A Dual Structured-Sparsity Model for Compressive-Sensed Video Reconstruction

Chen Zhao[#], Jian Zhang[#], Siwei Ma[#*], Ruiqin Xiong[#], Wen Gao[#*]
[#]*Institute of Digital Media & Cooperative Medianet Innovation Center, Peking University, Beijing, China*
[*]*Peking University Shenzhen Graduate School, Shenzhen, China*
{zhaochen, jian.zhang, swma, rqxiong, wgao}@pku.edu.cn

Abstract— **The compressive sensing theory indicates that robust reconstruction of signals can be obtained from far fewer measurements than those required by the Nyquist theorem. Thus, it has great potential in video acquisition and processing in that it can tremendously save the complex compression required by traditional video coding standards. In this paper, we consider reconstruction of compressive-sensed videos and propose a novel structured-sparsity model with a dual prediction strategy. This structured-sparsity model goes beyond simple sparsity and characterizes the intrinsic structure within the transform coefficients. Also, it exploits the sparsity of the residual between the current patch and its prediction. The prediction process is comprised of a dual strategy, which integrates the advantages of the ambient pixel domain and the measurement domain. In addition, an effective optimization method is designed for solving the formulated problem derived from the model. Experiments demonstrate that the proposed algorithm outperforms the state-of-the art methods for compressive-sensed video reconstruction in both subjective and objective quality.**

***Index Terms*—Compressive sensing, video reconstruction, structured sparsity, split Bregman iteration, dual domain**

I. INTRODUCTION

Compressive sensing (CS) is an emerging technology that suggests an intriguing new paradigm for signal acquisition and compression. It makes it possible for sparse signals to be sampled at a sub-Nyquist rate while ensuring their robust reconstruction under certain conditions [1]-[2]. The video signal, which has massive spatial and temporal redundancies, has good sparse representation in some transform domain. Thus, CS is potentially applicable to acquiring and processing the video signal. Compared to conventional video processing, which requires sophisticated predictive coding after video acquisition [1], [4], video compressive sensing integrates sampling and compression and tremendously reduces the computational resource consumption in the encoder.

The sampling process is usually in the unit of a block in order to reduce the memory burden in the encoder [5]. With very few measurements for each block, the burden of acquiring a high quality signal for compressive-sensed videos is shifted to the decoder side. Thus effective sparse representations of videos and corresponding recovery algorithms are desired to reconstruct the videos.

A straightforward approach is to independently reconstruct each video frame using a 2D sparsity transform (e.g., 2D DCT, 2D DWT [6]). This approach can serve as a baseline against which other techniques are compared. As well as spatial redundancies, video signals also have plenty of temporal correlations, which can thus be exploited to enhance the sparsity representation. Wakin et al. propose to jointly reconstruct consecutive frames using a 3D wavelet transform in [7].

A further effort is to take motion between frames into consideration. A variety of works propose to perform motion compensation and estimation (MC/ME) at the decoder (e.g., [8]-[9]). In these works, a block in the current frame can be predicted by one or more blocks in the neighboring frames for further modelling. The prediction accuracy of the blocks is essential for building an accurate temporal model. Unlike standard video coding, however, original values of video frames are not explicitly available at the decoder to perform motion estimation. Only random CS measurements of the underlying frames are present. A typical solution is to first obtain an initial recovery of each frame using some existing method, and then do motion estimation on these initial recoveries for further enhancement. Unfortunately, this strategy can only find a noisy prediction for blocks in the current frame, since the initial recovery usually has low fidelity to the actual frame values.

Based on the MC/ME idea, E. W. Tramel et al. propose to estimate the motion-compensated frame in the measurement domain [10]. Instead of directly performing block matching in the ambient pixel domain, it calculates the matching error in the measurement domain for finding the optimal predictive block. No initial recovery is required, and thus better prediction can be achieved in general. By this means, however, the block partitioning for reconstruction has to be coincide with that of sampling. Consequently, this restricts quality enhancement that can be achieved otherwise by using a smarter way of block partitioning at the decoder, e.g., a smaller block size, overlapping blocks.

