Denoising via truncated sparse decomposition*

Xie Zong-Bo(谢宗伯) and Feng Jiu-Chao(冯久超)†

School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, *China*

(Received 26 March 2010; revised manuscript received 20 December 2010)

This paper proposes a denoising algorithm called truncated sparse decomposition (TSD) algorithm, which combines the advantage of the sparse decomposition with that of the minimum energy model truncation operation. Experimental results on two real chaotic signals show that the TSD algorithm outperforms the recently reported denoising algorithms—non-negative sparse coding and singular value decomposition based method.

Keywords: denoising, truncated sparse decomposition, sparse decomposition, chaotic signals
PACS: 05.40.Ca, 05.45.–a, 05.45.Tp **DOI:** 10.1088/1674-1056/20/5/050504

1. Introduction

In the past decade, the study of denoising algorithms has been an important issue in acoustics signal processing. Various denoising algorithms such as kernel principal component analysis,[1] Wavelet Shrinkage (WS),[2,3] Non-Negative Sparse Coding (NNSC),[4] and Singular Value Decomposition (SVD)[5,6] have been proposed and utilized in many applications. An important objective of them is to extract the ideal signal from the additive white Gaussian noise (AWGN) corrupted signal. The NNSC[4] and SVD[5,6] are the most recently reported of these algorithms. The idea behind NNSC is to represent the input data approximately by a weighted linear combination of a small number of basis vectors, which contain high-level patterns in the input data, while the idea behind SVD is that by selecting the proper singular values representing signal and discarding the ones representing noise, the additive noise of an signal can be eliminated effectively. However, it is difficult to make the performance of NNSC or SVD be improved further. To overcome this limit, we propose the Truncated Sparse Decomposition (TSD) algorithm in this paper. Its basic idea is applying Sparse Decomposition (SD) to project the input data onto two sparse spaces and then choose the proper diagonal values representing the signal and discard the ones representing the noise. Experimental results on two real chaotic signals show that the TSD is more efficient for AWGN suppressing than NNSC and SVD.

2. The TSD algorithm

The pioneering work on SVD can be traced to Ref. [7] in which a method for representing the observed data by using the product of three matrices is presented. The SVD and NNSC enlighten us to develop the TSD.

SD Let $\boldsymbol{A}$ be an $m \times n$ observed data corrupted by AWGN. The $\boldsymbol{X}$ is an $m \times k$ matrix ($m > k$) and $\boldsymbol{Y}$ is an $n \times k$ matrix, which both impose sparsity requirements and the $\boldsymbol{D}$ is a positive diagonal matrix. The linear representation of the observed data is then given as follows:

$$\boldsymbol{A} = \boldsymbol{X}\boldsymbol{D}\boldsymbol{Y}^{\mathrm{T}}. \tag{1}$$

Let $\{u, \sigma, v\}$ be the largest singular triple of $\boldsymbol{A}$ and ε the tolerance for reconstruction error. We define the vector $\boldsymbol{w} = \left[u^{\mathrm{T}}, v^{\mathrm{T}}\right]^{\mathrm{T}}$ by $\tilde{w}$ in non-increasing order

$$[\tilde{w}(1)] \geq [\tilde{w}(2)] \geq \cdots \geq [\tilde{w}(m+n)]. \tag{2}$$

This procedure helps us to determine two permutation matrices $\boldsymbol{P}_u$ and $\boldsymbol{P}_v$ being in non-increasing orders in

*Project supported by the National Natural Science Foundation of China (Grant No. 60872123), the Joint Fund of the National Natural Science Foundation and the Guangdong Provincial Natural Science Foundation (Grant No. U0835001), the Doctorate Foundation of South China University of Technology, the Post-Doc Foundation of South China University of Technology, the Basic Scientific Research Fund of South China University of Technology for Youth, the Natural Science Fund of South China University of Technology for Youth, the Natural Science Foundation of Guangdong Province, China, and the China Postdoctoral Science Foundation (Grant No. 20100480049).

