PAPER
Multi-Hypothesis Prediction Scheme Based on the Joint Sparsity Model

Can CHEN[†], **Chao ZHOU**[†], *Nonmembers*, **Jian LIU**[††], *and* **Dengyin ZHANG**[†††,††††a)], *Members*

SUMMARY Distributed compressive video sensing (DCVS) has received considerable attention due to its potential in source-limited communication, e.g., wireless video sensor networks (WVSNs). Multi-hypothesis (MH) prediction, which treats the target block as a linear combination of hypotheses, is a state-of-the-art technique in DCVS. The common approach is under the supposition that blocks that are dissimilar from the target block are given lower weights than blocks that are more similar. This assumption can yield acceptable reconstruction quality, but it is not suitable for scenarios with more details. In this paper, based on the joint sparsity model (JSM), the authors present a Tikhonov-regularized MH prediction scheme in which the most similar block provides the similar common portion and the others blocks provide respective unique portions, differing from the common supposition. Specifically, a new scheme for generating hypotheses and a Euclidean distance-based metric for the regularized term are proposed. Compared with several state-of-the-art algorithms, the authors show the effectiveness of the proposed scheme when there are a limited number of hypotheses.

key words: distributed compressive video sensing (DCVS), multi-hypothesis (MH) reconstruction, joint sparsity model (JSM), wireless video sensor networks (WVSNs)

1. Introduction

Compressed sensing (CS) [1], [2], which involves signal sampling with far fewer measurements than required by Nyquist theory [3], is emerging as a desirable framework for signal acquisition. Under certain conditions [4], a signal $x \in R^{n \times 1}$ can be reconstructed with a high probability from its measurement vector $y \in R^{m \times 1}$ ($m < n$) by solving an optimization problem. According to different sparse models, two categories of CS recovery exist: (1) single subspace models, in which there are k non-zero coefficients in the representation vector corresponding to x which lies in a k-dimensional single subspace spanned by k basis

vectors [2], [5], and (2) union of subspace models, for which x lies in a union of subspaces [6]–[8] corresponding to a block-sparse structure [9]. Several convex or greedy algorithms can be employed to iteratively solve these optimization problems [36]–[39]. However, these algorithms suffer from high computational complexity issues. Thus, several works have introduced deep learning networks to recover signals from measurements [10]–[13]. Moreover, unrolling methods [14], [15] combine deep learning networks and iterative algorithms to learn more realistic signals from training data.

Recently, CS has been widely adopted in video acquisition and reconstruction because it can alleviate the burden on encoders. Combining CS and DVC [16], distributed compressive video sensing (DCVS) [27], [41] has received considerable attention from researchers. In DCVS, complex motion estimation (ME) and motion compensation (MC) processes are shifted from the encoder to the decoder, making the method suitable for resource-limited applications, e.g., wireless video sensor networks (WVSNs) [17]. To reduce the storage requirements of encoders, Lu [18] presented a block-based scheme for DCVS instead of sampling signals in a global manner. Various signal models have been proposed to reconstruct video sequences, such as two-dimensional (2D) models [19], [20] and three-dimensional (3D) models [21]. However, 2D models do not exploit the correlations among frames; and 3D models suffer from high computational complexity issues. By incorporating ME/MC techniques, Mun [22] presented a residual reconstruction method for block-based DCVS that efficiently exploits temporal and spatial correlations. Several schemes [23], [24] that mine sparsity in the residual domain have been proposed to enhance the reconstruction quality.

Multi-hypothesis (MH) prediction algorithms [27]–[31], which utilize a linear combination of hypotheses to predict the target block, can yield state-of-the-art results for DCVS. Although extensive efforts have been made to improve these algorithms, there are still some challenges that must be addressed:

1). *Generate the hypothesis set.* Generally, blocks extracted directly from side information (SI) within a search window centered on the target block are stacked to generate the hypothesis set. Several adaptive schemes have been proposed to optimize the hypothesis set [25], [26]. However, determining how to utilize these blocks and further improve the reconstruction quality remains a research question.

2). *Regularize the optimization problem.* Regularized

terms play important roles in MH prediction. By assuming that the weight vector of hypotheses is sparse, Do proposed the DISCOS approach [27]. Moreover, Fowler [28] presented a Tikhonov-regularized MH scheme for sparsity that was irrelevant. Additionally, several derivatives of this method [29], [30] have been proposed. However, these regularizations share the common assumption that blocks that are dissimilar from the target block should be given lower weights than blocks that are more similar to the target, limiting the reconstruction performance of MH prediction.