As discussed above, either the ambient pixel domain or the measurement domain is an optimal way for prediction. In this paper, we integrate both of their advantages and propose a dual structured-sparsity model for CS video reconstruction. We model the structured-sparsity of each block using the predictions obtained in both domains, and characterize both considerations in a unified statistical manner. Additionally, we propose an effective algorithm to solve the formulated optimization problem based on the split Bregman iteration algorithm.

The remainder of this paper is organized as follows. Section II introduces the background of compressive sensing. Section III details the proposed model for CS video reconstruction. Section IV presents the proposed method for solving the optimization problem. Section V demonstrates the experimental results of the proposed algorithm and Section VI concludes this paper.

II. Background

A. The Compressive Sensing Theory

The compressive sensing theory states that a signal with sparse representation in a transform domain can be accurately recovered from a small number of measurements [11]. Concretely, let us consider a signal of a finite dimension $\boldsymbol{x} \in \mathbb{R}^N$. The acquisition process of CS is expressed as

$$\boldsymbol{y} = \boldsymbol{\Phi x}, \tag{1}$$

where $\boldsymbol{\Phi}$ is an $M \times N$ random projection matrix and $\boldsymbol{y} \in \mathbb{R}^M$ represents the acquired linear measurements. *M/N* is called the sampling rate or subrate, which is typically very small. Thus the inverse problem is highly ill-posed.

We represent the signal $\boldsymbol{x}$ in an appropriate $N \times N$ basis $\boldsymbol{\Psi}$, i.e., $\boldsymbol{x} = \boldsymbol{\Psi\alpha}$. If at most $K<<N$ entries of the coefficients $\boldsymbol{\alpha}$ are nonzero, we say $\boldsymbol{x}$ is *K*-sparse in the domain $\boldsymbol{\Psi}$. Although many natural signals are not strictly sparse, they can be approximated as such; we call them compressible signals. According to the CS theory, sparse and compressible signals can be reconstructed by solving the following minimization problem:

$$\min_{x} \|\boldsymbol{\alpha}\|_0, \quad \text{s.t. } \boldsymbol{y} = \boldsymbol{\Phi x}. \tag{2}$$

Considering that the above ℓ_0 problem is generally NP-hard, it is usually converted to an ℓ_1 unconstrained problem for easier optimization, i.e.,

$$\min_{x} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{\Phi x}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1. \tag{3}$$

B. Overview of CS-based Video Sampling and Reconstruction

For the sake of encoder simplicity, we sample the video signal frame by frame in the unit of a block. We divide the video sequence into groups of pictures (GOP), each of which contains a key frame and several non-key frames. A key frame has a high sampling rate and a non-key frame has a relatively low sampling rate.

For reconstruction, each frame is first recovered independently using a 2D sparsity model to get an initial value. Then we divide each non-key frame into overlapped patches (patch overlapping is effective to avoid blocking artifacts) and apply our proposed dual structured-sparsity model to each patch. Our proposed solving method is utilized for optimization of the problem derived from the model. Hence, non-key frames with enhanced quality are finally reconstructed.

III. Proposed Dual Structured-sparsity Model

This section discusses the modelling process of each overlapped patch in a non-key frame. We present our structured-sparsity model in the first subsection, the dual-domain prediction in the second subsection, and formulate the reconstruction algorithm derived from this model in the third subsection.

A. Structured-Sparsity Model

Rather than enforcing simple sparsity of each patch's transform coefficients, we consider the intrinsic structure existing within these coefficients by utilizing the temporal correlations between frames [11].

