Journal of Information Hiding and Multimedia Signal Processing ©2016 ISSN 2073-4212
Ubiquitous International Volume **7**, Number **5**, September **2016**

Truncation Strategy of Tensor Compressive Sensing for Noisy Video Sequences

Qing-Zhu Wang

School of Information Engineering
Northeast Dianli University
169 Changchun Road, Jilin, China
150681573@qq.com

Wan-Jun Kang

School of Information Engineering
Northeast Dianli University
169 Changchun Road, Jilin, China
464760918@qq.com

Received March, 2016; revised June, 2016

Abstract. *Tensor Compressive Sensing (TCS) is an emerging approach for higher order data representation, such as medical imaging, video sequences, and hyperspectral images. In this paper, we present a Truncated TCS (TTCS) for noisy Three-dimensional (3D) video sequences. Inspired by the Marcenko-Pastur Probability Density Function (MPPDF) for the estimation of an unknown noise level, we extend a MPPDF-based optimal hard thresholding of singular value for noisy matrices into its tensor counterpart. Further, an asymptotic estimation of the optimal hard threshold is used for recovery of low-rank tensors from noise data. Finally, we apply this truncation strategy to a Tucker-based TCS for representation of noisy 3D video sequences and demonstrate experimentally that it outperforms state of the art.*
Keywords: Tensor compressive sensing; Optimal hard threshold; Truncated higher-order singular value decomposition

1. **Introduction.** During the last years there has been an increased interest in Compressed Sensing (CS) , which provides a general signal acquisition framework that enables the reconstruction of sparse signals from a small number of linear measurements [1]. Most of the development of CS was focused on problems involving 1D signal or 2D image data encoded in vectors [2]. However, many important applications involve higher dimensional signals or tensors such as video sequences, 3D medical images and hyperspectral images. In areas other than CS, a lot of interests have been conducted on tensor-based approaches for higher dimensional data analysis. Therefore, the higher-order extension of CS theory for multidimensional data has become an emerging topic. Recently, Cesar et al [3] provides a direct reconstruction formula to recover a tensor from a set of multilinear projections that are obtained by multiplying the data tensor by a different sensing matrix in each mode. Compared to existing sparsity-based CS methods [4-6], this Tucker-based tensor CS (Tucker-TCS) does not require to assume sparsity neither a dictionary based representation. In addition, it is super fast because it does not involve iterations making it potentially suitable for large-scale problems.

Above multidimensional CS are all for clear images. However, real world images may suffer from serious random noise, which affects the image quality and decrease the reliability of low-rank representation. As current TCS methods do not consider the effect of noise on low-rank representation of tensors, we develop a Truncation strategy of TCS (TTCS) for 3D noisy video sequences: (1) Inspired by the Marcenko-Pastur Probability Density Function (MPPDF) for the estimation of an unknown noise level, we extend a MPPDF-based optimal hard thresholding of singular value for noisy matrices into its tensor counterpart; (2) We use an asymptotic estimation of the optimal hard threshold to design TTCS for recovery of low-rank tensors from noise data.

This paper is organized as follows: in Section 2, the related notation, definition and basic results used throughout the paper, are introduced; in section 3, TTCS are proposed along with its detailed profs; in Section 4, several numerical results based on 3D video sequences are provided, validating our theoretical results and evaluating the stability and robustness of our proposed scheme, in section 5, the main conclusions of the present work are outlined.

2. Notations.

2.1. **Compressive Sensing.** Traditional CS is a framework for reconstruction of signals that have sparse representations [7]. A vector $\mathrm{x} \in R^M$ is called $s-$sparse if it has nonzero entries. The CS measurement protocol measures the signal x with the measurement matrix $\mathrm{A} \in R^{m\times \mathrm{M}}$ where $m < M$ and encodes the information as

$$\mathrm{y} = \mathrm{Ax} \tag{1}$$

The decoder knows A and attempts to recover x from y .