†Corresponding author. E-mail: fengjc@scut.edu.cn

© 2011 Chinese Physical Society and IOP Publishing Ltd

absolute values, respectively. For a given $\xi > 0$, let $k = k_\xi \leq m + n$ be the smallest integer satisfying

$$\tilde{w}^2(1) + \tilde{w}^2(2) + \cdots + \tilde{w}(k) \geq 2(1-\xi) \tag{3}$$

and let $i = i_\xi$ and $j = j_\xi = k_\xi - i_\xi$ be the numbers of u-components and v-components in the subset $\{\tilde{w}(1), \ldots, \tilde{w}(k)\}$ respectively. Then $\boldsymbol{x}_u$ and $\boldsymbol{y}_v$ denotes the sparse vectors only involving the i u-components $\tilde{u}(1), \ldots, \tilde{u}(i)$ and j v-components $\tilde{v}(1), \ldots, \tilde{v}(j)$ in the same component positions, respectively and zeros elsewhere,

$$\begin{aligned} x_u &= \boldsymbol{P}_u^{\mathrm{T}} \begin{bmatrix} \tilde{u}(1:i) \\ 0 \end{bmatrix}, \\ y_v &= \boldsymbol{P}_v^{\mathrm{T}} \begin{bmatrix} \tilde{v}(1:j) \\ 0 \end{bmatrix}. \end{aligned} \tag{4}$$

According to the definition and demonstration in Ref. [8], the sparsest decomposition for $\boldsymbol{A}$ equals a penalized optimization problem

$$F(\alpha, \beta, p) = \min_{\xi} \left\{ \alpha * i_\xi + \beta * j_\xi - \left(\frac{h(\xi)}{\|\boldsymbol{A}\|_2} \right)^p \right\}, \tag{5}$$

where

$$j_\xi = d(x_u, y_v) = \frac{x_u^{\mathrm{T}} \boldsymbol{A} y_v}{\|x_u\| \, \|y_v\|}$$

depends on the parameter ξ and i_ξ, j_ξ are the numbers of non-zeros in x_u and y_v, respectively, α and β are penalty constants. An efficient numerical method for finding the numerical solution of Eq. (5) is also available in Ref. [8]. From Ref. [8], we find that SD technique is not only fit for handling sparse observed matrices, but also providing similar approximation accuracy to the SVD.

Based on the above analysis, it is self-evident that: Given an observed data $\boldsymbol{A} = \tilde{\boldsymbol{A}} + \boldsymbol{N}$, where $\boldsymbol{N}$ is the AWGN, independent of the $\tilde{\boldsymbol{A}}$, then it is easy to obtain that

$$\begin{aligned} \boldsymbol{A} &= \tilde{\boldsymbol{A}} + \boldsymbol{N} \\ &= \boldsymbol{X}_{\mathrm{S}} \boldsymbol{D}_{\mathrm{S}} \boldsymbol{Y}_{\mathrm{S}}^{\mathrm{T}} + \boldsymbol{X}_{\mathrm{N}} \boldsymbol{D}_{\mathrm{N}} \boldsymbol{Y}_{\mathrm{N}}^{\mathrm{T}} \\ &= [\boldsymbol{X}_{\mathrm{S}}, \boldsymbol{X}_{\mathrm{N}}] \begin{bmatrix} \boldsymbol{D}_{\mathrm{S}} & 0 \\ 0 & \boldsymbol{D}_{\mathrm{N}} \end{bmatrix} [\boldsymbol{Y}_{\mathrm{S}}, \boldsymbol{Y}_{\mathrm{N}}]^{\mathrm{T}}, \end{aligned} \tag{6}$$

where S denotes signal subspace and N denotes noise subspace.

The main steps in the TSD-based denoising algorithm can be summarized as follows:

(A) The SD operation to the observed data $\boldsymbol{A}$;

(B) Determination of the optimal value of t for the best separating of the signal and noise;

(C) Reconstruction of the observed data $\boldsymbol{A}_t = \boldsymbol{X} \boldsymbol{D}_t \boldsymbol{Y}^{\mathrm{T}}$.

Given that r is the rank of $\boldsymbol{A}$, $r \leq \min(m, n)$, then in SVD, $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}$, t $(1 \leq t \leq r)$ denotes the truncated rank of $\boldsymbol{S}$ by setting the trailing $n - t$ singular values to 0.

Minimum energy model truncate operation Similar to the SVD-based denoising method first proposed in Ref. [9], the TSD-based denoising algorithm also depends on the accurate estimation of the value t. In recent years, variational and partial differential equation (PDE) methods have also been widely used in many applications.[10,11] There are several methods for solving a PDE and the principle behind them is to minimize energy functionality. We assume a variational approach to signal processing formulated as

$$\arg \left\{ \min_a E(a) \right\}, \tag{7}$$

where $[a_{ij}] \in \boldsymbol{A}$, $i = 1, \ldots, m$, $j = 1, \ldots, n$.

