Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **3**, May **2018**

Joint Spatio-Temporal Domain for Adaptive Kronecker Compressive Sensing

Hui Luo[1], Yuandong Wang[2*], Ning Zhang[3], Wenyuan Yu[4], Huan Zhang[5]

School of Information Engineering,
East China Jiaotong University
Nanchang 330013, China
[2*]Corresponding author: wyd_jxnc@163.com
[1]lh_jxnc@163.com, [3]zn_jxnc@163.com,
[4]ywy_jxnc@163.com [5]zh_jxnc@163.com

Received April, 2017; revised November, 2017

Abstract. *Compressive Sensing (CS) theory tells that a sparse or compressible signal can be recovered exactly from a few sampling measurements via solving a convex problem although the sampling rate is lower than the Nyquist rate. However, the CS literature has mostly focused on problems involving 1-D signals and 2-D images, many important applications involve multidimensional signals such as video signal. In this paper, based on the Kronecker compressive sensing (KCS), a novel joint spatio-temporal domain for Adaptive Kronecker Compressive Sensing (AKCS) method is proposed. Firstly, video sequences are divided into space domain and time domain. Secondly, in the space domain, adaptive sampling scheme is formulated based on statistical information of the image block, and a new measurement matrix is designed by using Kronecker product. Lastly, high-quality video can be reconstructed by applying AKCS in spatio-temporal domain. The improved BCS-IST algorithm is used to calculate the residual error of each block, and the current frame is combined with the predicted frame to reconstruct the current frame. When reconstructing the video signal, the time correlation between frames should be fully utilized besides considering the spatial correlation in the frame. Kronecker product can well describe the correlation between pixels in spatio-temporal domain, and is often used in frame rate lifting algorithm and edge information estimation algorithm.The experimental results demonstrate that the proposed method can effectively improve the peak-signal-to-noise ratio (PSNR) and running efficiency compared with other methods.*
Keywords: Compressed sensing; Kronecker product; Spatio-temporal domain; Adaptive sampling; Video reconstruction

1. **Introduction.** Compressive Sensing has attracted considerable attention for its capability of simultaneous sampling and compression over the past few years [1-4]. Despite measuring a signal with a much smaller number of measurements than the Nyquist sampling rate, CS can still perfectly reconstruct the signal by relying on the sparsity /compressibility property of natural signals in some sparse domains. The CS literature has mostly focused on problems involving single sensors and one-dimensional (1-D) or 2-D data. However, many important applications that hold the most promise for CS involve signals that are multidimensional such as video signal, and the coordinates of these signals may span several temporal or spectral dimensions. In this regard, Duarte et al. [5, 6] introduced an innovative CS scheme that employs the Kronecker product to meet the

spatio-temporal structure of video. The Kronecker Compressive Sensing method pointed out that the signal could be jointly reconstructed in the video sampling based on CS [7].

The high-dimensional sparsity basis can be obtained by using Kronecker product as well as measurement matrix can be designed by a formula. It means the two matrices can realize the joint reconstruction according to the measured values without changing the traditional signal sampling. After that, Xinwei Ye et al. [8] proposed optimal spatio-temporal projections with Holo-Kronecker Compressive Sensing (HKCS) of video acquisition. In their sampling scheme, the original identity matrix, a factor of the conventional Kronecker product measurement matrix, was replaced by an ill-posed matrix which enables compressive samplings along the temporal dimension, which can optimally utilize the redundancy spanning all the dimensions of the signal and the necessary measurements for exact recovery are significantly reduced. Although those methods can effectively improve the quality of the reconstructed video, the storing space and computing burden will increases sharply as the dimensionality grow, and this is the key issue impeding its application.