2.2. **Tensor notations.** A tensor is a multidimensional array [8-9]. The order of a tensor is the number of modes. For instance, tensor $\mathcal{X} \in R^{M_1\times\cdots\times M_d}$ has order d and the dimension of its n th mode is M_n. The decomposition and reconstruction of can be written as follows:

$$\begin{cases} \mathcal{W} = \mathcal{X} \times_1 \Phi^{(1)T} \times_2 \Phi^{(2)T} \times_3 \Phi^{(3)T} \\ \mathcal{X} = \mathcal{W} \times_1 \Phi^{(1)} \times_2 \Phi^{(2)} \times_3 \Phi^{(3)} \end{cases} \tag{2}$$

3. Proposed TTCS system.

In this section, we would like to introduce our TTCS system. Before going into the details, we analyze the relationship between truncation and low-rank representation of the tensors from noise data. Then the whole TTCS system is introduced step by step. Note that TTCS does not require to assume sparsity neither a dictionary based representation similar to Tucker-TCS [3].

3.1. **Truncation strategy.** For a given 3D tensor $\mathcal{X} \in R^{M_1\times M_2\times M_3}$ we have $\mathcal{X} = \mathcal{X}_0+\mathcal{E}$, where $\mathcal{E}$ is a noise tensor with independent and identically distributed zero-mean Gaussian samples. Here we use the Tucker-TCS [3] which is more stable, robust and accuracy than HOSVD to decompose and reconstruct $\mathcal{X}$:

$$\begin{cases} \mathcal{W} = \mathcal{X} \times_1 \Phi^{(1)T} \times_2 \Phi^{(2)T} \times_3 \Phi^{(3)T} \\ \mathcal{X} = \mathcal{W} \times_1 \mathrm{Z}_1\mathrm{W}^{\dagger}_{(1)} \times_2 \mathrm{Z}_2\mathrm{W}^{\dagger}_{(2)} \times_3 \mathrm{Z}_3\mathrm{W}^{\dagger}_{(3)} \end{cases} \tag{3}$$

The truncated MP pseudo-inverse $W^{\dagger}_{(n)}$, i.e., the core part of the truncation strategy of TCS may eliminate the noise and preserve the desired signal, which is discussed below. By using the SVD, $W_{(n)}$ and $W^{\dagger}_{(n)}$ can be defined as follows:

$$W_{(n)} = U_n S_n V_n^T \tag{4}$$

with entries of the diagonal matrix S_n defined as $s_i^{(n)}$ and

$$W_{(n)}^{\dagger} = V_n \tilde{S}_n^{\tau^*} U_n^T \tag{5}$$

with entries of the diagonal matrix S_n defined as follows:

$$\tilde{s}_i^{(n)} = \begin{cases} \frac{1}{s_i^{(n)}}, & for \quad s_i^{(n)} > \tau_n^* \\ 0, & for \quad s_i^{(n)} \leq \tau_n^* \end{cases} \tag{6}$$

where τ_n^* is the hard threshold. The other parameters of equation (3) are as follows:

$$Z^{(n)} = \begin{cases} \mathcal{X} \times_2 \Phi_2 \times_3 \Phi_3, & for \quad n = 1 \\ \mathcal{X} \times_1 \Phi_1 \times_3 \Phi_3, & for \quad n = 2 \\ \mathcal{X} \times_1 \Phi_1 \times_2 \Phi_2, & for \quad n = 3 \end{cases} \tag{7}$$

$$Z_n = (Z_{(n)}^{(n)}) \tag{8}$$

Therefore, the design of TTCS will raise a question: how to choose an optimal hard threshold.

3.2. **Optimal hard threshold.** Many researchers assume Marcenko-Pastur Probability Density Function (MPPDF) for the estimation of an unknown noise level [10-11. [11] defines an optimal hard threshold for a 2D matrix $X \in R^{m \times M}$observed in unknown noise level, with $\beta = \frac{m}{M}$. We extend the method into its 3D tensor counterpart as follows:

$$\tau^{(n)*} = \frac{\lambda_n}{\sqrt{\mu_n}} \tau_{med}^{(n)} \tag{9}$$

where $\sigma^{(n)*}(n = 1, 2, 3)$is the estimated noise variance, $\sigma_m^{(n)}ed$ is the median singular value of $W_{(n)}$, and μ_n is the median of the Marcenko-Pastur distribution, namely, the unique solution in $\beta_{n,-} \leq x \leq \beta_{n,+}$ to the equation