The joint sparsity model (JSM) [40], which assumes that two successive frames or blocks in the same scene are visually similar and have similar common and unique portions, can effectively exploit the relevant correlations in DCVS. Based on JSM, a Tikhonov-regularized MH prediction scheme is presented in this study. The main contributions of this study are two-fold:

1). Based on JSM, the authors propose a new scheme for generating the hypothesis set and a Euclidean distance-based metric for the regularized term.

2). The proposed theory, that the most similar block provides the similar common portion and the others blocks provide respective unique portions, breaks with the common-sense supposition that blocks that are more similar to the target should be given higher weights than dissimilar blocks.

The remainder of this paper is organized as follows. Section 2 provides the research background, and we review JSM and the frameworks of MH prediction. In Sect. 3, we describe the relationships between JSM and MH prediction and present the proposed Tikhonov-regularized MH prediction scheme in detail. The experimental results and conclusions are provided in Sect. 4 and Sect. 5, respectively.

2. Background

2.1 Joint Sparsity Model (JSM)

JSM [40] assumes that two successive frames or blocks in the same scene are visually similar and they should have similar common and unique portions. Conceptually, the two vectored frames or blocks, $x_i \in R^{n \times 1}$ and $x_{i+1} \in R^{n \times 1}$, can be expressed as follows.

$$x_i = x_c + x_{u\_i} \tag{1}$$
$$x_{i+1} = x_c + x_{u\_i+1} \tag{2}$$

where $x_c \in R^{n \times 1}$ is the similar common portion and $x_{u\_i} \in R^{n \times 1}$ and $x_{u\_i+1} \in R^{n \times 1}$ denote the respective unique portions of x_i and x_{i+1}, respectively.

2.2 Multi-Hypothesis (MH) Prediction-Based Framework

Based on ME/MC, scholars proposed the MH prediction technique for DCVS to enhance the reconstruction quality. MH prediction uses a linear combination of blocks to estimate the target block. DCVS shifts complex ME/MC tasks

from the encoder to the decoder to alleviate the burden on the encoder. Thus, the vectored target block $x_i \in R^{n \times 1}$ is unavailable and only the measurement vector $y_i \in R^{m \times 1}$ is available at the decoder. Based on the JL lemma [34], MH recasts the optimization problem from the pixel domain to the measurement domain:

$$\omega_i = \arg\min_{\omega} \|y_i - \Phi H_i \omega\|_2^2 \tag{3}$$

where $\Phi \in R^{m \times n}$ represents the measurement matrix, $H_i \in R^{n \times p}$ denotes the set of p hypotheses directly extracted from the SI within a search window centered on the target block, and $\omega_i \in R^{p \times 1}$ is the weight vector. Several regularizations have been proposed to solve this optimization problem, e.g., sparse-based regularization [27], [30], Tikhonov regularization [28], [31], and elastic net-based regularization [29]. To implement DCVS in resource-limited WVSNs for real-time field environmental monitoring, the overriding objective is to not increase the computational complexity. Thus, we expand the Tikhonov-regularized MH prediction method [28] to obtain the proposed scheme due to its low computational complexity and acceptable performance. Using Tikhonov regularization, the calculation of the weight vector of the i-th block $\bar{\omega}_i \in R^{p \times 1}$ can be described as follows.

$$\bar{\omega}_i = \arg\min_{\omega} \|y_i - \Phi H_i \omega\|_2^2 + \lambda \|\Gamma \omega\|_2^2 \tag{4}$$

where λ is a non-negative real value parameter and $\Gamma \in R^{p \times p}$ is a Tikhonov regularization matrix. The closed form of (4) can be given as follows.

$$\bar{\omega}_i = \left((\Phi H_i)^T (\Phi H_i) + \lambda \Gamma^T \Gamma \right)^{-1} (\Phi H_i)^T y_i \tag{5}$$

Residual reconstruction [22] is performed after MH prediction to further enhance the reconstruction quality. Commonly, Γ is given as follows:

$$\Gamma = \text{diag}\left(\|y_i - \Phi H_{i,1}\|_2^2, \ldots, \|y_i - \Phi H_{i,p}\|_2^2 \right) \tag{6}$$

where $H_{i,j=1,2\ldots p}$ denotes each hypothesis of H_i. This approach is based on the common-sense assumption that blocks that are dissimilar from the target block should be given lower weights than blocks that are more similar. This assumption can yield acceptable reconstruction quality, but it is not suitable for scenarios with more details. Thus, we propose a multi-hypothesis prediction scheme based on the joint sparsity model to enhance the details.

