debbie image under ONSL0 reaches 40.695dB.

As can be seen from Table 1, for the all images, under the compression ratio 0.5 the reconstruction time of OSL0 is the shortest, and it is 4.464s on average; the reconstruction time of ONSL0 is 2.415s smaller than that of SL0 and is 3.090s higher than that of OSL0 on average.

It can also be seen from Table 1, for the all images, under the compression ratio 0.5 the reconstruction relative error of ONSL0 is the smallest. Especially for the Debbie image, the relative error of ONSL0 is only 0.0169, and this is the smallest oneof all the relative error values in Table 1.

Figure 4, Figure 5 and Figure 6 are all the contrast about reconstructed visual effects of SL0, OSL0 and ONSL0 under compression rate 0.5.

Table 1. Compression rate 0.5, the reconstruction performance of SL0, OSL0 and ONSL0 for Debbie, Peppers and Barbara

Image Name	Algorithm	PSNR/dB	Runtime/s	Relative Error
Barbara	SL0	29.077	11.079	0.0550
	ONSL0	30.262	9.701	0.0480
	OSL0	29.958	4.954	0.0497
Debbie	SL0	40.236	8.902	0.0180
	ONSL0	40.695	5.768	0.0169
	OSL0	40.523	3.847	0.0173
peppers	SL0	31.982	9.927	0.0417
	ONSL0	33.365	7.193	0.0353
	OSL0	32.840	4.590	0.0376

Figure 4. Barbara image, compression ratio 0.5, contrast about reconstructed visual effects of SL0, OSL0 and ONSL0

Figure 5. Debbie image, compression ratio 0.5,contrast about reconstructed visual effects of SL0, OSL0 and ONSL0

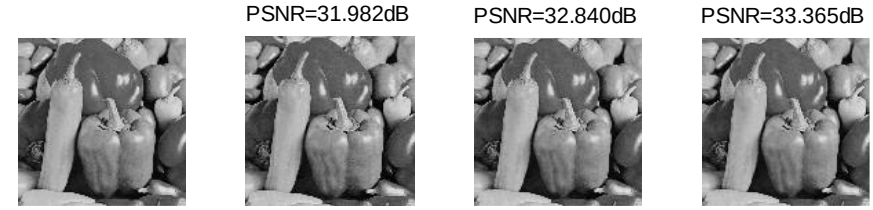

Figure 6. Peppers image, compression ratio 0.5, contrast about reconstructed visual effects of SL0, OSL0 and ONSL0

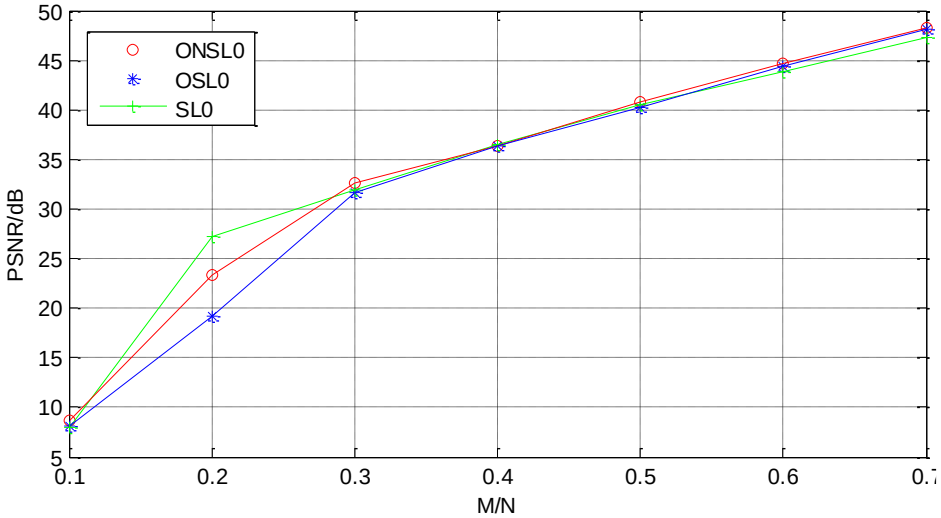

Figure 7. Debbie image, at different compression rates, psnr value comparison among SL0, OSL0 and ONSL0

