A MULTIHYPOTHESIS-BASED RESIDUAL RECONSTRUCTION SCHEME IN COMPRESSED VIDEO SENSING

Wen-Hao Li, Chun-Ling Yang, Li-Hong Ma*

School of Electronic and Information Engineering, South China University of Technology
Guangzhou 510640, China

ABSTRACT

A multihypothesis-based residual reconstruction scheme (MHRR) is presented in compressed video sensing (CVS). The residual is first predicted by a novel multihypothesis (MH) prediction method and the second prediction residual then is reconstructed independently. In the proposed MHRR, a method of generating hypothesis blocks in residual domain is offered, and the hypothesis blocks are obtained by pixel-domain ME technique and the linear prediction weights are calculated in measurement-domain, which can combine the advantages of pixel-domain ME technique and measurement domain MH prediction. Simulation results show that the proposed MHRR can achieve higher reconstruction performance than SPL (Smoothed Projected Landweber) for residual reconstruction.

Index Terms— compressed video sensing, residual reconstruction, multihypothesis, SPL algorithm

1. INTRODUCTION

Compressed sensing (CS) theory[1] has drawn great attention in the field of signal acquisition and data compression in the last decade. It expresses the idea that signals can be recovered from far fewer measurements than that suggested by the Nyquist sampling theory, as long as they are sparse or approximately sparse in some domain. Compared with conventional signal acquisition schemes which sample first and then compress, CS theory incorporates sampling with compressing, avoids unnecessary waste of sampling resource, and therefore has bright application prospect for resource-constrained circumstances such as wireless sensor networks and medical imaging[2].

In compressed video sensing (CVS) where CS theory is used to acquire video sequences, the chief objective is to reconstruct the video sequences as accurate as possible from far fewer measurements. To accomplish the objective, many reconstruction schemes are presented, among them the schemes that adopt the block-based multihypothesis (MH) prediction strategy[3-8] achieve the best reconstruction quality. Reference [3] recovers frames by performing iterative pixel-domain motion estimation (ME)/motion compensation (MC) and residual reconstruction, while reference [4] performs iterative measurement-domain ME/MC with a Tikhonov regularization and residual reconstruction to obtain higher reconstruction quality. On the basis of reference [4], an extra sparsity constraint in discrete cosine transform domain is attached to the MH prediction mode in [5]. Reference [6] presents a hybrid hypothesis prediction method which incorporates single-hypothesis (SH) and Elastic net-based MH mode, and reference [7] further develops several new hypothesis set optimization techniques to achieve higher reconstruction performance. A two-stage MH reconstruction scheme is developed in [8], which conducts measurement-domain ME/MC and residual reconstruction first and then applies pixel-domain ME/MC and residual reconstruction, and it also obtains the highest reconstruction quality among these schemes[3-8] adopting the block-based MH prediction strategy. It is noteworthy that generating hypothesis blocks and calculating prediction weights are implemented in the same domain, pixel-domain or measurement-domain, in all above MH prediction reconstruction schemes[3-8].

As a matter of fact, residual reconstruction is a key step for the block-based MH prediction strategy. But in all the relevant literatures[3-8], residual reconstruction is not the core and SPL[9] (Smoothed Projected Landweber) is always used for it. In fact, SPL is originally proposed for natural images reconstruction[10] which reconstructs each image blocks independently, thus fails to exploit temporal correlation and spatial correlation between frames and restricts the last reconstruction performance in CVS. With insight into the block-based MH prediction strategy which can obtain better reconstruction performance by well exploiting temporal and spatial correlation, a MH-based residual reconstruction scheme (MHRR) is presented in this paper. There are two contributions in the MHRR: 1) offer a method of generating hypothesis blocks for residual MH prediction because of no reference frames existing obviously; 2) novelly combine generating hypothesis blocks by pixel-domain ME technique with calculating prediction weights in measurement-domain, because this flexible combination can maximize the use of known information in the situation.

This work was supported by National Natural Science Foundation of China (grant 61471173) and Guangdong Natural Science Foundation (grant 2016A030313455).