3. The Proposed Tikhonov-Regularized MH Prediction Scheme

In this section, based on JSM, we propose a Tikhonov-regularized MH prediction scheme. Unlike the existing MH prediction schemes, which assume that blocks dissimilar to the target block are given lower weights than those more similar to the target block, the proposed scheme assumes that the most similar block provides a similar common portion, whereas the other blocks provide respective unique

portions. Specifically, a new scheme for generating hypotheses and a Euclidean distance-based metric for the regularized term are proposed.

According to (1) and (2), x_{i+1} can be predicted by x_i:

$$
\begin{aligned}
x_{i+1} = x_c + x_{u\_i+1} &= x_i + (x_{u\_i+1} - x_{u\_i}) \\
&= x_i + (x_{i+1} - x_i) \\
&= x'_c + x'_{u\_i+1}
\end{aligned}
\tag{7}
$$

where $x'_c = x_i$ and $x'_{u\_i+1} = x_{i+1} - x_i$. In other words, $x'_c \in R^{n \times 1}$ can be regarded as the similar common portion provided by the similar block x_i, and $x'_{u\_i+1} \in R^{n \times 1}$ as the respective unique portions provided by the residual block $x_{i+1} - x_i$. Based on JSM, the final MH prediction $\bar{x}_i \in R^{n \times 1}$ can be expressed as follows.

$$
\bar{x}_i = x'_c + x'_{u\_i}
\tag{8}
$$

Thus, we propose the assumption that the most similar block $\breve{H}_i \in R^{n \times 1}$ provides the similar common portion x'_c and the others blocks $\hat{H}_i \in R^{n \times (p-1)}$ provide respective unique portions x'_{u_i}:

$$
\bar{x}_i = x'_c + x'_{u\_i} = \breve{H}_i + \hat{H}_i \hat{\omega}_i
\tag{9}
$$

where $\hat{\omega}_i \in R^{(p-1) \times 1}$ represents the corresponding weight vector of $\hat{H}_i$.

3.1 Hypothesis Set Generation Scheme

We first generate the initial hypothesis set $\bar{H}_i \in R^{n \times p}$ by directly extracting blocks from the SI within a search window centered on the target block. Then, we use the Euclidean distance between measurement vectors to select the most similar block $\breve{H}_i$:

$$
\breve{H}_i = \underset{\bar{H}_{i,j}}{\arg \min} \|y_i - \Phi \bar{H}_{i,j}\|_2^2
\tag{10}
$$

where $\bar{H}_{i,j=1...p}$ denotes each hypothesis in $\bar{H}_i$. In (7), respective unique portions are provided by the residual block between the similar block and the other blocks. Thus, we generate $\hat{H}_i$ as follows.

$$
\hat{H}_i = \left(\bar{H}_{i,1} - \breve{H}_i, \ldots, \bar{H}_{i,p-1} - \breve{H}_i \right)
\tag{11}
$$

Note that the dimension of $\hat{H}_i$ is $p - 1$ because we remove $\breve{H}_i$ from $\bar{H}_i$ and each row of $\hat{H}_i$ is the residual between each hypothesis and $\breve{H}_i$.

3.2 Euclidean Distance-Based Metric

Unlike existing MH prediction schemes, the proposed scheme implements Tikhonov-regularized MH prediction for the measurement vector of respective unique portions $\hat{y}_i$ instead of y_i.

$$
\hat{y}_i = y_i - \Phi \breve{H}_i
\tag{12}
$$

In the proposed scheme, a similar common portion is provided by $\breve{H}_i$, and respective unique portions are provided by

$\hat{H}_i$. To preserve the respective unique portions over the similar common portion, we assign higher weights to blocks in $\hat{H}_i$ that are dissimilar from those in $\breve{H}_i$. Thus, we propose a Euclidean distance-based metric $\bar{\Gamma} \in R^{(p-1) \times (p-1)}$ for the Tikhonov-regularized MH prediction as follows.

$$
\bar{\Gamma} = \text{diag}\left(\frac{1}{\|\breve{H}_i - \hat{H}_{i,1}\|_2^2}, \ldots, \frac{1}{\|\breve{H}_i - \hat{H}_{i,p-1}\|_2^2} \right)
\tag{13}
$$

Note that residual blocks in $\hat{H}_i$ have less energy because the blocks in $\bar{H}_i$ are highly related. Thus, assigning higher weights to these dissimilar blocks will not affect the similar common portion but will enhance the respective unique portions. Furthermore, the measurement vector of the original block which can be regarded as a low-dimensional feature vector can be unreliable when used in similarity measurements. Thus, we use pixel-domain vectors instead of measurement-domain vectors to measure similarity. By substituting $\hat{y}_i$, $\hat{H}_i$, and $\bar{\Gamma}$ for y_i, H_i, and Γ in (5) and (9), we obtain $\hat{\omega}_i$ and $\bar{x}_i$.