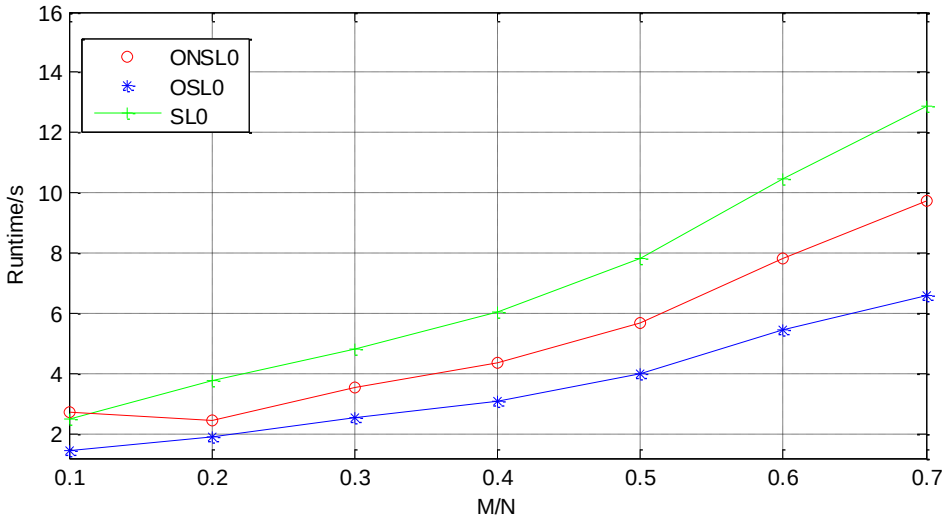

Figure 8. Debbie image, at different compression rates, reconstruction time(runtime) comparison among SL0, OSL0 and ONSL0

Under different compression ratios, Figure 7 and Figure 8 show the contrast curves about PSNR value and reconstructed time for the Debbie image with SL0, OSL0 and ONSL0. It can be seen from Fig. 7 that the PSNR value of ONSL0 is significantly higher than that of SL0 and OSL0 in the range of 0.1 to 0.3. Especially at compression ratio 0.2 difference value between SL0, OSL0 and ONSL0 is the maximum value. Difference value between SL0, OSL0 and ONSL0 is all very small where M/N varies in the range between 0.3 and 0.7.The PSNR values of these three algorithms all increase gradually along with addition in M/N, where M/N varies in the range between 0.1 and 0.7.

It can be seen from Fig. 8 that , when M/N ranges from 0.1 to 0.7,the reconstruction time (runtime) of these three algorithms all increase gradually along with addition in M/N. Besides, the difference value between OSL0 and ONSL0 increases gradually as M/N increases, and the difference value between SL0 and ONSL0 is same as that between the OSL0 and ONSL0. When there is a same M/N the runtime of SL0 is greater than that of the ONSL0 , and the runtime of ONSL0 is higher than the runtime of OSL0.

In summary, the reconstructed PSNR value of ONSL0 which is almost unchanged when M/N ranges from 0.3 to 0.7.

But the reconstruction time of ONSL0 is fairly longer than that of OSL0 and at the compression rate 0.7, it has almost reached a maximum 3s. This is equivalent to that the reconstruction time of ONSL0 is used in exchange for a small increase on the PSNR value of ONSL0.

However, it should be noted that as M/N increases, the increment of reconstruction time would increase gradually, while the increment of reconstructed PSNR value remains almost constant.

When M/N is relatively large, we should make a choice between the precision of reconstructed PSNR value and the reconstruction time.

Either we can select ONSL0 to further improve accuracy, or we can select OSL0 to shorten runtime while the reconstructed accuracy of PSNR is slightly lower.