Owing to the motion coherence of consecutive video frames [13], for the current patch, there exist similar patches in the neighboring key frames. We utilize these similar patches to generate predictions for the current patch (see the next subsection for the prediction process), and consider its residual sparsity in the transform domain. It is known that the residual is typically more compressible and thus exhibits higher sparsity levels than the original signal [5]. This residual sparsity is written as the following equation

$$\left\|\mathbf{W}\left(\boldsymbol{\Psi}^T\left(\boldsymbol{x} - \boldsymbol{x}_{pred}\right)\right)\right\|_1, \tag{4}$$

where $\boldsymbol{x}$ contains all the pixel values of the patch, $\boldsymbol{x}_{pred}$ represents a predictive patch obtained through our prediction strategy. Their difference is the residual. $\boldsymbol{\Psi}$ is a transform basis to sparsify the patch, and $\boldsymbol{\Psi}^T$ is its transpose. $\mathbf{W}$ is a weighting matrix, that consists of the weights $w_1, w_2, \ldots w_i, \ldots$ on the diagonal and zeroes elsewhere.

Eq. (4) indicates a structured-sparsity model, that not only specifies a more realistic fact of a patch by calculating the residual, but also discriminatively weights different coefficients to exhibit their structure.

B. Dual-Domain Prediction

We design a dual domain strategy for achieving the prediction $\boldsymbol{x}_{pred}$ for the patch: the pixel-domain prediction (P prediction) and the measurement-domain prediction (M prediction).

We do patch matching in the ambient pixel domain for P prediction. The patch size and patch partitioning can be set in any way that is beneficial for reconstruction and not influenced by the encoder. Thus P prediction is quite flexible. We find similar patches in the neighbouring key frames by directly comparing the sum of square errors (SSE) of the patch pixel values.

$$\boldsymbol{x}_{simP} = \arg\min_{\boldsymbol{x}_i \in P(\boldsymbol{x}_{key})} \left\|\boldsymbol{x}_i - \hat{\boldsymbol{x}}^0\right\|_2^2. \tag{5}$$

Eq. (5) shown above formulates the process of finding a similar patch. $P(\boldsymbol{x}_{key})$ is the set of all possible patches in the reference key frames, and $\boldsymbol{x}_i$ is a patch in the set. Note that since the original value of the current patch is not available, we use its initial recovery $\hat{\boldsymbol{x}}^0$ for this calculation. $\boldsymbol{x}_{simP}$ is the found patch that has the smallest SSE with the initial recovery. Once we have obtained such a patch as $\boldsymbol{x}_{simP}$, we remove it from the set $P(\boldsymbol{x}_{key})$ and calculate Eq. (5) again to find the second closest patch. By repeating this process, we find *C* similar patches and then compute their linear combination as a prediction $\boldsymbol{x}_{pred,1}$ as follows

$$\boldsymbol{x}_{pred,1} = \sum_{1 \le i \le C} \alpha_i \boldsymbol{x}_{simP,i}, \tag{6}$$

where $\boldsymbol{x}_{simP,i}$ represents the *i-th* similar patch, and α_i is its weight, which is determined by the SSE between the similar patch and the current patch as formulated in the following equation given a constant h

$$\alpha_i = \frac{\exp\left(-\left\|\boldsymbol{x}_i - \hat{\boldsymbol{x}}^0\right\|_2^2 / h\right)}{\sum_{1\le i\le C} \exp\left(-\left\|\boldsymbol{x}_i - \hat{\boldsymbol{x}}^0\right\|_2^2 / h\right)}. \quad (7)$$

Our second scheme is doing patch prediction in the measurement domain. M prediction is not influenced by the noisy initial recovery [10]. Compared to the P prediction, we recast the optimization into the measurement domain by applying the random projection matrix $\boldsymbol{\Phi}$ to the block difference as follows

$$\begin{aligned}\boldsymbol{\beta} &= \arg\min_{\boldsymbol{\beta}} \left\|\boldsymbol{\Phi}\left(\mathbf{X}\boldsymbol{\beta} - \boldsymbol{x}_{block}\right)\right\|_2^2 \\ &= \arg\min_{\boldsymbol{\beta}} \left\|\boldsymbol{\Phi}\mathbf{X}\boldsymbol{\beta} - \boldsymbol{y}\right\|_2^2.\end{aligned} \quad (8)$$