4. Results

In this section, we denote the proposed scheme as MH-JSM. We compare it with some other state-of-the-art MH algorithms experimentally and conceptually: (1). MH-TIK [28], which adopts a single Tikhonov regularization; (2). MH-wEnet [29], which associates a weighted l_1 regularization and a Tikhonov regularization as the reweighted elastic net regularization; and (3). MH-ST [30], which combines a Tikhonov regularization and a sparsity regularization on the frame. We evaluate MH-JSM based on the following classical CIF (352×288) sequences: Coastguard, Container, Foreman and Hall. We use the average peak signal-to-noise ratio (PSNR) of all non-key frames as the comparative index. We adopt the bilateral MC [35] algorithm to generate the SI for each non-key frame. The number of hypotheses is highly correlated with the MH prediction performance. Thus, we conduct experiments with different numbers of hypotheses (e.g. $p = 10$ and $p = 40$). In all experiments, other parameters are set as in [28] to maintain equality. Specifically, the sampling rate of key frames is set to 0.7, the group of pictures (GOP) is set to 8, and the size of block is set to 16. All experiments were conducted in MATLAB R2015b on a Dell laptop with an Intel (R) Core (TM) i7-4710HQ CPU (2.5 GHz).

Table 1 shows the average PSNR for various video sequences. In the region where $p = 40$, compared with MH-TIK, MH-wEnet, and MH-ST, the performance of MH-JSM is unstable and poor at low sampling rates but satisfactory at high sampling rates. This performance variation occurs because the similar common portion has a greater impact on the performance than the respective unique portions in MH-JSM. However, adopting a larger search window to generate more blocks can introduce a large number of uncorrelated blocks, resulting in high-energy residual blocks in $\hat{H}_i$. These high-energy residual blocks can influence the similar common portion. Furthermore, the Euclidean distance between

Table 1   Reconstruction performance (dB) comparisons

Sequence	Sampling rate	p=10				p=40			
		MH-JSM (proposed)	MH-TIK	MH-wEnet	MH-ST	MH-JSM (proposed)	MH-TIK	MH-wEnet	MH-ST
Container	0.1	**31.52**	29.78	30.02	29.82	28.24	29.56	29.82	29.64
	0.2	**32.56**	30.52	30.69	30.79	23.26	30.76	30.9	30.97
	0.3	**33.29**	31.25	31.35	31.41	31.59	31.7	31.69	31.79
	0.4	**33.94**	32.01	31.97	32.26	33.43	32.52	32.41	32.66
	0.5	34.71	32.95	32.76	33.07	**34.76**	33.49	33.29	33.7
Coastguard	0.1	**30.6**	30.04	30.44	30.15	27.62	30.06	30.41	30.25
	0.2	**31.85**	30.73	31.07	31.22	21.84	31.12	31.26	31.45
	0.3	**32.68**	31.44	31.51	31.76	30.41	31.94	31.96	32.29
	0.4	**33.48**	32.17	32.11	32.59	32.51	32.73	32.71	32.95
	0.5	**34.41**	33.11	32.99	33.34	34.04	33.67	33.48	33.88
Foreman	0.1	28.25	28.88	**29.03**	28.78	25.25	28.98	29	28.89
	0.2	29.65	29.7	30.1	30.03	19.23	29.99	30.32	**30.34**
	0.3	30.61	30.55	30.57	30.92	27.7	30.88	30.92	**31.25**
	0.4	31.43	31.37	31.29	31.4	29.83	31.73	31.68	**31.92**
	0.5	32.73	32.68	32.55	32.71	31.43	33.07	32.91	**33.24**
Hall	0.1	**33.56**	31.68	32.14	32.01	30.07	31.73	32.23	31.83
	0.2	**34.71**	32.53	33.09	32.89	25.12	32.93	33.25	33.17
	0.3	**35.58**	33.39	33.67	34.11	33.56	33.96	34	34.32
	0.4	**36.38**	34.22	34.15	34.77	35.62	34.91	34.86	35.16
	0.5	**37.33**	35.36	35.21	35.99	37.18	36.09	35.96	36.32