In order to further improve the quality of the reconstructed video, reduce the storing space and computing burden, we put forward a joint spatio-temporal domain for adaptive Kronecker compressive sensing,First, by measuring the Kronecker product matrix operations to structure suitable to the video signal, to achieve the overall compression measurement of video signals; then the multiple hypothesis (Multiple Hypothesis MH) model combined with minimum TV model, forecast the temporal and spatial characteristics of the construction of joint residual reconstruction model, and predicted frames of the current frame is obtained by iteration; finally calculation of the residual block using the improved BCS-IST algorithm proposed in this paper, and the predictive frame and the current frame are combined to reconstruct the current frame. The experimental results show that the proposed method can effectively improve the quality of video reconstruction while reducing the computational complexity and speeding up the algorithm.An improved BCS-IST algorithm is proposed. Compared with the traditional BCS-IST algorithm, this algorithm can better preserve the structure and texture information of the image, and achieve better image reconstruction effect, and accelerate the convergence speed to a certain extent.

The main contributions of this paper are as follows: 1) each block according to the statistical information of the image blocks are divided into complex fault block or simple block, so we can specify the adaptive sampling rate, the signal sampling rate does not depend on the bandwidth of the signal; 2) a new measurement matrix, designed by kronecder in space the domain, can use less number of measurement to acquire the same psnr. 3) the proposed method is combined with adaptive sampling with KCS to improve the reconstructed video quality first, so as to avoid the high frequency data sampling, the most important thing is to avoid a lot of discarded data sampling, the data only for important data that contains information of the nature of the signal sampling, but also reduce the storage space and the spatial and temporal computational burden.

2. **Background.**

2.1. **Compressive Sensing Theory.**

Supposing the original signal$x \in \mathbb{R}^{N\times 1}$is a discrete signal, $\psi \in \mathbb{R}^{N\times N}$is an orthogonal transform base, then can be sparsely represented as transform coefficient in sparse domainψ, $x = \psi\alpha$. The sparse signal x is sensed with only M measurements as $y = \Phi x = \Phi\psi\alpha$, where$\Phi = \{\varphi_1, \varphi_2, \cdots, \varphi_M\}^T$is a $M\times N$measurement matrix. Due to $M << N$, the reconstruction of x from y is generally ill-posed[9]. However, the compressive sensing theory points out that as long as x is sparse in some domains, we can reconstruct the original signal from the obtained measurements by solving l_p-norm optimization:

$$\hat{\alpha} = \arg\min_{\alpha} \|\alpha\|_p, s.t. y = \Phi\psi\alpha \quad (1)$$

2.2. **Kronecker Compressive Sensing.** To reduce complexity caused by a large sized measurement matrix in multi-dimensional signals, Duarte et al. presented the KCS, which jointly models the measurement matrix at each signal dimension. For a 2D signal $X \in \mathbb{R}^{N\times N}$, the measurement matrix is given as $\tilde{\Phi} = \Phi_1 \otimes \Phi_2$, where $\otimes$ denotes the Kronecker product, Φ_1 and $\Phi_2 \in \mathbb{R}^{M\times N}$ represent the measurement matrices respectively for corresponding dimensions. Therefore, the CS measurement is rewritten as $Y = \tilde{\Phi}X = \Phi_1 \otimes \Phi_2 \bullet X$,

where $y = vect(Y)$ is a vectorised version of the matrix$Y$. The measurement constraint is $\|\Phi x - y\|_2^2 = \left\|\tilde{\Phi}X - Y\right\|_2^2$. Then Equation (1) can be rewritten as follows:

$$\hat{\alpha} = \arg\min_{\alpha} \|\alpha\|_p, s.t. y = vect(\tilde{\Phi}X) \quad (2)$$

