*ICIC Express Letters*
*Part B: Applications*
Volume **7**, Number **8**, August **2016**

ICIC International ©2016 ISSN 2185-2766
pp. 1763–1768

# WMSN VIDEO RECONSTRUCTION METHOD BASED ON MODIFIED GRADIENT PROJECTION FOR SPARSE RECONSTRUCTION

Hui Luo[1], Gong Li[1], Yi Xu[1] and Manlin Luo[2]

[1]School of Information Engineering
East China Jiaotong University
No. 808, Shuanggang East Street, Nanchang 330013, P. R. China
{ lh_jxnc; lg_jxnc }@163.com

[2]School of Physics and Electronics
Hunan University
No. 2, Lushan South Road, Changsha 410000, P. R. China
lml_jxnc@163.com

Received December 2015; accepted March 2016

Abstract. *Wireless Multimedia Sensor Networks (WMSN) technology has been widely used in many fields. However, the huge amount of data that need to be processed has a serious impact on the effectiveness of WMSN. Compressive Sensing (CS) can solve this problem effectively. Video reconstruction method becomes the core problem when it comes to introducing CS into WMSN. In this paper, we proposed a WMSN video reconstruction method based on modified Gradient Projection for Sparse Reconstruction (GPSR). Firstly, low-resolution video frames can be reconstructed via matrix inverse transformation. A new measurement matrix is proposed for the convenience of inverse transformation. Then, motion vectors among video frames can be acquired from the low-resolution video frames. At last, high-resolution video can be reconstructed by applying modified GPSR with the motion vectors as constraints. Experimental results show that the proposed algorithm can effectively improve the quality of the reconstructed video and the execution time can be reduced. In addition, a rough description of the WMSN scene can be acquired in real-time.*
**Keywords:** WMSN, Video reconstruction, Compressive sensing, Motion vector, GPSR

1. **Introduction.** WMSN is widely used in various kinds of fields. Video, as an important data form in WMSN, consumes a lot of computation resource. However, the computation resource of WMSN nodes is limited. The recently developed CS theory provides a new idea for solving the above problem [1]. CS theory points out that the observation of original signal can be acquired by sparse projection. How to construct a robust reconstruction algorithm with better effect and lower computational complexity has been the essential issue.

Recently, many classical signal reconstruction algorithms have been proposed, such as Matching Pursuits (MP) [2,3] and Subspace Pursuit (SP) [4]. These algorithms have low computational complexity, but the reconstruction effect is not satisfactory. For more superior effect, convex optimization algorithm is proposed, such as Total Variation (TV) [5] and GPSR [6]. In [7], a novel image reconstruction algorithm based on intra prediction and total variation model is proposed for improving the reconstruction performance, but it takes a lot of time to reconstruct the signal. Besides the most popular methods we mentioned above, the iterative reweighting algorithm becomes a more efficient way to enhance the reconstruction quality [8]. However, that each weighting factor must be iteratively updated makes the algorithm complicated. What is worse, we cannot get the signal content intuitively before the reconstruction has been completed, which is negative to WMSN video applications.

Under the research of CS theory, we find that the video reconstruction can be divided into two stages. First, the preliminary results can be achieved by applying inverse transformation to CS measurements. Then, it may offer more information for subsequent processing to further reconstruct signal with high quality. Once the two-stage reconstruction algorithm is applied, a rough description of the WMSN scene can be acquired on the spot, and the computational complexity can be greatly reduced, which guarantees real-time monitoring of WMSN. GPSR algorithm can achieve accurate reconstruction with high probability compared to other algorithms [9], but it takes a lot of time to reconstruct the signal. According to the above analysis, this paper puts forward a two-stage reconstruction algorithm based on Modified GPSR for WMSN video.

The major contribution of this paper is twofold: 1) We design a measurement matrix which can realize inverse transformation quickly, so that a rough description of the WMSN scene can be acquired in a short time; 2) Our method is among the first to impose motion vector as constraints in the GPSR iteration process. The CS theory is introduced in Section 2. WMSN video reconstruction method based on Modified GPSR is presented in Section 3. Section 4 reports experimental results and discussions. Conclusions are given in Section 5.

2. **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, and then $x$ can be expressed as $x = \psi\alpha$, where $\alpha$ is the transform coefficient in sparse domain $\psi = [\psi_1, \psi_2, \ldots, \psi_N]$, and $\psi_i$ is a column vector with $N$ elements.