Note that in Eq. (8), the optimization unit is not a patch, but a block $\boldsymbol{x}_{block}$ that is corresponding with the measurements $\boldsymbol{y}$ acquired from the block-based CS sampling. $\mathbf{X}$ contains each possible block in the reference key frames as a column, and $\boldsymbol{\beta}$ is a column vector consisting of the weight for each block $\beta_1, \beta_2, \ldots \beta_i, \ldots$. We do this optimization as in [10] and obtain the optimal value for $\boldsymbol{\beta}$. Then we select C blocks that have the largest weights as the similar blocks of $\boldsymbol{x}_{block}$, and calculate a predictive block as follows

$$\boldsymbol{x}_{pred,\text{block}} = \frac{1}{\sum_{1\le i\le C} \beta_i} \sum_{1\le i\le C} \beta_i \boldsymbol{x}_{simB,i}, \quad (9)$$

where $\boldsymbol{x}_{simB,i}$ represents the *i-th* similar block, and β_i is its corresponding weight.

Here, in order to take advantage of patch overlapping also for the M prediction, we design a scheme to transfer predictive blocks to predictive patches. For each block in the current frame, we calculate a predictive block through Eq. (8) and Eq. (9), and thus for the entire frame, we can assemble all the predictive blocks to get a predictive picture. In the way that the current frame is partitioned, we partition the predictive picture into overlapped predictive patches. Thus, each predictive patch corresponds to a patch in the current frame. We denote the predictive patch obtained by the M prediction as $\boldsymbol{x}_{pred,2}$.

C. Formulation of the Reconstruction Problem

We incorporate our structured-sparsity model with the dual-domain prediction into the CS reconstruction paradigm and obtain the following optimization problem

$$\hat{\boldsymbol{x}}_f = \arg\min_{\boldsymbol{x}_f} \frac{1}{2}\left\|\boldsymbol{y}_f - \boldsymbol{\Phi}_f \boldsymbol{x}_f\right\|_2^2 + \lambda_1 \sum_{\boldsymbol{x}\in OP(\boldsymbol{x}_f)} \left\|\mathbf{W}_1 \boldsymbol{\Psi}^T\left(\boldsymbol{x} - \boldsymbol{x}_{pred,1}\right)\right\|_1 + \lambda_2 \sum_{\boldsymbol{x}\in OP(\boldsymbol{x}_f)} \left\|\mathbf{W}_2 \boldsymbol{\Psi}^T\left(\boldsymbol{x} - \boldsymbol{x}_{pred,2}\right)\right\|_1, \quad (10)$$

in which $\boldsymbol{x}_f$ denotes the entire original frame, $\boldsymbol{y}_f$ represents the corresponding measurements acquired by projecting $\boldsymbol{x}_f$ onto the random matrix $\boldsymbol{\Phi}_f$. $OP(\boldsymbol{x}_f)$ is a set of all the overlapped patches in the current frame $\boldsymbol{x}_f$. The entries of the weighting matrices $\mathbf{W}_1$ and $\mathbf{W}_2$ are calculated as inversely proportional to the variances of the respective predictive patches. λ_1 and λ_2 are parameters that balance the contributions of the two model terms.

IV. Optimization

We design a method to solve the optimization problem of Eq. (10) based on the split Bregman iteration algorithm (SBI) [14].