Table 2   Reconstruction time (s) comparisons

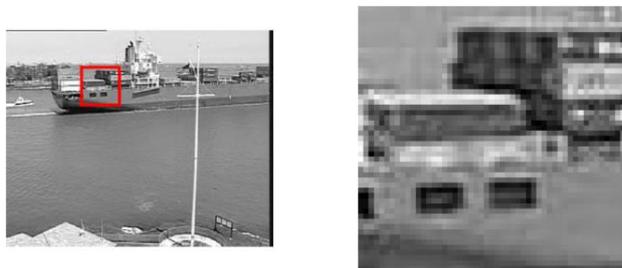| Sequence	Sampling rate	p=10				p=40			
		MH-JSM (proposed)	MH-TIK	MH-wEnet	MH-ST	MH-JSM (proposed)	MH-TIK	MH-wEnet	MH-ST
Container	0.1	2.1	1.9	1.1	2.7	7.8	7.6	3.9	9.9
	0.2	2	1.9	1.5	2.7	7.9	7.6	5.6	10.1
	0.3	2.2	2	1.9	2.8	8.4	8.1	8	10.5
	0.4	2.1	1.9	2.8	2.8	8.6	8.3	12.1	10.8
	0.5	2.1	2	4.9	3	8.7	8.2	20.5	11.1
Coastguard	0.1	2	1.9	1.3	2.8	7.8	7.7	4.1	10
	0.2	2	2	1.6	2.8	8.1	7.9	5.8	10.2
	0.3	2.1	2	2	3	8.6	8.4	8.5	10.8
	0.4	2.2	1.9	3.1	3.1	8.8	8.5	12.3	11
	0.5	2.2	2	4.7	3.1	8.9	8.5	20.2	11.2
Foreman	0.1	2	2	1.1	2.5	7.6	7.5	3.8	9.7
	0.2	2.1	1.9	1.5	2.7	8	7.7	5.5	10.1
	0.3	2	1.9	1.8	2.7	8.4	8.2	8.2	10.6
	0.4	2.2	1.9	2.8	3	8.7	8.3	12.7	10.8
	0.5	2.1	2	4.5	3.1	8.6	8.3	20.8	10.9
Hall	0.1	2.1	2	1.2	2.4	7.4	7.3	3.7	9.5
	0.2	2.1	2	1.6	2.6	7.7	7.5	5.5	9.7
	0.3	2.1	1.9	2.1	2.8	8.2	7.9	7.9	10.3
	0.4	2.2	2	3.1	2.9	8.5	8	11.7	10.4
	0.5	2.1	2	4.8	3	8.6	8.2	19.9	10.7

measurement vectors is used to select the most similar block $\bar{H}_i$. The measurement vector of the original block which can be regarded as a low-dimensional feature vector can be unreliable when used in similarity measurement applications, especially at low sampling rates, resulting in failure in selecting the most similar block. Table 1 shows that increasing in the number of hypotheses cannot guarantee a reconstruction improvement for each algorithm because it enhances the representation ability and introduces inaccurate hypotheses. Moreover, the computational complexity exponentially increases, as shown in Table 2. Thus, the ideal scheme improves the reconstruction quality with as few hypotheses as possible.

In the region where p = 10, MH-JSM stably outperforms the other algorithms for the Container, Coastguard, and Hall sequences because the blocks extracted from a small search window are highly correlated. We take experiments with a sampling rate of 0.5 as examples. For the Container sequence, MH-JSM outperforms MH-TIK, MH-wEnet, and MH-ST by 1.76 dB, 1.95 dB, and 1.64 dB, respectively. Moreover, for the Coastguard sequence, MH-JSM outperforms these methods by 1.3 dB, 1.42 dB, and 1.07 dB. Additionally, for the Hall sequence, MH-JSM outperforms the other methods by 1.97 dB, 2.12 dB, and 1.34 dB. However, MH-JSM does not obviously improve the reconstruction quality for the Foreman sequence. These results differ because the Foreman sequence is a shot-shifted video sequence, and the others are shot-fixed video sequences. In shot-shifted video sequences, the same portion of a frame in one shot can disappear in the next shot, making it difficult to select the most similar block and related blocks. However, this task is easy in shot-fixed video sequences. Figure 1 shows visual reconstruction comparisons of the first non-key frame in the Container sequence at a
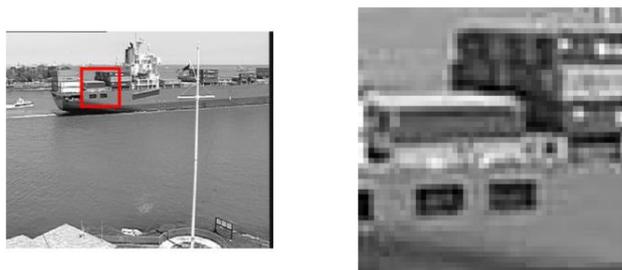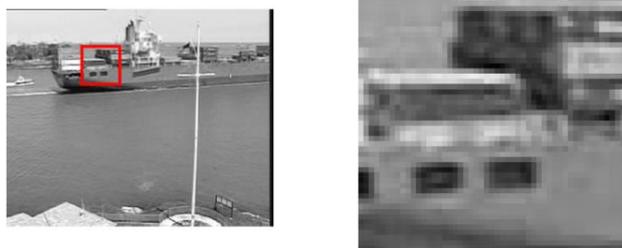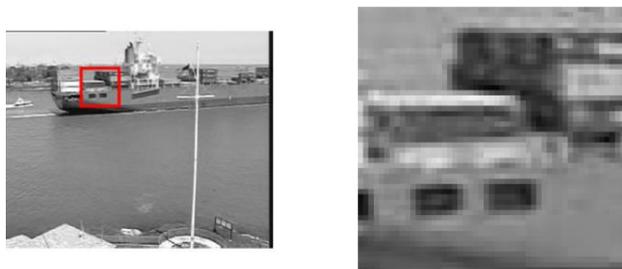
(a) Original