$$\int_{\beta_{n,-}}^{x} \frac{\sqrt{(\beta_{n,+} - t)(t - \beta_{n,-})}}{2\pi\beta_n t} dt = \frac{1}{2} \tag{10}$$

where $\beta_{n,\pm} = (1 \pm \sqrt{\beta_n})^2$. As described in 2.2 and 3.1, the dimension of W_n is

$$W_{(n)} \in \begin{cases} R^{m_1 \times m_2 m_3}, & for \quad n = 1 \\ R^{m_2 \times m_1 m_3}, & for \quad n = 2 \\ R^{m_3 \times m_1 m_2}, & for \quad n = 3 \end{cases} \tag{11}$$

The corresponding β_n is

$$\beta_n = \begin{cases} \frac{m_1}{m_2 m_3}, & for \quad n = 1 \\ \frac{m_2}{m_1 m_3}, & for \quad n = 2 \\ \frac{m_3}{m_1 m_2}, & for \quad n = 3 \end{cases} \tag{12}$$

Inspired by the optimal threshold coefficient for matrices [11], the optimal threshold coefficient λ_n for 3D tensors can be defined as follows:

$$\lambda_n = \sqrt{2(\beta_n + 1) + \frac{8\beta_n}{(\beta_n + 1) + \sqrt{\beta_n^2 + 14\beta_n + 1}}} \tag{13}$$

Algorithm 1. Chosen of optimal threshold

Input:

(1) β_n ($n = 1,2,3$);

(2) The number of steps K .

Output:

The optimal hard threshold τ_n^* .

Start:

for *n*=1,2,3, do

Initial sets *I* and *J*;

for *k*=1:*K* , do

compute $y_{n,k}$ as (14);

update sets *I* and *J* as (15) ;

update $\beta_{n,+}^{(k)}$ and $\beta_{n,-}^{(k)}$ as (16);

end for

update μ_β as (17);

Compute $\sigma^{(n)*}$ as (9), (13) and (17);

Compute τ_n^* as (18)-(20);

end for

End

Algorithm 2. TTCS system

Input:

(1) Noisy video sequences tensor $\mathcal{X} \in R^{M_1 \times M_2 \times M_3}$;

(2) Sensing matrices $\Phi^{(n)} \in R^{m_n \times M_n}$ ($n = 1,2,3$).

Output:

Low rank reconstruction $\widehat{\mathcal{X}}$.

Start:

(1) Compute $\mathbf{W}_{(n)}$ of $\mathcal{W}$ and the corresponding β_n according to formula (12) ;

(2) Compute the optimal hard threshold τ_n^* corresponding to $\mathbf{W}_{(n)}$ by algorithm 1;

(3) Compute the truncated MP Pseudo-Inverse $\mathbf{W}_{(n)}^{\dagger}$ according to τ_n^* by equations (4)-(6);

(4) Reconstruct $\widehat{\mathcal{X}}$ according to (3);

End

As the median μ_n is not available analytically, we make available an asymptotic method to solve it. Consider

$$y_{n,k} = 1 - \int_{\beta_{n,-}+(k-1)\Delta}^{\beta_{n,-}+k\Delta} \frac{\sqrt{(\beta_{n,+}-t)(t-\beta_{n,-})}}{2\pi\beta_n t} dt \tag{14}$$

where K is the number of steps, $\Delta = \frac{\beta_{n,+}-\beta_{n,-}}{K}$ and $k = 1, \cdots, K$. Let the initial sets $J = \{0\}\text{and}I = \{0\}$. Update J and I for y_n, k as follows:

$$\begin{cases} J = \{J, k\}, & for \quad y_{n,k} < \frac{1}{2} \\ I = \{I, k\}, & for \quad y_{n,k} > \frac{1}{2} \end{cases} \tag{15}$$

Further update $\beta_{n,-}^{(k)}$ and $\beta_{n,+}^{(k)}$as follows:

$$\begin{cases} \beta_{n,-}^{(k)} = \max J \\ \beta_{n,+}^{(k)} = \min I \end{cases} \tag{16}$$