We can implement non-adaptive linear observation by selecting the appropriate measurement matrix to obtain measurements $y$:

$$y = \Phi x = \Phi\psi\alpha \tag{1}$$

where $\Phi = \{\phi_1, \phi_2, \ldots, \phi_M\}^T$ is a $M \times N$ measurement matrix. Due to $M \ll N$, the reconstruction of $x$ from $y$ is generally ill-posed [10]. 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 $l_1$-norm optimization:

$$\min_{\alpha} \|\alpha\|_1, \text{ s.t. } y = \Phi\psi\alpha \tag{2}$$

3. **WMSN Video Reconstruction Based on Modified GPSR.** In order to apply the above-mentioned theory to actual video system, researchers in Rice University designed Single Pixel Camera (SPC) [11], which largely decreases the amount of data that need to be acquired. If we can employ SPC as video capture unit in WMSN, the amount of data can be reduced and the lifecycle can be extended in the network. As a research object, this paper mainly concentrates on the video captured by SPC in WMSN.

3.1. **Video data acquisition model of WMSN.** In WMSN, the compressive measurements that SPC captured at the sample instant $t$ can be modeled as $y_t = \langle \phi_t, x_t \rangle + e_t$, where $\phi_t$ is the measurement matrix, $x_t$ is the scene at sample instant $t$, and $e_t$ represents measurement noise. Assuming that there are $A = a \times a$ pixels in each frame, then $x_t \in \mathbb{R}^{A\times 1}$.

Suppose that we rewrite our time-varying scene $x_t$ for a window of $Q$ consecutive sample instants as $x_t = S + \Delta x_t$. Here, $S$ is the static component (assumed to be invariant for the considered window of $Q$ samples), and $\Delta x_t = x_t - S$ is the variant component. We use notation $y_{1:Q}$ to represent the matrix made of all measurements during the window

$Q$, and define $z_t = \langle \phi_t, \Delta x_t \rangle$, which represents the temporal-approximation error caused by assuming the scene remains static for $Q$ samples, then

$$y_{1:Q} = [y_1, y_2, \ldots, y_3]^T = \Phi S + z_{1:Q} + e_{1:Q} \tag{3}$$

where $e_{1:Q}$ is measurement error, and $\Phi \in \mathbb{R}^{Q \times A}$ is measurement matrix, whose $t$-th column is measurement vector $\phi_t$.

The following will discuss the errors of spatial down-sampling to time-invariant component $S$. We adopt $S_L$ ($S_L \in \mathbb{R}^{A_L \times 1}$) to represent the spatial down-sampling of $S$, where $A_L = a_L \times a_L$ ($A_L < A$). We define $H$ and $L$ as up-sampling and down-sampling operators, respectively. According to the definition, they satisfy $H \in \mathbb{R}^{A \times A_L}$ and $L \in \mathbb{R}^{A_L \times A}$. Then Equation (3) can be rewritten as follows:

$$y_{1:Q} = \Phi H S_L + \Phi(I - HL)S + z_{1:Q} + e_{1:Q} \tag{4}$$

Here, the first item can be seen as the measurements of static image, and the others can be seen as measurement error.

3.2. **Reconstructing the low-resolution video.** In order to analyze the trade-off that arises from the static scene hypothesis and down-sampling procedure, we consider the scenario where the effective matrix $\Phi H$ is of dimension $Q \times A_L$ ($Q \geq A_L$). If the matrix $\Phi H$ has full rank, then we can obtain an estimate $\hat{S}_L$ of the low-resolution static scene as:

$$\hat{S}_L = (\Phi H)^{-1} y_{1:Q} = S_L + (\Phi H)^{-1}(\Phi(I - HL)S + z_{1:Q} + e_{1:Q}) \tag{5}$$

where $(\cdot)^{-1}$ denotes the pseudo inverse. From Equation (5), we can reconstruct the low-resolution video under the hypothesis that the scene remains unchanged in the interval $Q$.

3.3. **Design of measurement matrix.** We note that there is a special requirement when obtain $\hat{S}_L$: the matrix $\Phi H$ is reversible. However, the measurement matrix commonly used in CS, such as i.i.d. Gaussian matrices, right multiplying them with $H$ often results in an ill-conditioned matrix. In order to improve the accuracy of the low-resolution video frames, we design a novel measurement matrix $\Phi$. It meets the condition that $M = \Phi H$, where $M$ is a $Q \times Q$ dimension Hadamard matrix. Hadamard matrix is chosen as the basis of transformation because of its low complexity of inverse transformation. The measurement matrix we designed can be represented as the following form:

$$\Phi = ML + \Gamma \tag{6}$$

where $\Gamma$ is an auxiliary matrix that obeys the following constraints: 1) The choice of auxiliary matrix should guarantee $\Phi \in \{+1, -1\}$; 2) The measurement matrix $\Phi$ should satisfy the Restricted Isometry Property (RIP) [12]; 3) $\Gamma$ should be chosen so that $\Gamma H = 0$.