(b) MH-JSM (proposed): 31.35 dB
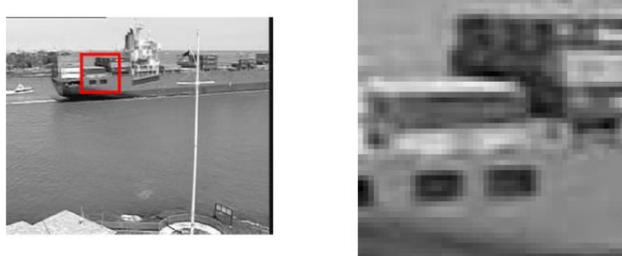
(c) MH-TIK: 29.71 dB

(d) MH-wEnet: 29.94 dB

(e) MH-ST: 29.87 dB

Fig. 1 Visual comparisons of the Container sequence. (a) Original; (b) MH-JSM (proposed); (c) MH-TIK; (d) MH-wEnet; (e) MH-ST

sampling rate of 0.1.

To investigate the complexity of different algorithms, we present the reconstruction time of each algorithm in Table 2. Overall, the reconstruction time of MH-JSM is slightly greater than that of MH-TIK. This extra time is required to divide blocks into two hypothesis sets instead of generating one hypothesis set. To implement DCVS in resource-limited WVSNs for real-time field environmental monitoring, the overriding objective is to not increase the computational complexity. Compared to other methods, MH-JSM achieves a better balance between the reconstruction quality and computational complexity.

5. Conclusion

To implement DCVS in resource-limited WVSNs for real-time field environmental monitoring, a Tikhonov-regularized MH prediction scheme in which the most similar block provides the similar common portion and the others blocks provide respective unique portions is proposed in this paper. The proposed scheme differs from the common-sense assumption. Moreover, a new scheme for selecting hypotheses and a Euclidean distance-based metric for the regularized term are proposed. Compared to several state-of-the-art algorithms, MH-JSM achieves a better balance between the reconstruction quality and computational complexity. In future work, strategies for selecting blocks in shot-shifted scenes should be further investigated to improve the applicability of MH-JSM.

Acknowledgments

References

[1] E.J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inf. Theory, vol.52, no.2, pp.489–509, 2006, 10.1109/TIT.2005.862083.

[2] D.L. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, vol.52, no.4, pp.1289–1306, 2006, 10.1109/TIT.2006.871582.

[3] H.J. Landau, "Sampling, data transmission, and the Nyquist rate," Proc. IEEE, vol.55, no.10, pp.1701–1706, 1967, 10.1109/PROC.1967.5962.

[4] E.J. Candes, "The restricted isometry property and its implications for compressed sensing," Comptes rendus mathematique, vol.346, no.9-10, pp.589–592, 2008, 10.1016/j.crma.2008.03.014.

[5] E.J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies," IEEE Trans. Inf. Theory, vol.52, no.12, pp.5406–5425, 2006, 10.1109/TIT.2006.885507.

[6] Y.M. Lu and M.N. Do, "A theory for sampling signals from a union of subspaces," IEEE Trans. Signal Process., vol.56, no.6,

pp.2334–2345, 2008, 10.1109/TSP.2007.914346.

[7] T. Blumensath, "Sampling and reconstructing signals from a union of linear subspaces," IEEE Trans. Inf. Theory, vol.57, no.7, pp.4660–4671, 2011, 10.1109/TIT.2011.2146550.