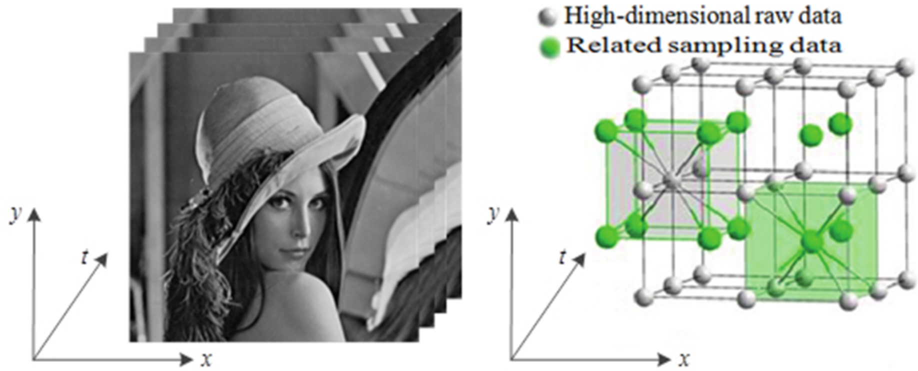

FIGURE 1. The diagram of the measurement matrix

3. **Joint Spatio-Temporal Domain for Adaptive Kronecker Compressive Sensing.** In this section, we study the current most widely used video signal processing. Different from traditional image signal sampling, the video signal on the basis of the former increase the concept of time dimension, and we can establish a three dimensional models after the video signal is generalized. As shown in figure 1, three dimensional video signal can be divided into two different dimensions: space dimension $\{(x, y)\}$ and time dimension $\{t\}$. The space dimension, expressed as each frame image of video sequences, is a two-dimentional data. It contains the surface physical properties of rules object and the background that generated the spatial redundancy. The time dimension refers to the different frame images of the different values of the same pixel location, it has a piecewise smooth feature. And each pixel is likely to be associated with stereo neighborhood pixels.

3.1. **Adaptive Sampling Based on Statistical Information of the Image Block.** In the space dimension, the video frame is first divided into small blocks, then the variance and discrete cosine transform (DCT) coefficients are obtained by calculating the statistical information of blocks, meanwhile a weighted assignment factor is introduced, finally the

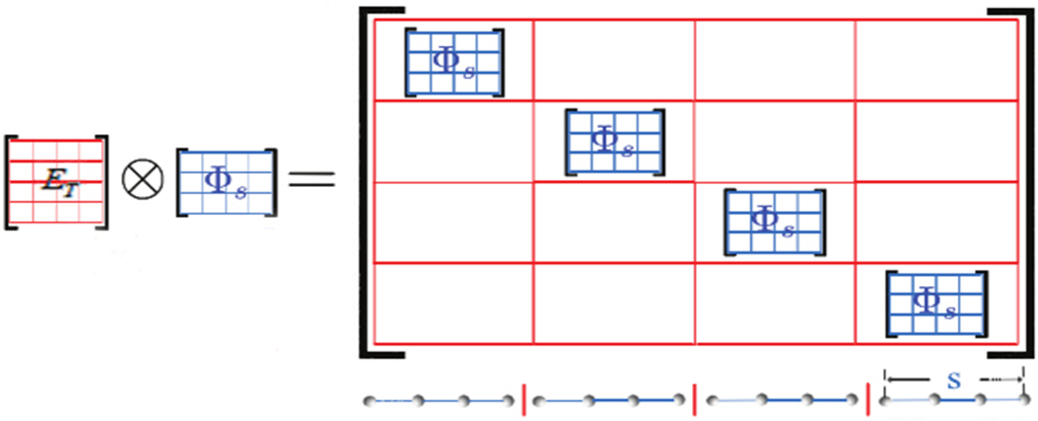

FIGURE 2. The diagram of the measurement matrix

concrete allocation method is formulated. The main process of adaptive sampling is as follow.

3.1.1. *Segmenting the Video Frame.* Before sampling, each frame image (the total pixels $N = I_c \times I_r$) is divided into small blocks with size of $B \times B$ each. And the corresponding output CS vector y_i can be written as $y_i = \Phi_{SB_i} x_i$. Where x_i represent the vectorized signal of the th block through raster scanning, Φ_{SB_i} is a $M_{SB_i} \times B^2$ measurement matrix in correspondence with the th block. And M_{SB_i} is the measurement number of image block x_i, it relies on the statistical information of the image block.