2. BLOCK-BASED MH PREDICTION STRATEGY

In CVS, the block-based MH prediction strategy is a very efficient reconstruction method. The video sequence to be processed is usually divided into groups of pictures (GOP), the first frame of each GOP treated as key frame and the others as non-key frames. Each frame is sampled by block-based compressed sensing (BCS) approach[10] which is widely used for its low requirement for storage capacity and computation. In BCS, each frame of a GOP is divided into $b \times b$ blocks, then each block is sampled with measurement matrix Φ, $y_{k,i} = \Phi x_{k,i}$, where $x_{k,i}$ denotes a rasterized vector of block i in frame k, Φ is a $m \times b^2$ orthonormal Gaussian random measurement matrix and $y_{k,i}$ is measurement such that sampling rate, or subrate, is m / b^2.

At the reconstruction side, key frames are recovered independently and non-key frames are reconstructed by the block-based MH prediction and residual reconstruction. Several patterns of reference-frame and MH predication methods are presented in reference [3-8] so as to sufficiently exploit temporal correlation.

2.1. Multi-reference-frame pattern

Multi-reference-frame pattern can provide better hypotheses for MH prediction, the kind of multi-reference-frame pattern in reference [8] which is adopted in this paper is illustrated as below, supposed that the size of GOP is 8.

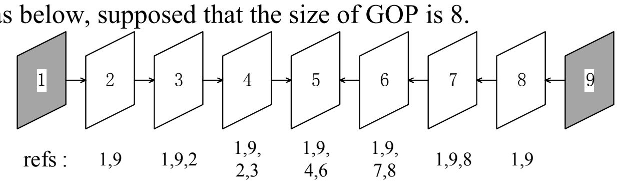


Fig. 1. Multi-reference-frame pattern in [8]

In Fig.1, the 9th frame is the key frame of the next GOP, the arrows indicate the order of reconstruction. The "refs" below the non-key frames is the reference-frame numbers corresponding to each non-key frame, and the largest reference-frame number is 4 for all GOPs.

2.2. Block-based MH prediction methods

Suppose x_{k-1} (the frame $k-1$) is a reference frame of x_k, there are two block-based MH prediction methods, pixel-domain ME/MC method and measurement-domain ME/MC method, for the recovery of $x_{k,i}$ using only available Φ and measurements $y_{k,i}$. In the pixel-domain ME/MC method[3], an approximation $\bar{x}_{k,i}$ of $x_{k,i}$ is first created from $y_{k,i}$, then hypothesis blocks matching $\bar{x}_{k,i}$ are selected in the searching window in x_{k-1} according to some criterions such as mean square error (MSE) or sum of absolute differences (SAD) to constitute hypothesis set $H_{k,i}$, finally the pixel-domain MH prediction is conducted as,

$$\tilde{x}_{k,i} = H_{k,i} w_{k,i}, \tag{1}$$

where $\tilde{x}_{k,i}$ is the MH prediction of $x_{k,i}$, $H_{k,i}$ is a matrix of dimensionality $b^2 \times P$ whose columns $h_1, \cdots, h_P$ are the rasterizations of the hypothesis blocks and $w_{k,i}$ is a linear prediction weight vector of dimensionality $P \times 1$ which is obtained by the following algorithm.

$$w_{k,i} = \underset{w_{k,i}}{\arg\min} \left\| \bar{x}_{k,i} - H_{k,i} w_{k,i} \right\|_2^2 \tag{2}$$

In the measurement-domain ME/MC method[4-8], the hypothesis blocks are first directly selected using $y_{k,i}$ in the searching window in x_{k-1} according to some measurement-domain criterion to constitute hypothesis set $H_{k,i}$, then the MH prediction is conducted as equation (1) and the linear prediction weights are obtained in measurement-domain attached with a Tikhonov regularization, i.e.,

$$w_{k,i} = \underset{w_{k,i}}{\arg\min} \left\| y_{k,i} - \Phi H_{k,i} w_{k,i} \right\|_2^2 + \lambda \left\| \Gamma w_{k,i} \right\|_2^2, \tag{3}$$

where λ is a scale factor and Γ is defined as:

$$\Gamma = diag\left(\left\| y_{k,i} - \Phi h_1 \right\|_2^2, \cdots, \left\| y_{k,i} - \Phi h_P \right\|_2^2 \right) \tag{4}$$

It is worth noting that generating hypothesis blocks and calculating $w_{k,i}$ are implemented in the same domain, pixel-domain or measurement domain, in the both methods, and the modes are used in all the relevant literatures. But in some situations, their combination and implementation in different domain can bring better prediction performance, this idea will be used in our proposed prediction scheme in this paper.