[8] Y.C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," IEEE Trans. Inf. Theory, vol.55, no.11, pp.5302–5316, 2009, 10.1109/TIT.2009.2030471.

[9] Y.C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: uncertainty relations and efficient recovery," IEEE Trans. Signal Process., vol.58, no.6, pp.3042–3054, 2010, 10.1109/TSP.2010.2044837.

[10] A. Mousavi, A.B. Patel, and R.G. Baraniuk, "A deep learning approach to structured signal recovery," 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE, pp.1336–1343, 2015, 10.1109/ALLERTON.2015.7447163.

[11] A. Mousavi and R.G. Baraniuk, "Learning to invert: Signal recovery via deep convolutional networks," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, pp.2272–2276, 2017, 10.1109/ICASSP.2017.7952561.

[12] X. Lu, W. Dong, P. Wang, et al., "ConvCSNet: A convolutional compressive sensing framework based on deep learning," arXiv preprint arXiv:1801.10342, 2018.

[13] H. Yao, F. Dai, D. Zhang, et al., "DR2-net: Deep residual reconstruction network for image compressive sensing," arXiv preprint arXiv:1702.05743, 2017.

[14] U.S. Kamilov and H. Mansour, "Learning optimal nonlinearities for iterative thresholding algorithms," IEEE Signal Process. Lett., vol.23, no.5, pp.747–751, 2016, 10.1109/LSP.2016.2548245.

[15] C.A. Metzler, A. Mousavi, and R.G. Baraniuk, "Learned D-AMP: Principled neural network based compressed image recovery," arXiv preprint arXiv:1704.06625, 2017.

[16] B. Girod, A.M. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," Proc. IEEE, vol.93, no.1, pp.71–83, 2005, 10.1109/JPROC.2004.839619.

[17] F. Pereira, L. Torres, C. Guillemot, T. Ebrahimi, R. Leonardi, and S. Klomp, "Distributed video coding: selecting the most promising application scenarios," Signal Processing: Image Communication, vol.23, no.5, pp.339–352, 2008, 10.1016/j.image.2008.04.002.

[18] L. Gan, "Block compressed sensing of natural images," Proc. international conference on digital signal processing, IEEE, pp.403–406, 2007, 10.1109/ICDSP.2007.4288604.

[19] J. Zhang, D. Zhao, and W. Gao, "Group-based sparse representation for image restoration," IEEE Trans. Image Process., vol.23, no.8, pp.3336–3351, 2014, 10.1109/TIP.2014.2323127.

[20] 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, 2014, 10.1016/j.sigpro.2013.09.025.

[21] P. Llull, X. Liao, X. Yuan, J. Yang, D. Kittle, L. Carin, G. Sapiro, and D.J. Brady, "Coded aperture compressive temporal imaging," Optics express, vol.21, no.9, pp.10526–10545, 2013, 10.1364/OE.21.010526.

[22] S. Mun and J.E. Fowler, "Residual reconstruction for block-based compressed sensing of video," Data Compression Conference (DCC), IEEE, pp.183–192, 2011, 10.1109/DCC.2011.25.

[23] C. Zhao, S. Ma, J. Zhang, R. Xiong, and W. Gao, "Video compressive sensing reconstruction via reweighted residual sparsity," IEEE Trans. Circuits Syst. Video Technol., vol.27, no.6, pp.1182–1195, 2017, 10.1109/TCSVT.2016.2527181.

[24] K. Chang, P.L.K. Ding, and B. Li, "Compressive sensing reconstruction of correlated images using joint regularization," IEEE Signal Process. Lett., vol.23, no.4, pp.449–453, 2016, 10.1109/LSP.2016.2527680.

[25] J. Chen, N. Wang, F. Xue, and Y. Gao, "Distributed compressed video sensing based on the optimization of hypothesis set update technique," Multimedia Tools and Applications, vol.76, no.14,

pp.15735–15754, 2017, 10.1007/s11042-016-3866-4.

[26] Y. Kuo, K. Wu, and J. Chen, "A scheme for distributed compressed video sensing based on hypothesis set optimization techniques," Multidimensional Systems and Signal Processing, vol.28, no.1, pp.129–148, 2017, 10.1007/s11045-015-0337-4.

[27] T.T. Do, Y. Chen, D.T. Nguyen, N. Nguyen, L. Gan, and T.D. Tran, "Distributed compressed video sensing," 16th IEEE International Conference on Image Processing (ICIP), IEEE, pp.1393–1396, 2009, 10.1109/CISS.2009.5054678.