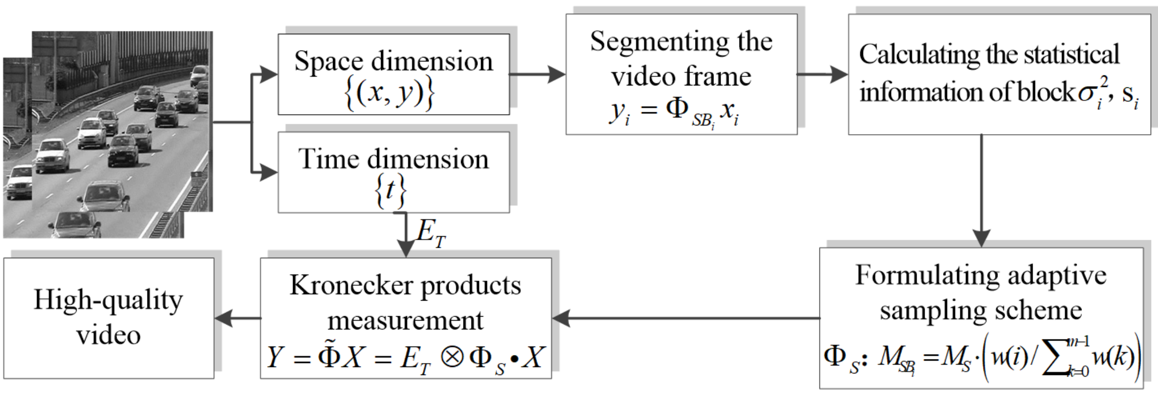

FIGURE 3. The process of the proposed method

3.1.2. *Calculating the Statistical Information of Block.* In the image pixel domain, the variance of image can be used to reflect the textures of image block. It is easy to know that a block with a bigger variance often means that it includes lots of complicated textures, thus more measurements are assigned to the image block. On the contrary, the blocks with smaller variance only need to allocate a small amount of measurement. The variance of image block can be calculated as:

$$\sigma_i^2 = \frac{1}{B \times B} \sum_{j=0}^{B^2-1} (b_j - \mu_i)^2 \tag{3}$$

Where μ_i is the mean value of image block x_i, and b_j is the th pixel in x_i. In the image transform domain, the number of significant DCT coefficients represents the specific content of the image. The DCT coefficients of texture blocks tend to have larger magnitudes, and the smooth blocks have relatively small magnitudes. In the process of adaptive sampling, threshold value T_i is set first, then the DCT coefficients are served as the important coefficients when the absolute value of the image block DCT coefficients are greater than the threshold value. This threshold value is defined as:

$$\mathrm{T}_i = \frac{1}{B \times B} \sum_{j=0}^{B^2-1} |D_j| \tag{4}$$

Where D_j is the DCT coefficient of image block x_i. Here, we use s_i represent the total number of significant DCT coefficients.

3.1.3. *Formulating Adaptive Sampling Scheme.* The specific allocation method is as follows:

Step-1: Based on the statistical information of image block in both pixel domain and transform domain, weighted assignment factor $w(i)$ is first introduced.

$$w(i) = \mathrm{c}\sigma_i^2 + (1-c)s_i \tag{5}$$

Where c is a control parameter and $0 < c < 1$.

Step-2: Developing allocation method. Assuming that the number of whole image measurement $M_S = \mu\mathrm{B}^2$, μ is a constant, allocated M_{SB_i} for image block x_i as:

$$M_{SB_i} = M_S \cdot \left(w(i) / \sum_{k=0}^{m-1} w(k) \right) \tag{6}$$

Where $m = \lfloor N/B^2 \rfloor$. All the blocks are classified, If $M_{SB_i} \geq B^2$, image block denotes as complex block , otherwise, referred to as simple block P_1. In P_0, if $M_{SB_i} > B^2$, setting $M_{SB_i} = B^2$.

Step-3: Fine-tuning the assignment method. First calculate the remaining number of sampling. And then adjust the number of sampling for .

$$\Delta = M_S - \sum_{i \in P_0} M_{SB_i} - \sum_{i \in P_1} M_{SB_i} \tag{7}$$

$$M_{SB_i} = M_{SB_i} + \Delta \cdot \frac{B^2 - M_{SB_i}}{\sum\limits_{i \in P_1} (B^2 - M_{SB_i})} \tag{8}$$

Step-4: For image blocks in P_1, put the ones with $M_{SB_i} > B^2$ into P_0, repeat step-3 to step-4 until $\Delta = 0$.