In order to restore the signal through the way of TSD, we establish a minimum energy model in matrix space as follows:

$$E(a) = \alpha \int_{\Omega} \|\nabla a\| \, \mathrm{d}i\mathrm{d}j + (1-\alpha) \int_{\Omega} \|a - a_0\| \, \mathrm{d}i\mathrm{d}j, \tag{8}$$

where $\|\cdot\|$ denotes the ℓ^2 norm, a_0 denotes noisy data and a denotes restored data here. Our objective is to find the optimal $a_{\mathrm{opt}} = a_t$ which makes the following formula into existence:

$$E(a_{\mathrm{opt}}) = \min \{E(a_i)\}, \ i = 1, 2, \ldots, r. \tag{9}$$

The model established above consists of two items, which can be denoted as follows:

$$\begin{cases} E_1(a) = \int\limits_{\Omega} \|\nabla a\| \mathrm{d}i\mathrm{d}j, \\ E_2(a) = \int\limits_{\Omega} \|a - a_0\| \, \mathrm{d}i\mathrm{d}j. \end{cases} \tag{10}$$

The first item $E_1(a)$ in the model presents the smoothness of data, called the smoothness energy and the second item $E_2(a)$ presents the disparity between the restored data and the noisy data, called the disparity energy.[9] Here we use the following calculation method:

$$\begin{aligned} \|\nabla a\| &= \sqrt{a_{ij}^2 + a_{i+1j+1}^2}, \\ \|a - a_0\| &= \sqrt{(a - a_0)^2}. \end{aligned} \tag{11}$$

This model established above avoids the overloaded computation for calculating PDE, because it is designed to find the optimal a_t in the finite steps.

3. Experiments

To evaluate the proposed algorithm, the signal-to-noise ratio (SNR) is adopted as a measure of noise in this paper. Its definition is given as follows:

$$\mathrm{SNR} = 10\lg\sigma_x^2/\sigma_e^2, \tag{12}$$

where σ_x and σ_e are the standard deviations of the signal and noise, respectively.

Two chaotic signals 'logistic map' and 'coupled map lattices (CMLs)'[12−15] are given as follows: (i) logistic map

$$a(l) = ja(l-1)(1-a(l-1)), \quad 0 \le j \le 4, \tag{13}$$

where j is the parameter of logistic map set as $j = 3.9$, which guarantees the chaotic motion of the map; (ii) CMLs combines a closed loop structure with length R and an open loop structure with length $L - R$,

$$a_p(t) = \begin{cases} (1-\varepsilon) f(a_p(l-1),\mu) + \varepsilon f(a_{p-1}(l-1),\mu), & 1 \le p \le R, \\ (1-0.5\varepsilon) f(a_p(l-1),\mu) + 0.5\varepsilon f(a_{p-1}(l-1),\mu), & R < p \le L, \end{cases} \tag{14}$$

where $a_0(l) \equiv a_R(l)$, the subscript $p = 1, 2, \ldots, L$ which denotes the lattice site index, l is the discrete time index, ε is the coupling constant and μ is the parameter of the nonlinear function $f(a_p, \mu)$. In this study, we adopt the tent map to represent the nonlinear function as follows:

$$f(a_p,\mu) = \begin{cases} 2\mu a_p, & 0 < a_p \le 0.5, \\ 2\mu(1-a_p), & 0.5 < a_p < 1. \end{cases} \tag{15}$$

The parameters of CMLs are set as $L = 188, R = 28, \mu = 0.66, n = 3000$ and $\varepsilon = 0.8$. The original logistic map and the original high dimensional CMLs are shown in Fig. 1.

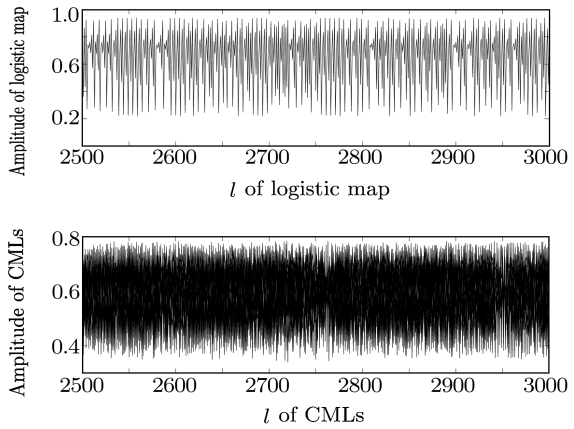


Fig. 1. The original logistic map (above) and the original CMLs (below).