First, we convert Eq. (10) to a constrained problem by introducing two more variables $\boldsymbol{z}$ and $\boldsymbol{v}$ [15] as follows

$$\begin{aligned}\left(\hat{\boldsymbol{x}}_f, \hat{\boldsymbol{z}}_f, \hat{\boldsymbol{v}}_f\right) = \arg\min_{\boldsymbol{x}_f, \boldsymbol{z}_f, \boldsymbol{v}_f} \frac{1}{2}\left\|\boldsymbol{y}_f - \boldsymbol{\Phi}_f \boldsymbol{x}_f\right\|_2^2 + \lambda_1 \sum_{\boldsymbol{x}\in OP(\boldsymbol{x}_f)} \left\|\mathbf{W}_1 \boldsymbol{\Psi}^T\left(\boldsymbol{z} - \boldsymbol{x}_{pred,1}\right)\right\|_1 \\ + \lambda_2 \sum_{\boldsymbol{x}\in OP(\boldsymbol{x}_f)} \left\|\mathbf{W}_2 \boldsymbol{\Psi}^T\left(\boldsymbol{v} - \boldsymbol{x}_{pred,2}\right)\right\|_1, \quad \text{s.t.}\ \boldsymbol{x}_f = \boldsymbol{z}_f, \boldsymbol{x}_f = \boldsymbol{v}_f.\end{aligned} \quad (11)$$

Then we apply the Bregman algorithm to Eq. (11) and arrive at the following five iterative steps:

$$\boldsymbol{x}_f^{(j+1)} = \arg\min_{\boldsymbol{x}_f} \frac{1}{2}\left\|\boldsymbol{y}_f - \boldsymbol{\Phi}_f \boldsymbol{x}_f\right\|_2^2 + \eta_1 \left\|\boldsymbol{x}_f - \boldsymbol{z}_f^{(j)} - \boldsymbol{b}^{(j)}\right\|_2^2 + \eta_2 \left\|\boldsymbol{x}_f - \boldsymbol{v}_f^{(j)} - \boldsymbol{c}^{(j)}\right\|_2^2, \quad (12)$$

$$\boldsymbol{z}_f^{(j+1)} = \arg\min_{\boldsymbol{z}_f} \lambda_1 \sum_{\boldsymbol{x}\in OP(\boldsymbol{x}_f)} \left\|\mathbf{W}_1 \boldsymbol{\Psi}^T\left(\boldsymbol{z} - \boldsymbol{x}_{pred,1}\right)\right\|_1 + \eta_1 \left\|\boldsymbol{x}_f^{(j+1)} - \boldsymbol{z}_f - \boldsymbol{b}^{(j)}\right\|_2^2, \quad (13)$$

$$\boldsymbol{v}_f^{(j+1)} = \arg\min_{\boldsymbol{v}_f} \lambda_2 \sum_{\boldsymbol{x}\in OP(\boldsymbol{x}_f)} \left\|\mathbf{W}_2 \boldsymbol{\Psi}^T\left(\boldsymbol{v} - \boldsymbol{x}_{pred,2}\right)\right\|_1 + \eta_2 \left\|\boldsymbol{x}_f^{(j+1)} - \boldsymbol{v}_f - \boldsymbol{c}^{(j)}\right\|_2^2, \quad (14)$$

$$\boldsymbol{b}^{(j+1)} = \boldsymbol{b}^{(j)} - \left(\boldsymbol{x}^{(j+1)} - \boldsymbol{z}^{(j+1)}\right), \quad (15)$$

$$\boldsymbol{c}^{(j+1)} = \boldsymbol{c}^{(j)} - \left(\boldsymbol{x}^{(j+1)} - \boldsymbol{v}^{(j+1)}\right). \quad (16)$$

In the above equations, the superscript j denotes the iteration number. We can see that the problem of Eq. (10) is split into three subproblems: the $\boldsymbol{x}$ problem, the $\boldsymbol{z}$ problem, and the $\boldsymbol{v}$ problem. The $\boldsymbol{x}$ problem is a strictly convex function, which can be directly solved by derivative calculation. For solving the $\boldsymbol{z}$ problem and the $\boldsymbol{v}$ problem, we utilize the strategy in [16] and can also achieve concrete solutions for them.

