Side Information Extrapolation Using Motion-aligned Auto Regressive Model for Compressed Sensing based Wyner-Ziv Codec

Ran Li, Zongliang Gan, Ziguan Cui, Minghu Wu and Xiuchang Zhu

Jiangsu Province Key Lab on Image Processing & Image Communication,

Nanjing University of Posts and Telecommunications, China

[e-mail: *liran358@163.com*]

*Corresponding author: Xiuchang Zhu

Received November 28, 2012; revised January 18, 2013; accepted February 6, 2013; published February 26, 2013

Abstract

In this paper, we propose a compressed sensing (CS) based Wyner-Ziv (WZ) codec using motion-aligned auto regressive model (MAAR) based side information (SI) extrapolation to improve the compression performance of low-delay distributed video coding (DVC). In the CS based WZ codec, the WZ frame is divided into small blocks and CS measurements of each block are acquired at the encoder, and a specific CS reconstruction algorithm is proposed to correct errors in the SI using CS measurements at the decoder. In order to generate high quality SI, a MAAR model is introduced to improve the inaccurate motion field in auto regressive (AR) model, and the Tikhonov regularization on MAAR coefficients and overlapped block based interpolation are performed to reduce block effects and errors from over-fitting. Simulation experiments show that our proposed CS based WZ codec associated with MAAR based SI generation achieves better results compared to other SI extrapolation methods.

Key words: *Compressed sensing, Wyner-Ziv codec, side information, motion-aligned auto regressive model*

This work was supported by the Graduate Student Innovation Project of Jiangsu Province China (CXZZ12_0466 and CXZZ11_0390), National Natural Science Foundation of China (61071091 and 61271240), Natural Science Foundation of the Higher Education Institutions of Jiangsu Province China (12KJB510019), and Technology Research Program of Hubei Provincial Department of Education (D20121408).

http://dx.doi.org/10.3837/tiis.2013.02.011

1. Introduction

With rapid advances of multimedia communication, new requirements of video applications come to the stage, such as wireless video surveillance, wireless PC camera, mobile video calls, wireless multimedia sensor network, and so on. However, these video applications require a low-cost encoder since the computational power, memory and/or battery capacities are scarce at the encoder. Traditional video encoder based on hybrid coding framework (e.g. H.26x, MEPG) is commonly 5 to 10 times more complex than the decoder on account of performing motion estimation (ME) and motion compensation (MC) to fully exploit redundancy existing in the video. Therefore, the traditional video coding faces many challenges in these new video applications.

In order to alleviate the complexity burden of the encoder, distributed video coding (DVC) [1], which enables to explore the video statistics, partially or totally, at the decoder only, relying on a low encoding complexity, has received wide attention. The mathematics philosophy of DVC was proposed and discussed by Slepian and Wolf [2] according to the Information Theory. The complement of Slepian-Wolf coding for lossy compression is the Wyner-Ziv (WZ) coding [3] which deals with lossy source coding of $\boldsymbol{X}$ with side information (SI) $\boldsymbol{Y}$ at the decoder and can reduce complexity of the encoder by shifting ME and MC to the decoder. The SI is usually interpreted as an attempt made by the decoder to obtain an estimate of the original frame. In the WZ coding scenario, error correcting codes are used to improve the quality of SI until a target quality for the final decoded frame is achieved. One of the most interesting DVC systems is the asymmetric WZ coding scheme [4] proposed by Aaron et al. in which the key frames are encoded by H.263+ intra frame mode and the WZ frames are encoded by Slepian-Wolf codec based turbo codes.

From the above-mentioned content, we can see that performance of DVC system depends on two factors: the first one is the quality of SI, and the second one is the correction capability of WZ coding. Both are quite difficult since original frames are not available at the decoder and the statistics of video source are dynamically varying in spatial and temporal domain.

For the way SI generated, DVC can be categorized into interpolation and extrapolation case. In interpolation case, SI is generated by the interpolating between the previous and following reconstructed WZ/key frames [5]. On the contrary, in the extrapolation case, the SI is generated by referring only the previous reconstructed frame [6-7], similar to the P frame coding in hybrid video coding. Although the SI generated by interpolating has superior performance than that generated by extrapolating since the former can use the future information to generate SI, the extrapolation DVC is very desirable for low latency cases since the decoding process begins as soon as it receives the previous reconstructed frame without waiting for the arrival of the following reconstructed key frame. To improve the compression performance of low-delay DVC, many extrapolation schemes have been

proposed to develop the quality of SI. Natario et al. [7] proposed a robust extrapolation module to generate SI based on translational motion model. In this method, the extrapolation is completed by ME, motion field smoothening, motion projection as well as overlapping and uncovered areas. However, this translational motion model is not always satisfied, especially for video sequences with high motion. Zhang et al. [8] proposed an auto regressive (AR) model to replace the translational motion model and improve the quality of SI. In AR model, the SI of each pixel within the current WZ frame t is generated as a linear weighted summation of pixels within a window in the previous reconstructed WZ/key frame t-1 along the motion trajectory. The method based on AR model regards SI extrapolation as an adaptive filtering problem and implicitly embeds motion information into the filter coefficients, but several unreasonable assumptions on motion trajectory and AR coefficients also lead to appearance of block effects and over-fitting.