4. Conclusion and Future Work

This paper based on SL0 and OSL0 has proposed ONSL0 with regard to the shortcomings on estimated function of $\ell 0$ norm and iterative termination condition. It uses a hyperbolic tangent function which is of better quality to approximate the $\ell 0$ norm and applies the modified Newton method to achieve the optimal value in solution. Through the previous numerical simulation experiments, we can see that the reconstructed quality of ONSL0 is superior to that of SL0 and OSL0

regarding objective data, but we can hardly see any difference in subjective visual effect. However, ONSL0 involves the Hessian matrix, that is, the inverse problem of $\nabla F_\sigma(s)$, which will waste a certain amount of runtime. If we can further find a sequence of functions which can better approximate the ℓ0 norm and replace $f_\sigma(s_i)$ to reduce the cost of runtime, then this new algorithm will become a reconstruction algorithm with a good performance. This will be one of the next steps that needs to be studied.

References

[1]B.F.Logan,Properties of high-pass signals [Ph.D.thesis],Columbia University,1965.

[2]F.SantosaandW.W.Symes,"Linear inversion of band-limited reflection seismograms,"Society for Industrial and Applied Mathematics,vol.7,no.4,pp.1307–1330,1986.

[3]D.L.DonohoandP.B.Stark,"Uncertainty principles and signal recovery,"SIAM Journal on Applied Mathematics,vol.49,no.3,pp.906–931,1989.

[4]E.J.Candès,J.Romberg,andT.Tao,"Robust uncertainty principles:exact signal reconstruction from highly incomplete frequency information,"IEEE Transactions on Information Theory, vol.52, no.2, pp. 489–509,2006.

[5]D.L.Donoho,"Compressed sensing,"IEEE Transactions on Information Theory, vol.52, no.4, pp. 1289–1306,2006.

[6]E.J.CandèsandT.Tao,"Near-optimal signal recovery from random projections:universal encoding strategies?"IEEE Transactions on Information Theory,vol.52,no.12,pp.5406–5425,2006.

[7] M. Elad, Sparse and Redundant Representations. New York, NY, USA:Springer, 2010.

[8]G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximations," Construct. Approx., vol. 13, no. 1, pp. 57–98, 1997.

[9]J.A.Tropp and S. J. Wright,"Computational methods for sparse solution of linear inverse problems,"Proc.IEEE,vol.98,no.6,pp.948–958,2010.

[10] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in Proc. IEEE ICASSP, 2008, pp. 3869–3872.

[11] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," Appl. Comput. Harmon. Anal., vol. 27, no. 3,pp. 265–274, 2009.

[12] H. Mohimani, M. Babaie-Zadeh, and Ch. Jutten, "A fast approach forovercomplete sparse decomposition based on smoothed ℓ0 norm," IEEETrans. Signal Process., vol. 57, 1, pp. 289–301, 2009.

[13]S.BoydandL.Vandenberghe,Convex Optimization,Cambridge University Press, 2004.

[14]H.Mohimani,M.Babaie-Zadeh,andC.Jutten,"A fast approach for overcomplete sparse decomposition based on smoothed ℓ0norm,"IEEE Transactions on Signal Processing, vol. 57, no. 1, pp. 289–301, 2009.

[15]Mostafa S.and Massoud B.Z.Iterative Sparsification-Projection: Fast and Robust Sparse Signal Approximation, IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 64, NO. 21, NOVEMBER 1, 2016.

[16] H. Mohimani, M. Babaie-Zadeh, and Ch. Jutten, "Fast sparse representation based on smoothed ℓ0 norm," in Proceedings of 7th InternationalConference on Independent Component Analysis and Signal Separation(ICA2007), Springer LNCS 4666, London, UK, September 2007, pp.389–396.

[17]Fang X F, Zhang J S, Li Y Q. Sparse Signal Reconstruction Based on Multiparameter Approximation Function with Smoothed Norm[J]. Mathematical Problems in Engineering, 2015, 2014(6):1-9.

[18]Hyder M M, Mahata K. An Improved Smoothed ℓ0 Approximation Algorithm for Sparse Representation[J]. IEEE Transactions on Signal Processing, 2010, 58(4):2194-2205.