Therefore, once M_{SB_i} is determined, the corresponding measurement matrix Φ_S=diag$\{\Phi_{SB_1}, \Phi_{SB_2}, \cdots, \Phi_{SB_i}, \cdots\}$ is also determined.

3.2. **Design of measurement Matrix by Kronecker Product.** On the basis of part 3.1, the synthetic measurement matrix can be designed by using Kronecker product. Such matrices correspond to measurement processes that operate individually on a single d-section of the multidimensional signal. For video signal, synthetic measurement matrix can be expressed as

$$\tilde{\Phi} = E_T \otimes \Phi_S \tag{9}$$

where $E_T \in \mathbb{R}^{Q \times Q}$ denote identity matrix under the time domain and Φ_S is a $M_S \times I^2$ measurement matrix under the space domain, respectively. The diagram of the measurement matrix is shown in figure 2.

3.3. **Joint Spatio-Temporal Domain for video Reconstruction.** After we get the synthetic measurement matrix $\tilde{\Phi}$, the sampling scheme can be represented as:

$$Y = \tilde{\Phi}X = E_T \otimes \Phi_S \bullet X \Leftrightarrow \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_Q \end{bmatrix} = \begin{bmatrix} \Phi_S & 0 & \cdots & 0 \\ 0 & \Phi_S & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Phi_S \end{bmatrix} \bullet vec(X) \tag{10}$$

Here, X is a $I \times I \times Q$ video signal, it can be written as:

$$vec(X) = (vec(F_1)^T, vec(F_2)^T, \cdots, vec(F_Q)^T)^T \in \mathbb{R}^{I^2 Q} \tag{11}$$

where $vec(F_i)$ denote column reordering of the $i-th$ frame of X. Equation (10) shows that each measurement matrix Φ_S has a frame image that corresponds to that Φ_S, which realizes compressive sample with the corresponding space dimension signal. And the total sampling value Y is each frame mosaic image after compressive sampling.

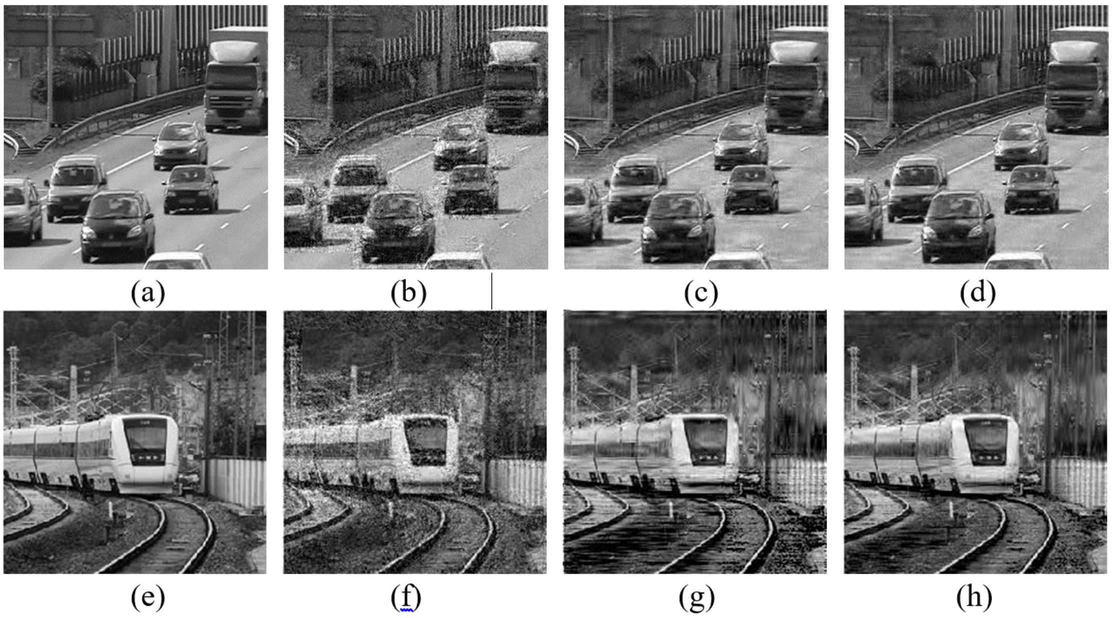

Figure 4. Visual comparison of reconstruction results for Bus and Train of different methods: (a) and (e) original image, (b) and (f) KCS, (c) and (g) HKCS, (d) and (h) proposed method (AKCS).