V. Experimental Results

In this section, we present the experimental results of our algorithm. We show the results of four standard test sequences of the 'CIF' size (352×288): 'Akiyo', 'Foreman', 'Carphone', and 'Paris'. Each video sequence is sampled frame by frame at a block level using a Gaussian random projection matrix. The block size is set 32×32. We compare our proposed algorithm to four representative CS reconstruction methods in the literature: 2D DWT [6], 3D DCT, MC/ME [9], and Video MH [10]. 2D DWT is a baseline to compare the performances of all the CS video methods.

In the implementation of our algorithm, the transform matrix $\boldsymbol{\Psi}$ is made the DCT basis. We use the method of video MH to initialize each frame. The two key frames before and after the current non-key frame are used as references for prediction. The subrate of the key frames and the non-key frames are 0.7 and 0.2, respectively. All the parameters of the model are set empirically for all test sequences. Concretely, the size of each patch is 8×8 pixels, the number of selected similar patches/blocks C is 10. The values of λ_1 and λ_2 are both set to be 0.001, and the values of η_1 and η_2 are both set to be 0.01.

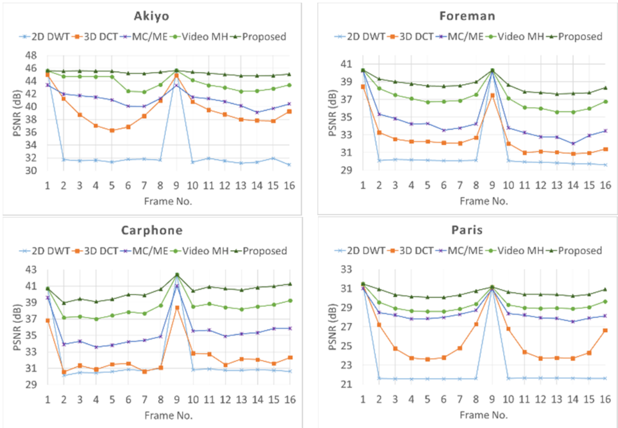

Fig. 1 CS reconstruction performance of the four video sequences using the proposed algorithm and the four comparative methods. The GOP size is 8. The first frame in a GOP is a key frame, and the others are non-key frames

(a) Original

(b) 2D DWT (30.15 dB)

(c) 3D DCT (32.27 dB)

(d) MC/ME (34.28 dB)

(e) Video MH (36.72 dB)

(f) Proposed (38.54 dB)

Fig. 2 Visual comparison of CS reconstruction results for the 5-*th* frame of 'Foreman' by different methods.

Fig. 1 illustrates the CS reconstruction performance of two GOPs of the four test sequences using the five algorithms. It is obvious that the proposed algorithm (dark green) achieves the highest PSNR performance for all the test video frames. On average, the proposed algorithm outperforms 2D DWT, 3D DCT, MC/ME, and video MH by 8.84 dB, 6.18 dB, 3.83 dB, 1.52 dB respectively.

The visual results of the recovered 5-*th* frame of 'Foreman' using the five algorithms are presented in Fig. 2. Obviously, our proposed algorithm not only yields the highest objective score in PSNR, but also preserves fine details in the frames and has much clearer visual results than the other comparative methods.

The computational complexity of the proposed algorithm is O (N) and it takes about 6 minutes to reconstruct a video frame.

VI. Conclusions

In this paper, we propose an effective CS video reconstruction algorithm via a novel dual structured-sparsity model. In this model, we design a strategy to take advantage of two prediction techniques for modelling the structured-sparsity of natural video signals. Owing to the collaborative design, the temporal correlations existing between video frames can be fully exploited. The experimental results demonstrate that the proposed algorithm outperforms the state-of-the-art methods in the literature. This work can be used in scenarios that require a simple encoder or high-quality video reconstruction. Possible future work concerning CS-based video coding includes design of the projection matrix, compression of the measurements, etc.