2.3. Residual reconstruction

While $\tilde{x}_{k,i}$ is here, the relationship between measurement-domain residual $y^r_{k,i}$ and $\tilde{x}_{k,i}$ can be expressed as,

$$y^r_{k,i} = \Phi x^r_{k,i} = \Phi(x_{k,i} - \tilde{x}_{k,i}) = y_{k,i} - \Phi \tilde{x}_{k,i}, \tag{5}$$

where $x^r_{k,i}$ is the residual block and $y^r_{k,i}$ is the corresponding measurement. SPL is widely used for residual reconstruction for its simplicity of implementation[9], but its independent recovery fails to exploit temporal and spatial correlation of prediction residuals and restricts the reconstruction quality.

3. MH-BASED RESIDUAL RECONSTRUCTION

Through a lot of simulation, it is found that the measurement domain prediction residuals (first prediction residual) are not always small-energy signal or sparse signal, inspired by the block-based MH prediction strategy which can well exploit temporal and spatial correlation, this paper proposed a MH-based residual reconstruction scheme (MHRR). First, the residual is predicted by a novel block-based MH prediction method; Then, the second prediction residual is recovered using SPL. Temporal and spatial correlation between the first prediction residuals is used in the proposed MHRR, so the second prediction residual has higher sparsity. In the proposed MHRR, hypothesis blocks are obtained by pixel-domain ME technique and the linear prediction weights are calculated in measurement-domain, which can combine the advantages of pixel-domain ME technique and measurement domain MH prediction to get better prediction accuracy. The difficulty in the MHRR is how to obtain suitable hypothesis blocks because there exists no obvious reference frames for

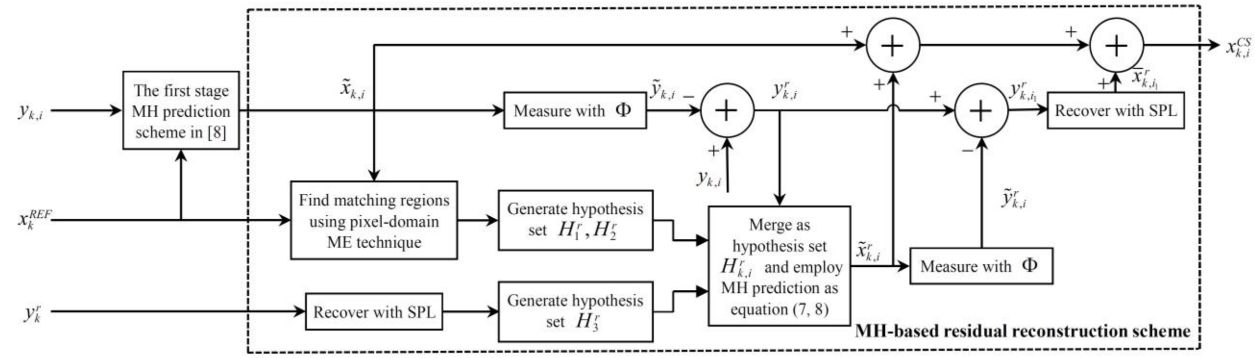


Fig. 2. MH-based residual reconstruction scheme

residual MH prediction.

In the following subsection, the overall framework of the proposed MHRR is given first and then we describe the method of generating hypothesis blocks and calculating the linear prediction weights in details.