After K steps, μ_n could be obtained as

$$\mu_n = \frac{\beta_{n,-} + \beta_{n,+}}{2} \tag{17}$$

The $\sigma^{(n)*}$ can be solved through equations (9)-(17). We compute the optimal hard threshold through the proportion of the singular values which are lager than $\sigma^{(n)*}(j = 1, \cdots, m_n)$. Let

$$l_n = \sum_{j=1}^{m_n} v_j \tag{18}$$

where

$$v_j = \begin{cases} 1, & for \quad \sigma_j^{(n)} > \sigma^{(n)*} \\ 0, & for \quad \sigma_j^{(n)} \leq \sigma^{(n)*} \end{cases} \tag{19}$$

The optimal threshold can be defined as follows:

$$\tau_n^* = \mathrm{s}_n([\frac{l_\mathrm{n}}{L_\mathrm{n}}]) \tag{20}$$

where $\mathrm{s}_n = \mathrm{diag}(\mathrm{S}_n)$, L_n is the number of all singular values and $[\cdot]$ is rounding function. The chosen of the optimal threshold and complete TTCS system are shown in algorithms1 and 2, respectively.

4. **Simulation Results and Analysis.**

4.1. **Experiments setup.** KCS [4] outperforms several other methods including independent measures and partitioned measurements in terms of reconstruction accuracy in tasks related to compression of multidimensional signals. CP-TCS-P [5] stands out for its reconstruction efficiency compared with a Multi-way CS (MWCS) [6]. Tucker-TCS [3] has significant advantages compared to existing sparsity based CS methods. For the above reasons, we experimentally compare TTCS with KCS, CP-TCS and Tucker-TCS on the reconstruction of 13 commonly used video test sequences in QCIF formats [12]. Each frame of the video sequence is preprocessed to have 128128 and we choose 128 frames. To evaluate the performance under different noise levels, the videos are added with varying levels of Gaussian noise (mean zero, the variance varies from 0 to 40 with an increase of 1). The quantitative measure is the Peak Signal Noise Ratio (PSNR).

Tab.1 PSNR (dB) comparison of TTCS and state of the art methods

Video Name	Sampling ratio	Methods	Noise Level			
			$\sigma_v = 0$	$\sigma_v = 10$	$\sigma_v = 20$	$\sigma_v = 30$
Mother and Daughter	$\frac{(120)^3}{(128)^3} = 0.82$	KCS	46.65	28.19	22.37	18.93
		CP-CS	30.58	26.22	21.60	18.43
		Tucker-CS	49.18	29.39	23.20	19.60
		TTCS	***49.18***	***34.06***	***29.78***	***26.66***
	$\frac{(112)^3}{(128)^3} = 0.67$	KCS	42.58	28.19	22.61	19.24
		CP-CS	30.50	26.19	21.59	18.42
		Tucker-CS	44.87	29.64	23.80	20.32
		TTCS	***44.87***	***33.24***	***30.12***	***27.05***
	$\frac{(104)^3}{(128)^3} = 0.54$	KCS	40.51	27.64	22.61	19.24
		CP-CS	30.36	26.12	21.55	18.42
		Tucker-CS	41.98	29.88	24.36	20.98
		TTCS	***41.98***	***32.70***	***29.05***	***26.96***
	$\frac{(96)^3}{(128)^3} = 0.42$	KCS	33.47	27.64	22.94	19.80
		CP-CS	30.26	26.09	21.54	18.39
		Tucker-CS	39.46	29.95	24.89	21.71
		TTCS	***39.46***	***32.00***	***29.03***	***26.90***
Foreman	$\frac{(120)^3}{(128)^3} = 0.82$	KCS	42.58	27.93	22.35	19.09
		CP-CS	24.26	27.87	22.19	18.76
		Tucker-CS	44.96	28.97	22.96	19.43
		TTCS	***44.13***	***30.61***	***28.50***	***26.39***
	$\frac{(112)^3}{(128)^3} = 0.67$	KCS	37.23	27.56	22.40	19.29
		CP-CS	24.79	27.94	22.08	18.76
		Tucker-CS	40.48	28.79	23.38	19.95
		TTCS	***40.48***	***29.17***	***27.96***	***25.88***
	$\frac{(104)^3}{(128)^3} = 0.54$	KCS	33.61	27.02	22.39	19.45
		CP-CS	24.53	27.72	22.18	18.75
		Tucker-TCS	37.71	28.52	23.61	20.49
		TTCS	***37.71***	***29.11***	***27.03***	***25.41***
	$\frac{(96)^3}{(128)^3} = 0.42$	KCS	30.73	26.26	22.24	19.52
		CP-TCS	24.46	27.94	22.18	18.76
		Tucker-TCS	34.92	28.20	23.74	20.97
		TTCS	***34.92***	27.81	26.38	24.93