Acknowledgment

This work was supported in part by the National Basic Research Program of China (973 Program, 2015CB351800), National Natural Science Foundation of China (61322106, 61421062), and Shenzhen Peacock Plan, which are gratefully acknowledged.

References

[1] J. Zhang, C. Zhao, D. Zhao, and W. Gao, "Image Compressive Sensing Recovery Using Adaptively Learned Sparsifying Basis via L0 Minimization," Signal Processing, vol.103, pp. 114-126, Oct. 2014.

[2] J. Zhang, D. Zhao, W. Gao, "Group-based Sparse Representation for Image Restoration," IEEE Transactions on Image Processing, vol. 23, no. 8, pp. 3336-3351, Aug. 2014.

[3] G. J. Sullivan, J. Ohm, W. J. Han, and T. Wiegand. (2012, Dec.). Overview of the high efficiency video coding (HEVC) standard. IEEE Trans. Circuits Syst. Video Technol. 22(12), pp. 1649-1668.

[4] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra. (2003, July). Overview of the H. 264/AVC video coding standard. IEEE Transactions on Circuits and Systems for Video Technology. 13(7), pp. 560-576.

[5] J. E. Fowler, S. Mun, and E. W. Tramel, "Block-Based Compressed Sensing of Images and Video," Foundations and Trends in Signal Processing, vol. 4, no. 4, pp. 297-416, March 2012.

[6] S. Mun and J. E. Fowler, "Block Compressed Sensing of Images Using Directional Transforms," in *Proc. IEEE International Conference on Image Processing*, Cairo, 2009, pp. 3021 - 3024.

[7] M. B. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "Compressive imaging for video representation and coding," in Picture Coding Symp. (PCS), Beijing, China, 2006.

[8] J. Park and M. Wakin, "A Multiscale Framework for Compressive Sensing of Video," in Proc. Picture Coding Symp. (PCS), Chicago, IL, US, 2009, pp. 197-200.

[9] S. Mun and J. E. Fowler, "Residual reconstruction for block-based compressed sensing of video," in Proc. IEEE Data Compression Conf., Snowbird, UT, 2011, pp. 183–192.

[10] E. W. Tramel and J. E. Fowler, "Video Compressed Sensing with Multihypothesis," in Proc. IEEE Data Compression Conf., Snowbird, UT, 2011, pp. 193 - 202.

[11] J. Zhang, D. Zhao, C. Zhao, R. Xiong, S. Ma, W. Gao, "Image Compressive Sensing Recovery via Collaborative Sparsity," IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 2, no. 3, pp. 380–391, Sep. 2012.

[12] C. Zhao, S. Ma, and W. Gao. "Video compressive sensing via structured Laplacian modelling." in Proc. IEEE Visual Communications and Image Processing Conference (VCIP), Dec. 2014.

[13] X. Zhang, R. Xiong, S. Ma, G. Li and W. Gao, "Video super-resolution with registration-reliability regulation and adaptive total variation, Journal of Visual Communication and Image Representation, Volume 30, Pages 181-190, July 2015.

[14] T. Goldstein and S. Osher. (2009, Apr.). The Split Bregman Method for L1-Regularized Problems. SIAM J. Imaging Sci.. 2(2), pp. 323–343.

[15] J. Zhang, D. Zhao, R. Xiong, S. Ma, W. Gao. (2014, Jun.). Image Restoration Using Joint Statistical Modeling in Space-Transform Domain. IEEE Trans. Circuits Syst. Video Technol.. 24(6), pp.915-928.

[16] C. Zhao, S. Ma and W. Gao, "Image Compressive-Sensing Recovery Using Structured Laplacian Sparsity in DCT Domain and Multi-Hypothesis Prediction", IEEE International Conference on Multimedia & Expo (ICME), Chengdu, China, July 2014.