3.1. The overall framework

The overall framework of the proposed MHRR is presented in Fig.2, where x_k^{REF} denotes reference frames of x_k , y_k^r is the measurement of x_k^r , $\tilde{y}_{k,i}$ is the measurement of $\tilde{x}_{k,i}$, H_1^r, H_2^r, H_3^r are hypothesis sets of current residual block $x_{k,i}^r$, $\tilde{x}_{k,i}^r$ is the MH prediction of block $x_{k,i}^r$, y_{k,i_1}^r is the second prediction residual measurement, $\bar{x}_{k,i_1}^r$ is the independent recovery from y_{k,i_1}^r with SPL and $x_{k,i}^{cs}$ represents the final result of $x_{k,i}$.

3.2. Hypothesis set generation and MH prediction for the first prediction residual frame

The whole MHRR can be implemented by the following 4 steps and the reconstruction of residual block $x_{k,i}^r$ is taken as an example.

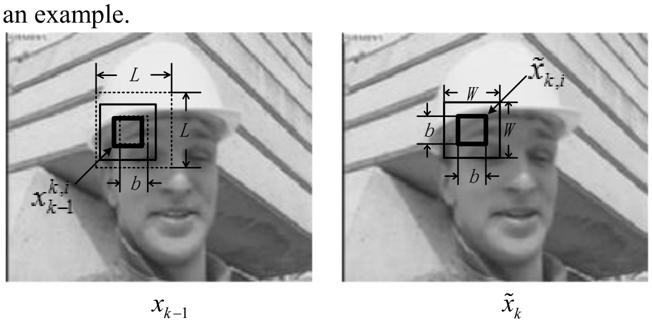


Fig. 3. Illustration to pixel-domain ME

<u>Step 1</u>: Generate hypothesis set H_1^r from the adjacent reference frame using pixel-domain ME technique

<u>Step1.1</u>: Search the best matching block in pixel domain

For the block $\tilde{x}_{k,i}$ in $\tilde{x}_k$ (see Fig.3), a best matching block $x_{k-1}^{k,i}$ is selected from the adjacent reference frame x_{k-1} at the criterion of SAD between block $\tilde{x}_{k,i}$ and all the matching blocks from the searching window of a $L \times L$ square region centered at the same space position of $\tilde{x}_{k,i}$. Because $\tilde{x}_{k,i}$ as a good approximation can estimate motion more accurately than $\tilde{y}_{k,i}$ does.

<u>Step1.2</u>: Generate a matching region for $x_{k,i}^r$

Select image block $Win\left(x_{k-1}^{k,i}\right)$, a $W \times W$ square region centered at the best matching block $x_{k-1}^{k,i}$ in x_{k-1} , and select image block $Win\left(\tilde{x}_{k,i}\right)$ with the same size as $Win\left(x_{k-1}^{k,i}\right)$ centered at $\tilde{x}_{k,i}$ in $\tilde{x}_k$, the matching region Mat_reg of $x_{k,i}^r$ is obtained as

$$Mat\_reg = Win\left(x_{k-1}^{k,i}\right) - Win\left(\tilde{x}_{k,i}\right) \tag{6}$$

<u>Step 1.3</u>: Design the hypothesis set H_1^r of $x_{k,i}^r$

The hypothesis set H_1^r consists of blocks extracted by carrying out one-pixel shifts in vertical and horizontal directions in the matching region Mat_reg .

<u>Step 2</u>: Generate hypothesis set H_2^r from other reference frames

The multi-reference-frame pattern in Fig. 2 for its good performance is still extended to the MHRR, and repeat Step 1 for the other reference frames (see Fig.1) to obtain another corresponding hypothesis set H_2^r .

<u>Step 3</u>: Generate hypothesis set H_3^r from the independent recovery $\bar{x}_k^r$ of x_k^r

An independent recovery $\bar{x}_k^r$ of x_k^r that is obtained from y_k^r by SPL[9] is used to generate another hypothesis set H_3^r because the spatial correlation can be fairly exploited from it. H_3^r is constituted by selecting all hypothesis blocks from the $W \times W$ square region centered at the position of $x_{k,i}^r$.