Tab.2 Computation time (s) comparison of TTCS and state to the art methods

Video Name	Sampling ratio	Methods	Average computation time (s)
Mother and Daughter	$\frac{(120)^3}{(128)^3}=0.82$	KCS	1923
		CP-CS	1146
		Tucker-CS	15
		TTCS	***16***
	$\frac{(112)^3}{(128)^3}=0.67$	KCS	2907
		CP-CS	1399
		Tucker-CS	9
		TTCS	***10***
	$\frac{(104)^3}{(128)^3}=0.54$	KCS	2493
		CP-CS	1565
		Tucker-CS	9
		TTCS	***9***
	$\frac{(96)^3}{(128)^3}=0.42$	KCS	2916
		CP-CS	1140
		Tucker-CS	8
		TTCS	***9***
Foreman	$\frac{(120)^3}{(128)^3}=0.82$	KCS	1901
		CP-CS	1347
		Tucker-CS	14
		TTCS	***15***
	$\frac{(112)^3}{(128)^3}=0.67$	KCS	4117
		CP-CS	1312
		Tucker-CS	12
		TTCS	***13***
	$\frac{(104)^3}{(128)^3}=0.54$	KCS	2586
		CP-CS	1396
		Tucker-TCS	13
		TTCS	***13***
	$\frac{(96)^3}{(128)^3}=0.42$	KCS	2631
		CP-TCS	1432
		Tucker-TCS	8
		TTCS	***10***