The optimization problem formulated as a total variation (TV)-based CS recovery [10] is given as:

$$\min_X TV(X) + \frac{\mu}{2}\left\|\tilde{\Phi}X - Y\right\|_2^2 \tag{12}$$

Where μ is a constant parameter, and $TV(X) = \|\nabla_x X\|_1 + \|\nabla_y X\|_1$, whose ∇_x and ∇_y denote the gradient operator in the horizontal and vertical directions, respectively. In conclusion, the process of the proposed method can be shown in Figure 3.

4. **Experimental Simulation and Results Analysis.** In order to verify the effectiveness and the superiority of the method proposed in this article, we use Daubechies (db8) wavelet as the sparsity base and build on the ℓ_1-norm minimization solvers from [11] for the video data. Meanwhile, we compare our method with the KCS and the HKCS methods. The simulation results were generated from two videos of spatial resolution 256×256

TABLE 1. PSNR at the sampling rate of 0.2, 0.35 and 0.5

Video	Sample rate	PSNR(dB)		
		KCS	HKCS	Proposed(AKCS)
Bus	0.2	21.2519	22.8741	23.3182
	0.35	22.7546	24.0127	25.9276
	0.5	23.4150	25.3904	27.1603
Train	0.2	21.5578	22.6958	23.7695
	0.35	23.0137	24.5269	26.1240
	0.5	23.8725	26.1573	27.5086

pixels with 100 frames, called "Bus" and "Train", respectively. In our experiment, the original video of "Bus" and "Train" are shown in Fig.4(a) and Fig.4(e). We assume that the size of block is 16×16 for each frame. The optimal value for c in (5) is between 0.5 0.7 in general, it hoped to get a better result by choosing $c = 0.65$ for all experiments. For more rigorous, all the results are averaged over 5 independent trials. Due to the limitation of space, we only demonstrate the visual quality comparison shown in Figure 4 at the sample rate of 0.35. Table 1 tabulates average PSNR values of the results for video at the sampling rate of 0.2, 0.35 and 0.5. Figure 5 shows the Running time of KCSBHKCS and proposed method. The subjective vision, PSNR and running time are employed as performance metrics.

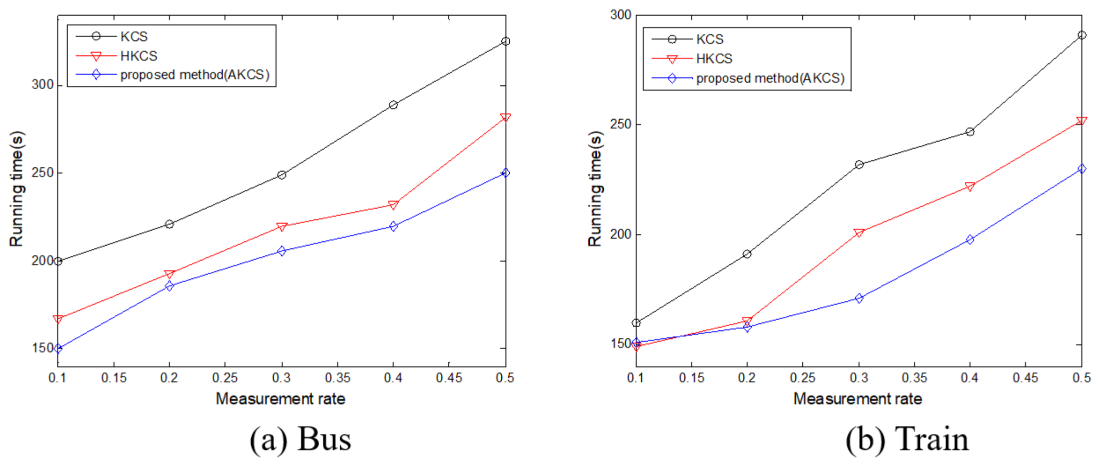

FIGURE 5. Running time of different method at the sampling rate of 0.1 0.5.