<u>Step 4</u>: Generate the final hypothesis set $H_{k,i}^r$ and conduct optimal MH prediction in measurement-domain

The whole hypothesis set $H_{k,i}^r$ can be generated by merging hypothesis set H_1^r, H_2^r and H_3^r , then the optimal linear prediction weights are calculated in measurement-domain to obtain better prediction residual blocks, because the residual measurements y_k^r are the only real information at reconstruction side.

$$w_{k,i}^r = \underset{w_{k,i}^r}{\arg\min} \left\| y_{k,i}^r - \Phi H_{k,i}^r w_{k,i}^r \right\|_2^2 + \lambda \left\| \Gamma w_{k,i}^r \right\|_2^2 , \tag{7}$$

$$\tilde{x}^r_{k,i} = H^r_{k,i} w^r_{k,i} \tag{8}$$

In equation (7), Γ is the same as equation (3).

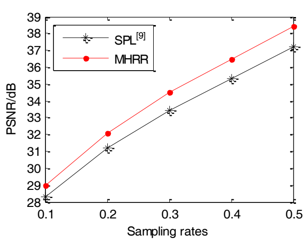
4. SIMULATIONS

In this section, simulation results are reported in two parts. The first part gives the comparison of the proposed MHRR with SPL[9] for 1s_wMRMH[8] which is the first stage MH prediction scheme in [8]. The second part compares the proposed MHRR with both state-of-the-art algorithms Up-Se-AWEN-HHP[7] and 2sMHR[8].

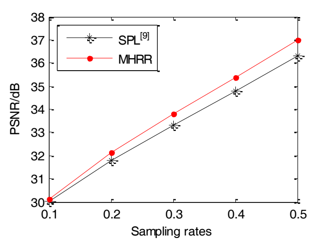
4.1. Comparison with SPL-DDWT for 1s-wMRMH[8]

In this part, measurement-domain MH prediction adopts 1s_wMRMH[8] and all parameters are same as [8]. With regards to residual reconstruction, SPL is used as benchmark and the parameters are same as reference [8]; the proposed MHRR has 5 reference frames at most, the block size b is 16, the size of matching region W is 32 and the size of searching window L is 48, all parameters of SPL in MHRR for recovering y^r_k and y^r_{k,i_1} is same as reference [8].

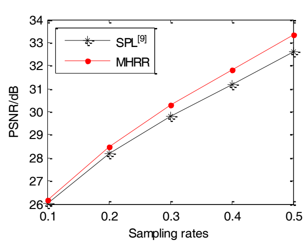
The video sequences *soccer, coastguard, football and foreman* in QCIF (quarter common intermediate format) are used in the simulation, and the first 97 frames (6 GOPs and the key frame of the 7th GOP) of each video sequence are used to test our proposed scheme. Fig. 4 depicts the curves of reconstruction performance (in PSNR) versus subrates.



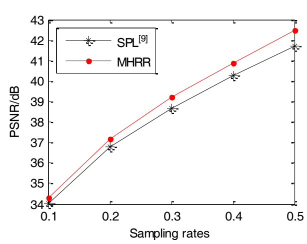
(a) *soccer*



(b) *coastguard*



(c) *football*



(d) *foreman*

Fig. 4. Average PSNR versus sampling rates of different residual reconstruction scheme for 1s_wMRMH[8]

It is obvious that the proposed MHRR gains a better performance over SPL for residual reconstruction in PSNR no matter for slowly moving *coastguard* and *foreman* sequences or fiercely moving *soccer* and *football* sequences, because the proposed MHRR well combines the advantages of pixel-domain ME technique and measurement-domain MH prediction for residual reconstruction, which can exploit temporal and spatial correlation as well as make the second prediction residual sparser to improve last performance.

4.2. Comparison with Up-Se-AWEN-HHP and 2sMHR

In this part, the proposed MHRR is deployed after the first stage 1s_wMRMH in 2sMHR[8], named as 3sMHR. And all parameters of 3sMHR are set as Up-Se-AWEN-HHP[7], the 2sMHR[8] uses the same parameters. The parameters for both MHRR and SPL used after 1s_wMRMH in 2sMHR[8] are same as that subsection 4.1.