[28] E.W. Tramel and J.E. Fowler, "Video compressed sensing with multihypothesis," 2011 Data Compression Conference, pp.193–202, 2011, 10.1109/DCC.2011.26.

[29] J. Chen, Y. Chen, D. Qin, and Y. Kuo, "An elastic net-based hybrid hypothesis method for compressed video sensing," Multimedia Tools and Applications, vol.74, no.6, pp.2085–2108, 2015, 10.1007/s11042-013-1743-y.

[30] M. Azghani, M. Karimi, and F. Marvasti, "Multihypothesis compressed video sensing technique," IEEE Trans. Circuits Syst. Video Technol., vol.26, no.4, pp.627–635, 2016, 10.1109/TCSVT.2015.2418586.

[31] C. Chen, D. Zhang, and J. Liu, "Resample-based hybrid multi-hypothesis scheme for distributed compressive video sensing," IEICE Trans. Inf. & Syst., vol.100, no.12, pp.3073–3076, 2017, 10.1587/transinf.2017EDL8133.

[32] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," IEEE Trans. Pattern Anal. Mach. Intell., vol.35, no.11, pp.2765–2781, 2013, 10.1109/TPAMI.2013.57.

[33] Y. Li, W. Dai, J. Zou, H. Xiong, and Y.F. Zheng, "Structured sparse representation with union of data-driven linear and multilinear subspaces model for compressive video sampling," IEEE Trans. Signal Process., vol.65, no.19, pp.5062–5077, 2017, 10.1109/TSP.2017.2721905.

[34] W.B. Johnson and J. Lindenstrauss, "Extensions of Lipschitz mappings into a Hilbert space," Contemporary mathematics, vol.26, no.1, pp.189–206, 1984, 10.1090/conm/026/737400.

[35] B.-T. Choi, S.-H. Lee, and S.-J. Ko, "New frame rate up-conversion using bi-directional motion estimation," IEEE Trans. Consum. Electron., vol.46, no.3, pp.603–609, 2000, 10.1109/30.883418.

[36] C.A. Metzler, A. Maleki, and R.G. Baraniuk, "From denoising to compressed sensing," IEEE Trans. Inf. Theory, vol.62, no.9, pp.5117–5144, 2016, 10.1109/TIT.2016.2556683.

[37] D. Needell and J.A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," Applied and computational harmonic analysis, vol.26, no.3, pp.301–321, 2009, 10.1016/j.acha.2008.07.002.

[38] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM journal on imaging sciences, vol.2, no.1, pp.183–202, 2009.

[39] Z. Wen, B. Hou, and L. Jiao, "Joint sparse recovery with semisupervised MUSIC," IEEE Signal Process. Lett., vol.24, no.5, pp.629–633, 2017, 10.1109/LSP.2017.2680603.

[40] M.F. Duarte, M.B. Wakin, D. Baron, et al., "Universal distributed sensing via random projections," Proc. 5th International Conference on Information Processing in Sensor Networks, pp.177–185, 2006.

[41] J. Prades-Nebot, Y. Ma, and T. Huang, "Distributed video coding using compressive sampling," Picture Coding Symposium, pp.1–4, 2009.

Can Chen received the B.S. degree in Nanjing University of Posts and Telecommunications, Nanjing, China, in 2015. He is currently pursuing the Ph.D. degree in signal and information processing in the College of Telecommunications and Information Engineering of Nanjing University of Posts and Telecommunications, Nanjing, China. His research interest include image and video coding, image and video processing, machine learning, and compressive sensing.

Chao Zhou received the B.S. degree in Wuhan University of Technology, Wuhan, China in 2016. He is currently pursuing the Ph.D. degree signal and information processing in the College of Telecommunications and Information Engineering of Nanjing University of Posts and Telecommunications, Nanjing, China. His research interest includes image and video coding, image and video processing, and compressive sensing.

Jian Liu was born in China in 1990. He received the Ph. D. degree in signal and information processing at Nanjing University of Posts and Telecommunications (NJUPT), Nanjing, China, in 2018. Since 2018, he works in Nanjing University of Finance and Economics (NUFE), Nanjing, China, and is currently a lecturer of College of Information Engineering. His research interests are in the areas of nonlinear stochastic resonance (SR) and its applications, including signal detection, signal transmission, and digital communication system.

Dengyin Zhang received the B.S. degree, M.S. degree and Ph.D. degree in Nanjing University of Posts and Telecommunications, Nanjing, China, in 1986, 1989 and 2004 respectively. He is currently a Professor of the School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing, China. He was in Digital Media Lab at Umea University in Sweden as a visiting scholar from 2007 to 2008. His research interests include signal and information processing, networking technique, and information security.