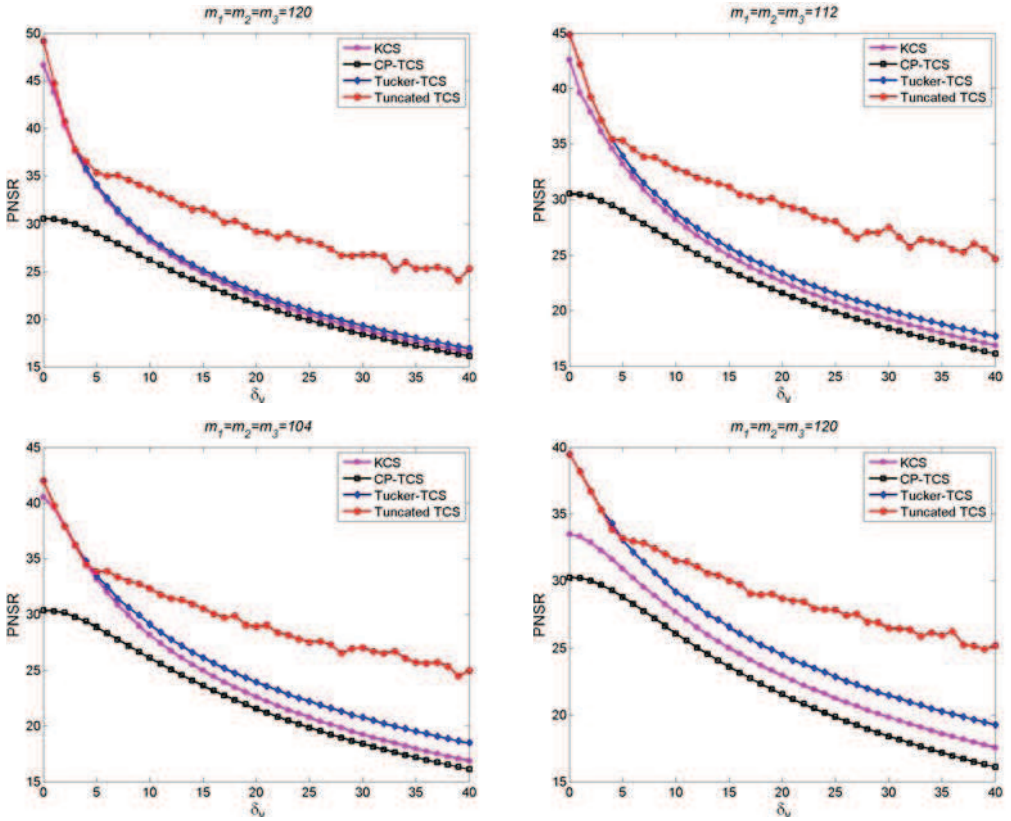

Fig.1 Illustration of reconstruction performances for video with different sampling ratio and noise level

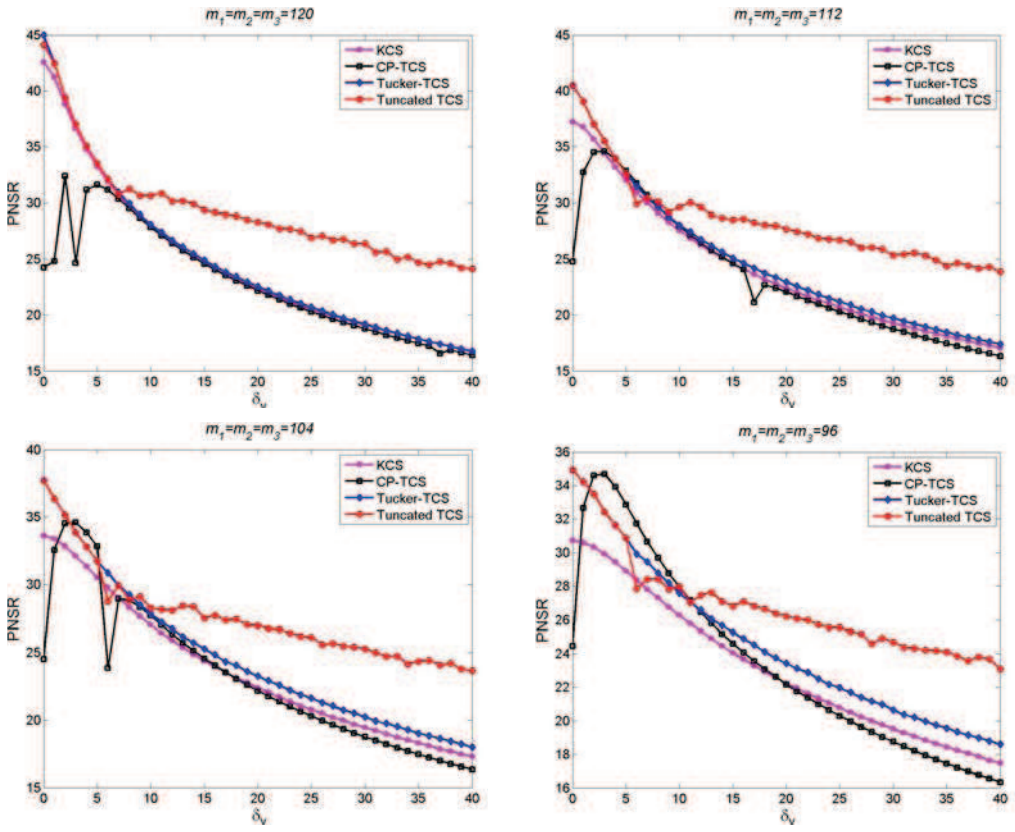

Fig.2 Illustration of reconstruction performances for “Foreman” with different sampling ratio and noise level