3.4. **Extracting motion vector based on optical flow.** On the basis of an appropriate measurement matrix, we can easily reconstruct the low-resolution video frames, and then compute motion vector between consecutive frames based on optical flow. Assuming that $\hat{S}_L^i$ and $\hat{S}_L^j$ are the $i$-th and $j$-th low-resolution video frames, respectively, and they are consecutive. $\hat{S}^i = H\hat{S}_L^i$ and $\hat{S}^j = H\hat{S}_L^j$ are the up-sampling results of the above low-resolution video frames. Then, optical flow between the two frames can be written as:

$$\hat{S}^i(x, y) = \hat{S}^j(x + u_{x,y}, y + v_{x,y}) \tag{7}$$

where $\hat{S}^i(x, y)$ denotes the pixel $(x, y)$ in the $a \times a$ plane of $\hat{S}^i$. $u_{x,y}$ and $v_{x,y}$ correspond to the translation of the pixel $(x, y)$ between frame $i$ and $j$.

3.5. **Reconstructing the high-resolution video.** After we get the motion vector, we can transform the non differentiable $l_1$-norm minimization problem to a differentiable objective function, and then reconstruct the high-resolution video by the gradient projection iteration. In the process of reconstruction, we consider the following two constraints: the consistency with the acquired CS measurements, and the estimated optical flow constraints between consecutive frames. Together, we arrive at the flowing convex optimization problem:

$$\min_{\alpha} \|\alpha\|_1 \,, \text{ s.t. } y = \Phi\psi\alpha;\ \alpha_i(x,y) = \alpha_j(x+u_x, y+v_y) \tag{8}$$

The above problem can be solved by the standard convex optimization techniques. Setting a threshold, Equation (8) can be converted into:

$$\min_{\alpha} \|\alpha\|_1 \,, \text{ s.t. } \|y - \Phi\psi\alpha\|_2 \le \sigma_1;\ \|\alpha_i(x,y) - \alpha_j(x+u_x, y+v_y)\|_2 \le \sigma_2 \tag{9}$$

The parameters $\sigma_1$ and $\sigma_2$ are indicative of the measurement noise and the inaccuracies in the brightness constancy, respectively. In conclusion, the process of the proposed method can be shown in Figure 1.

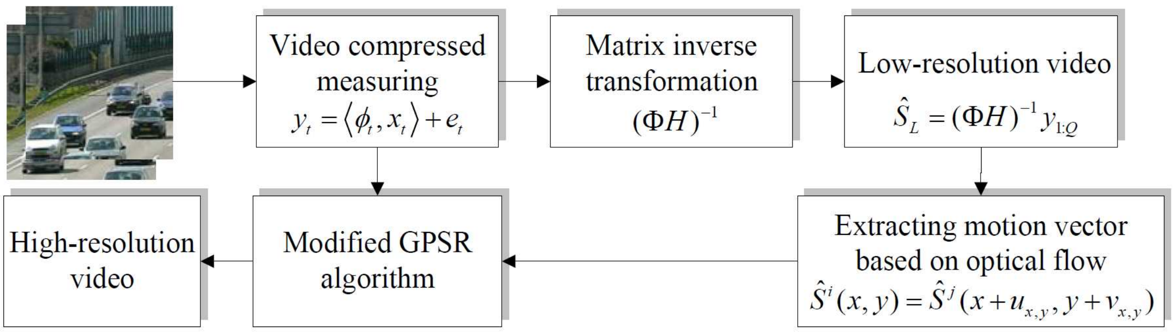


FIGURE 1. The process of the proposed method

4. **Experimental Results and Discussions.** This section describes some experiments to testify the performance of the method mentioned above. The simulation results were generated from two videos of 256 × 256 pixels with 256 frames. One of them is an "Intelligent vehicle video", and the other is "Traffic monitoring video". The low-resolution video has a spatial resolution of 64 × 64 pixels. A Gaussian white noise with a standard deviation of 10 was added to the compressive measurements. The maximum number of iterations is set to 100. For comparison, the original GPSR algorithm is used to reconstruct the same video with the same noise level.