After generating SI, the WZ coding is performed to correct some errors existing in SI. The conventional WZ encoder applies a channel code (usually a turbo code or a LDPC code) to the pixels or transform coefficients of the frame, and transmits a portion of the resulting parity bits. The decoder uses the received parity bits to correct errors in SI and controls the bit-rate by a feedback channel. Recently, the appearance of compressed sensing (CS) offers a new idea for WZ coding. CS theory demonstrates that signals which have sparse representation under some transform domain can be sampled at sub-Nyquist rates via linear projection onto a random basis while still enabling exact reconstruction of the original signal, which provides the potential of dramatic reduction of computation complexity in video compression. Based on CS theory, some practical DVC systems have been presented. Do et al. proposed distributed compressed video sensing (DISCOS) [9] to perform WZ coding. In their framework, the key frames are encoded by traditional H.26x intra mode and WZ frames are encoded by CS measurement matrix (e.g. Structurally Random Matrices, SRMs [10]). The WZ decoder utilizes a CS reconstruction algorithm to recover the original frame by exploiting the temporal correlation with neighboring key frame. Liu et al. proposed block based adaptive compressed sensing [11] to allocate CS measurements for every video frame which further improves the performance of WZ coding. Besides, Baig et al. provided CS based WZ coding scheme [12] by incorporating SI generation into the decoder, and the SI generation scheme exploits the correlation between CS measurements of nearby frames. Although these CS based DVC systems can obtain a good quality of decoder frame, they fail to take quantization and entropy coding into account and therefore lose practical engineering significance. In addition, the crucial SI in conventional WZ coding is not fully utilized in CS reconstruction.

To advance the compression performance of low-delay DVC, we improve the correction capability of WZ coding and the quality of SI respectively in this paper. Firstly, a CS based WZ codec is proposed to implement error correction and rate control in the popular DVC architecture proposed by [4]. The proposed WZ encoder uses the same random measurement

matrix to sample every block of the original frame and saves these CS measurements in a buffer after quantization and entropy coding. Depending on the assumed correlation channel model and SI, the decoder detects some blocks having a certain amount of errors in SI and requires the encoder to transmit a portion of CS measurements saved in the buffer by a feedback channel. Then, a specific CS reconstruction algorithm uses these CS measurements to correct the errors in the SI. With accurate prediction of the SI and the high-efficient CS based error-correction algorithm, our WZ codec can effectively reduce the bit rate of a DVC system while enabling exact reconstruction of the original frames. Secondly, in order to obtain higher quality SI in a low delay DVC, we also improve the AR model proposed by [8] and introduce the motion-aligned auto regressive model (MAAR) which appears in [13] to acquire more accurate motion trajectory. The MAAR model can refine the inaccurate motion field in AR model to improve the resulting block effects. In SI extrapolation based on AR model, the AR coefficients are computed by the Least Mean Square (LMS) algorithm. However, without prior knowledge of the AR coefficients, over-fitting happens in the estimated SI. In order to overcome over-fitting, we perform LMS algorithm with smooth constraint to compute MAAR coefficients. The smooth constraint fully utilizes a prior knowledge on similarity between the target pixel value and the training pixel samples which is measured by Tikhonov matrix. Besides, the overlapped block is also proposed to reduce block effects and over-fitting. To verify the performance of the proposed CS based WZ codec using MAAR model based SI extrapolation for low-delay DVC, various experiments are conducted. The simulation results have confirmed that our SI extrapolation is able to achieve SI with much higher accuracy compared with other existing methods, especially for AR model based SI extrapolation, and the proposed CS based WZ codec can also obtain a good error correction capability.

The reminder of this paper is as follows. The overall architecture of the proposed system is first presented in Section 2. Then the CS based WZ codec is described in detail in Section 3. The SI extrapolation using MAAR model is presented in Section 4 followed by the experimental results and analysis in Section 5. Finally the conclusions are provided in the last section.

2. Framework Overview

The block diagram of the low-delay DVC composed of the proposed CS based WZ codec and MAAR model based SI extrapolation is depicted in **Fig.1**. The coding process starts by dividing the input frames into key frames and WZ frames. At the encoder side, the key frames are encoded using the H.264/AVC intra coding scheme. The WZ frames are divided into small blocks and sampled with the same random measurement matrix. After uniform quantization and entropy coding, the bits encoded by CS measurements are stored in the buffer and transmitted in small amount upon decoder request.