The first 88 frames of video sequences *soccer* and *coastguard* in CIF (common intermediate format) are used in the simulation. Table 1 lists the average PSNR versus sampling rates for different residual reconstruction schemes, and the reconstruction performance of Up-Se-AWEN-HHP[7] is obtained from reference [7] directly.

TALBE 1 Comparison of different reconstruction schemes

sequences	methods	sampling rates		
		0.1	0.3	0.5
coastguard	Up-Se-AWEN-HHP	30.86	34.67	37.99
	2sMHR	31.49	35.24	38.57
	3sMHR	31.81	36.09	39.47
foreman	Up-Se-AWEN-HHP	36.17	39.65	42.08
	2sMHR	37.11	40.33	42.54
	3sMHR	37.29	40.52	42.74

It can be observed that 3sMHR obtains higher PSNR than both 2sMHR and Up-Se-AWEN-HHP. For *coastguard* sequence, 3sMHR has 0.85dB and 0.9dB gains at sampling rate 0.3 and 0.5 respectively compared with 2sMHR, as well as has 1.42dB and 1.48dB gains compared with Up-Se-AWEN-HHP. Because the last reconstruction performance of 2sMHR depends on the first stage measurement-domain MH perdition 1s_wMRMH[8] and residual reconstruction, therefore the proposed MHRR improves last reconstruction performance through a novel residual reconstruction scheme. Note that, the proposed MHRR has good compliance and can be employed in the block-based MH prediction strategy to improve the last reconstruction performance.

5. CONLUSIONS

A multihypothesis-based residual reconstruction scheme (MHRR) is presented in this paper. Since the measurement-domain prediction residuals are not always sparse signal or small-energy signal, the residual is first predicted by a novel MH prediction method, which makes the second prediction residual sparser, and then the second prediction residual is reconstructed. The proposed MHRR novelly combines the advantages of pixel-domain ME technique and measurement domain MH prediction, because hypothesis blocks obtained by pixel-domain ME technique are more accurate as well as the measurement-domain MH prediction offers the optimal linear prediction weights by straight using the only real residual measurement at reconstruction side. The proposed MHRR can achieve higher reconstruction performance than SPL, the advance of 3sMHR is up to 0.9dB and 1.48dB at subrate 0.5 respectively compared to 2sMHR and Up-Se-AWEN-HHP.

6. REFERENCES

[1] D.L. Donoho, "Compressed sensing," IEEE Trans. on Information Theory, 52(4): 1289-1306, 2006.
[2] Qaisar S., Bilal R.M., Iqbal W., et al. "Compressive sensing: From theory to applications, a survey", Journal of Communications and Networks, 15(5): 443-456, 2013.
[3] S. Mun and J.E. Fowler, "Residual reconstruction for block-based compressed sensing of video," Data Compression Conference, IEEE, pp. 183-192, 2011.
[4] Tramel E.W. and J.E. Fowler, "Video compressed sensing with multihypothesis," Data Compression Conference, IEEE, pp. 193-202, 2011.
[5] M. Azghani, M. Karimi and F. Marvasti, "Multihypothesis compressed video sensing technique," IEEE Trans. on Circuits and Systems for Video Technology, vol. 26, no. 4, pp. 727-635, 2016.
[6] Chen Jian, Chen Y., and Qin D., et al. "An elastic net-based hybrid hypothesis method for compressed video sensing," Multimedia Tools and Applications, 74(6): 2085-2108, 2015.
[7] Kuo Y., Kai Wu, and Jian Chen, "A scheme for distributed compressed video sensing based on hypothesis set optimization techniques," Multidimensional Systems and Signal Processing, pp. 1-20, 2015.
[8] Ou W.F., Yang C.L., Li W.H., et al. "A two-stage multi-hypothesis reconstruction scheme in compressed video sensing," IEEE International Conference on Image Processing (ICIP), pp. 2494-2498, 2016.
[9] Mun S. and J.E. Fowler, "Block compressed sensing of images using directional transforms," IEEE International Conference on Image Processing, pp. 3021-3024, 2009.
[10] Gan L. "Block compressed sensing of natural images," IEEE International Conference on Digital Signal Processing, pp. 403-406, 2007.