In our experiment, the two videos are averagely divided into three parts, respectively. One frame is selected from each part for display. The original video of "Intelligent vehicle video" and "Traffic monitoring video" are shown in Figure 2(a) and Figure 3(a), respectively. The reconstruction results of "Intelligent vehicle video" are shown in Figure 2. The reconstruction results of "Traffic monitoring video" are shown in Figure 3. Table 1 tabulates the running time of low-resolution reconstruction, high-resolution reconstruction and GPSR. For more rigorous effects, each video has been processed three times on the same computer.

Through analyzing Figure 2 and Figure 3, we can find: 1) The low-resolution video, an intermediate of our method, has PSNR of no less than 20dB. It can provide a rough description of the scene; 2) The proposed method achieves better reconstruction performance in terms of PSNR and video details to the comparison. From Table 1, it is clear that the low-resolution video can be reconstructed in no more than 7 seconds with the proposed algorithm. As for the high-resolution reconstruction, the proposed algorithm

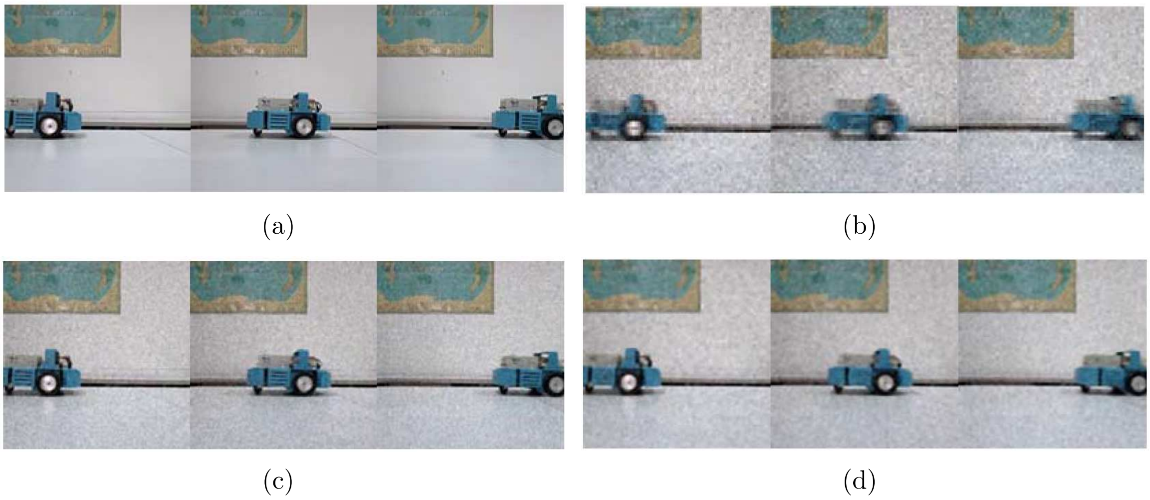


Figure 2. Reconstruction results of "Intelligent vehicle video": (a) original video, (b) proposed method (low-resolution, PSNR: 21.12dB), (c) proposed method (high-resolution, PSNR: 28.25dB), (d) GPSR (PSNR: 25.16dB)