Figure 4 shows the reconstruction visual result at the rate of 0.35, it can be seen that the proposed method provides better visual result than the other methods. Besides, through analyzing Table 1 and Figure 5, we can find: 1) With the increase of sampling rate, PSNR and Running time are also increasing; 2) The proposed method (AKCS) has PSNR of more than 23 dB, it could obtain a better reconstruction performance in terms of PSNR and less Running time than the other methods in the same sampling rate; 3) Under the same sampling rate, AKCS has higher PSNR, in other words, AKCS can use less number of measurement to acquire the same PSNR. No matter from the visual effect (Figure 4), PSNR or Running time, it is noticeable that the proposed method is superior to KCS and HKCS.

5. **Conclusion and Future Works.** In this paper, based on the Kronecker compressive sensing, a novel joint spatio-temporal domain for adaptive Kronecker compressive sensing method is proposed, which consists of three stages, namely adaptive sampling based on statistical information of the image block, design of measurement matrix by Kronecker product and joint spatio-temporal domain for video reconstruction. Compared with KCS and HKCS, our method achieves greater reconstruction quality image with less number of measurement and better running efficiency. For future work, the performance of our method may be further improved by seeking more efficient ways to adaptively allocate random measurements to image blocks.

Acknowledgment. This work is supported by the Major Research and Development Plan Projects of Jiangxi province of China (NO.20161BBE50083), the Major Research Plan Projects of Jiangxi province of China (NO.20171ACE50026), S&T Projects of Transport Department of Jiangxi Province of China (NO.2016D0036), and S&T Projects of Education Department of Jiangxi Province of China(NO.GJJ150532).

REFERENCES

[1] D. L. Donoho, Compressed sensing, *Journal of IEEE Transactions on Information Theory*, vol. 52, no. 4, , pp. 1289-1306, 2006.

[2] E. J. Cand?, M. B. Wakin, An Introduction To Compressive Sampling, *Journal of IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21-30, 2008.

[3] G. Houkun, et al, Background Subtraction Based on Sparse Representation and Dictionary Learning, *Journal of East China Jiaotong University*, vol. 29, no. 1, pp. 43-47, 2012.

[4] L. Hui, et al, WMSN Video Encoding/Decoding Based on Compressed Sensing in Traffic Surveillance, *Journal of East China Jiaotong University*, vol. 30, no. 1, pp. 52-57, 2013.

[5] M. F. Duarte, R. G. Baraniuk, Kronecker product matrices for compressive sensing, *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3650-3653, 2010.

[6] M. F. Duarte, R. G. Baraniuk, Kronecker Compressive Sensing, *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 21, no. 2, pp. 494-504, 2012.

[7] C. F. Caiafa, A. Cichocki, Computing sparse representations of multidimensional signals using Kronecker bases, *Journal of Neural Computation*, vol. 25, no. 1 pp. 186-220, 2012.

[8] X.Ye, H. Xiong, Optimal Spatio-Temporal Projections with Holo-Kronecker Compressive Sensing of Video Acquisition, *IEEE Data Compression Conference*, pp. 412-412, 2012.

[9] G. Kutyniok, Theory and applications of compressed sensing, *GAMM-Mitteilungen*, vol. 36, no. 1, pp. 79-101 , 2013

[10] T. N. Canh, K. Q. Dinh, and B. Jeon, Total variation for Kronecker compressive sensing with new regularization, *in Proc. Pic. Coding Symp. (PCS)* , pp. 261-264, 2013.

[11] E. van den Berg and M. P. Friedlander, Probing the Pareto frontier for basis pursuit solutions, *Journal of SIAM J. Sci. Comput*, vol. 31, no. 2, pp. 890-912, 2008.