These signals with varying AWGN are used to test the performance of TSD. The results are depicted in Figs. 2 and 3 respectively.

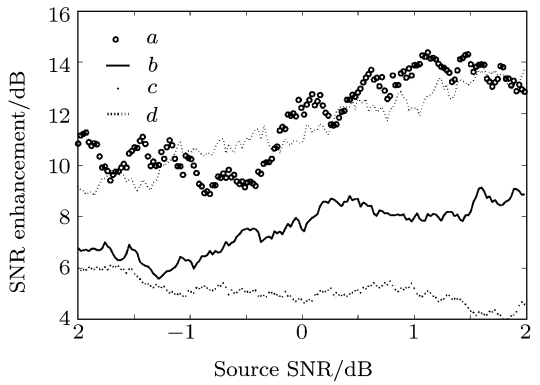


Fig. 2. Denoising results of noisy logistic map, in which curves a, b, c and d are for TSD, SVD, WS and NNSC, respectively.

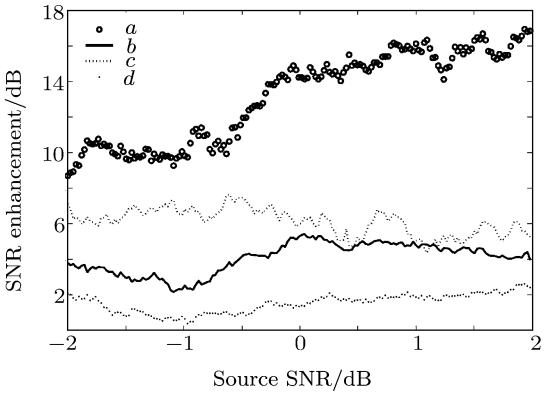


Fig. 3. Denoising results of noisy CMLs, in which curves a, b, c and d are for TSD, SVD, NNSC and WS, respectively.

The results indicate that the TSD is more efficient than NNSC[4] and SVD-based denoising algorithm.[9] The WS has the worst average performance in both experiments, thus it might be not suitable for denoising noisy chaotic signals.

Based on these results, we believe that the advantage of the TSD results from two aspects: First,

the SD is a good decomposition technique, which can decompose the input data by a more suitable and generalized criterion. Second, the used minimum energy model truncate operation helps us to recover ideal manifolds' structure, which also improves the denoising performance.

4. Conclusion

The TSD algorithm has been proposed in this paper, which is a combination of SD and the minimum energy model truncation operation. The principle is to apply the SD to project the observed data onto two sparse spaces, in which ideal manifolds' structure can be explored by the minimum energy model truncate operation and then the ideal data can be extracted. Experiments confirm the high performance of TSD in the case of real chaotic signals.

Acknowledgement

The authors thank Ms. Zheng S.Q. and her research team for their remarkable suggestions.

References

[1] Chin T J and Suter D 2007 *IEEE Trans. Image Process.* **16** 1662
[2] Balster E J, Zheng Y F and Ewing R L 2006 *IEEE Trans. Circuits Syst. Video Technol.* **16** 220
[3] Han J J and Fu W J 2010 *Chin. Phys.* B **19** 010205
[4] Hoyer P O and Hyvuarinen A 2002 *Vision Research* **42** 1593
[5] Manfredi C, Aniello M D and Bruscaglioni P 2001 *Logopedics Phoniatrics Vocology* **26** 179
[6] Li E Z, Ling B L, Liu Y, Ti A, Hu L Q and Gao X 2010 *Chin. Phys.* B **19** 035203
[7] Onn R, Steinhardt A O and Bojanczyk A W 1991 *IEEE Trans. Signal Process.* **39** 1575
[8] Zhang Z Y and Zha H Y 2004 *Science in China* Ser. A *Mathematics* **47** 908
[9] Zhang Z, Zhang H, Sang H and Yuan W 2008 *Proc. 2008 Congress on Image and Signal Processing* **1** 411
[10] Momania S and Odibatb Z 2007 *Phys. Lett.* A **365** 345
[11] Odibata Z and Momanib S 2008 *Appl. Math. Modelling* **32** 28
[12] Xie Z B and Feng J C 2009 *Chin. Phys. Lett.* **26** 030501
[13] Liu Z C 2009 *Chin. Phys.* B **18** 636
[14] Cui D, Gao Z Y and Zhao X M 2008 *Chin. Phys.* B **17** 1703
[15] Xie Y H, Yang W and Zhang Y 2010 *Acta Phys. Sin.* **59** 8255 (in Chinese)