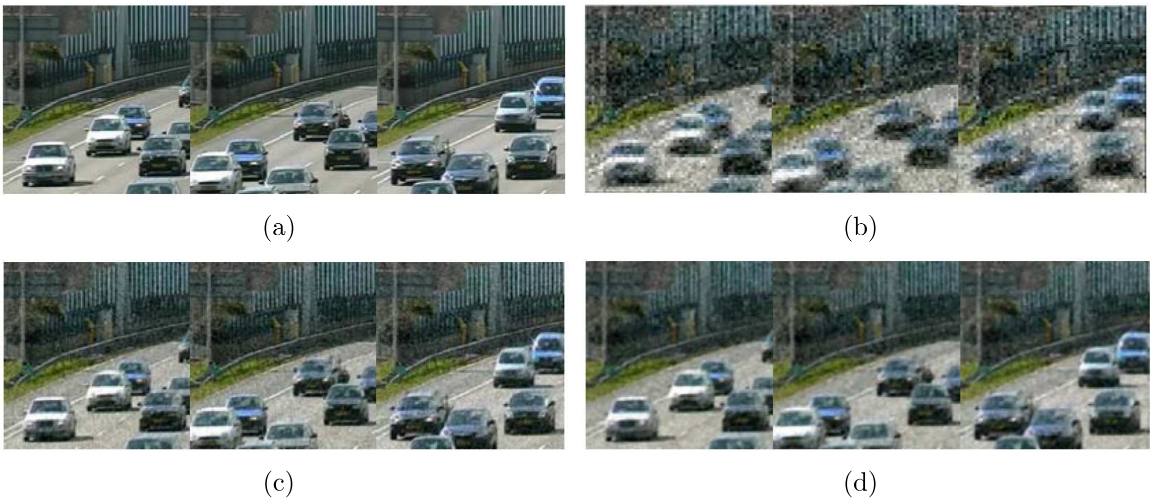


Figure 3. Reconstruction results of "Traffic monitoring video": (a) original video, (b) proposed method (low-resolution PSNR: 20.38dB), (c) proposed method (high-resolution PSNR: 26.51dB), (d) GPSR (PSNR: 25.38dB)

Table 1. Comparison of running time

| Video | Running time (in seconds) | | |
|---|---|---|---|
| | Low-resolution | High-resolution | GPSR |
| Intelligent vehicle | 5.23 | 7538.64 | 7603.58 |
| | 4.89 | 7536.29 | 7605.47 |
| | 5.30 | 7584.22 | 7612.34 |
| Traffic monitoring | 6.47 | 8225.35 | 8548.78 |
| | 6.21 | 8212.86 | 8553.45 |
| | 6.59 | 8238.57 | 8498.43 |

reduces execution time to some extent compared to GPSR. No matter from the visual effect of PSNR or running time, it is noticeable that the proposed algorithm is superior to GPSR. Therefore, the proposed method is more suitable for the WMSN video reconstruction.

5. **Conclusions.** This paper proposes an effective WMSN video reconstruction method. This method includes a novel measurement matrix that enables the efficient reconstruction of a low-resolution video. Meanwhile, the high-resolution video reconstruction can be achieved by using convex optimization. Comparing with the original GPSR algorithm, our method achieves better reconstruction quality, both in PSNR performance and in visual result. In addition, a rough description of the WMSN scene can be acquired in real-time and the execution time can be reduced. For future work, the performance of our method may be further improved by using multi-frame optical-flow estimation.

**Acknowledgement.** This paper is supported by the National Natural Science Foundation of China (No. 61261040).

## REFERENCES

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

[2] T. T. Do, L. Gan and N. Nguyen, Sparsity adaptive matching pursuit algorithm for practical compressed sensing, *IEEE Conference on Signals, Systems and Computers*, pp.581-587, 2008.

[3] J. A. Tropp and A. C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, *IEEE Transactions on Information Theory*, vol.53, no.12, pp.4655-4666, 2007.

[4] W. Dai and O. Milenkovic, Subspace pursuit for compressive sensing signal reconstruction, *IEEE Transactions on Information Theory*, vol.55, no.5, pp.2230-2249, 2009.

[5] Y. Liu, Y. Li, J. Lou, W. Wang and M. Zhang, Omni-total variation algorithm for the restoration of all-focused catadioptric image, *Optik*, vol.125, no.14, pp.3685-3689, 2014.

[6] Z. Liu, Z. Wei and W. Sun, An iteratively approximated gradient projection algorithm for sparse signal reconstruction, *Applied Mathematics and Computation*, vol.228, no.1, pp.454-462, 2014.

[7] Y. Song, W. Cao, Y. Shen and G. Yang, Compressed sensing image reconstruction using intra prediction, *Neurocomputing*, vol.151, no.3, pp.1171-1179, 2015.

[8] S. Zhu, B. Zeng and M. Gabbouj, Adaptive sampling for compressed sensing based image compression, *Journal of Visual Communication and Image Representation*, vol.30, pp.94-105, 2015.

[9] J. K. Liu and S. J. Li, A projection method for convex constrained monotone nonlinear equations with applications, *Computers and Mathematics with Applications*, vol.70, no.10, pp.2442-2453, 2015.

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

[11] D. Takhar, J. N. Laska, M. Wakin et al., A new compressive imaging camera architecture using optical domain compression, *Symposium on Electronic Imaging: Computational Imaging*, pp.43-52, 2006.

[12] S. Qaisar, R. M. Bilal, W. Iqbal et al., Compressive sensing: From theory to applications, a survey, *Journal of Communications and Networks*, vol.15, no.5, pp.443-456, 2013.