Fig.3 The original Mother and Daughter video frames

Fig.4 Noisy video frames

Fig.5 Reconstructed video frames by KCS

Fig.6 Reconstructed video frames by CP-TCS

Fig.7 Reconstructed video frames by Tucker-TCS

Fig.8 Reconstructed video frames by proposed TTCS

4.2. **Results analysis.** Tables 1 and 2 compare the PNSR and the computation time at some noise levels of above methods on two of the videos mentioned in 4.1 respectively. The superiorities of TTCS are highlighted. TTCS shows outstanding performance in comparison with KCS and CP-TCS, in terms of both accuracy and speed. What needs to be stressed is that Tucker-TCS and TTCS are super fast compared with KCS and CP-TCS. TCS requires slightly more computation time in comparison with Tucker-TCS and can be safely ignored. It can also be found that the advantage of TTCS over Tucker-TCS increase with the noise levels. Figs.1-2 compare the PSNR for the two videos when , , and . As noise level increases, TTCS tends to outperform others in term of reconstruction accuracy. Figs.3-8 provide a visual evaluation of reconstruction results under video "Mobile and Daughter" with noise level. Original and noisy frames are shown in Figs.3-4. We specifically look into the recovered frames of all methods when . Recovered frames 1, 33, 65, 97, 128 of KCS, CP-TCS, Tucker-TCS and TTCS are depicted as examples in Figs.5-8.

5. **Conclusion.** TTCS is a novel truncation strategy of TCS that takes advantages of both MPPDF and Tucker-TCS. The former efficiently eliminates the noise; the later improves the reconstruction performance of desired signal. TTCS is applied and compared with current state-of-the art methods on commonly used video test sequences. The experimental results demonstrate that the proposed TTCS outperforms other multidimentional CS algorithms in noisy 3D video sequences.

6. **Acknowledgment.** This work was supported by National Natural Science Foundation of China (61301257).

REFERENCES

[1] Y. Song, W. Cao, Y. F. Shen,eds. Compressed Sensing Image Reconstruction Using Intra Prediction, *Neurocomputing*,vol.151, pp.1171–1179, 2006.

[2] J. B. Yang, X. J. Liao, X. Yuan, eds. Compressive Sensing by Learning a Gaussian Mixture Mode From Measurements, *IEEE Transactions on Image Processing*, vol.24, no.1, pp.106–119, 2015.

[3] F. C. Cesar,C. Andrzej, Stable, Robust, and Super Fast Reconstruction of Tensors Using Multi-Way Projections, *IEEE Transactions on Signal Processing*, vol.63, no.3, pp.780–793, 2015.

[4] F. D. Marco, G. B. Richard, Kronecker Compressive Sensing, *IEEE Transactions on Image Processing*, vol.21, no.2, pp.494–504, 2012.

[5] S. Friedland, Q. Li, D. Schofeld, Compressive Sensing of Sparse Tensors, *IEEE Transactions on Image Processing*, vol.23, no.10, pp.4438–4446, 2014.

[6] N. D. Sidiropoulos,A. Kyrillidis, Multi-way Compressed Sensing for Sparse Low-rank Tensors, *IEEE Signal Processing Letters*,vol.19, no.11, pp.757–760, 2012.

[7] D. L. Donoho, Compressed Sensing, *IEEE Transactions Information Theory*,vol.52, no.4, pp.1289–1306, 2006.

[8] N. Vannieuwenhoven, A New Truncation Strategy for the Higher-order Singular Value Decomposition, *Siam Journal on Scientific Computing*,vol.34, no.2, pp.1027–1052, 2012.

[9] P. Gastaldo, L. Pinna, L. Seminara,(eps.), Computational Intelligence Techniques for Tactile Sensing Systems, *Sensors*,vol.14, no.6, pp.10952–10976, 2014.

[10] G. Matan, L. D. David, The Optimal Hard Threshold for Singular Values is $\frac{4}{\sqrt{3}}$, *IEEE Transactions on Information Theory*,vol.60, no.8, pp.5040–5053, 2014.

[11] R. N. Raj, OptShrink: An Algorithm for Improved Low-Rank Signal Matrix Denoising by Optimal, Data-Driven Singular Value Shrinkage, *IEEE Transactions on Information Theory*,vol.60, no.5, pp.3002–3018, 2014.

[12] P. Seeling, M. Reisslein, Video Traffic Characteristics of Modern Encoding Standards: H.264/AVC with SVC and MVC Extensions and H.265/HEVC, *The Scientfic World Journal*,vol.2014, pp.1–16, 2014.