At the decoder side, the key frames are decoded using H.264/AVC intra decoding scheme.

For the WZ frames, the SI is first generated by the proposed MAAR model. As shown in **Fig.1**, the SI generation consists of forward and backward MAAR models whose coefficients are computed using similar derivations proposed by [8]. Different from the computing of AR coefficients in [8], we use LMS algorithm with smooth constraint to compute MAAR coefficients. In order to reduce block effects and over-fitting, the overlapped block is used to interpolate SI. Results of the two models are then averaged to generate the final SI. Then the iterative WZ decoder receives an amount of CS measurements by feedback channel to correct the SI errors and generate the WZ decoded frame using the specific CS reconstruction algorithm.

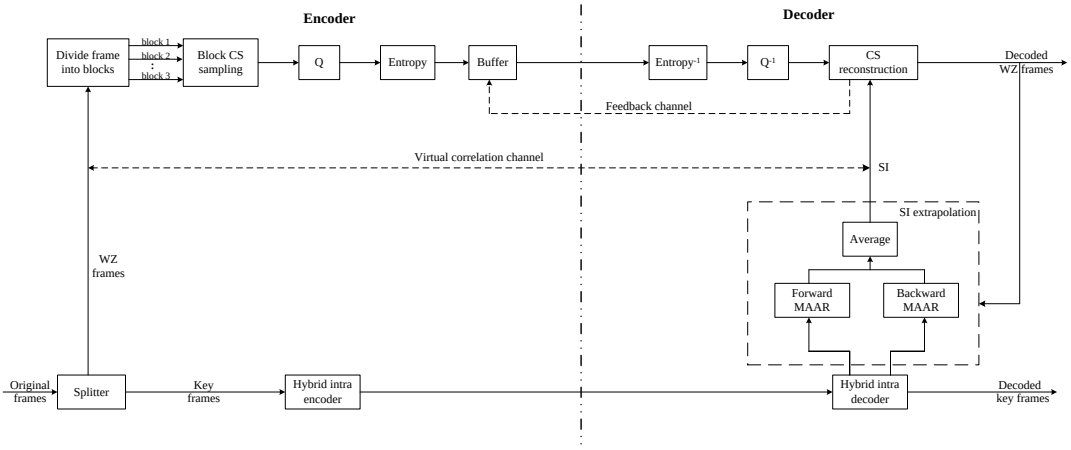


Fig. 1. The block diagram of the low-delay DVC composed of the proposed CS based WZ codec and MAAR model based SI extrapolation

3. WZ Video Codec Based on CS

3.1 Encoding Using Block CS Sampling

At the WZ encoder side, the WZ frame $\boldsymbol{X}$ is divided into small blocks with size $B\times B$ each and sampled with the same measurement matrix. Let $\boldsymbol{x}_i$ represent the vectorized signal of the i-th block through raster scanning. The corresponding output CS vector $\boldsymbol{y}_i$ can be written as

$$\boldsymbol{y}_i = \boldsymbol{A}\boldsymbol{x}_i \tag{1}$$

where $\boldsymbol{A}$ is an $M_B\times B^2$ matrix, and M_B is CS measurements of each block. In general, image has a sparse representation in the known DCT transform domain, and the visual property of human eye determines that the low and medium frequency DCT coefficients are more important than the high frequency ones. Therefore, we only need to sample the low and medium frequency DCT coefficients by a random measurement matrix and perform CS reconstruction algorithm to recovery them. For this purpose, the matrix $\boldsymbol{A}$ consists of two parts: the orthonormalized i.i.d Gaussian matrix $\boldsymbol{\Phi}$ and the special DCT transform matrix $\boldsymbol{\Psi}$ which is used to extract low and medium frequency DCT coefficients of each image block, that is, $\boldsymbol{A}=\boldsymbol{\Phi}\boldsymbol{\Psi}$. As shown in **Fig. 2**, we regard the first 70 DCT coefficients in zigzag order

as the low and medium frequency of each image block with size 16×16. From empirical studies, we suggest block size B = 16 for video sequences hereafter.

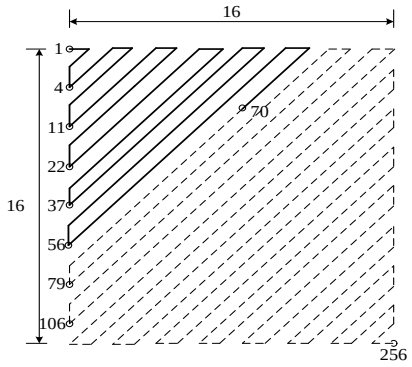


Fig. 2. The low and medium frequency DCT coefficients of each image block with size 16×16

Finally, the CS measurements acquired by block CS sampling are uniformly quantified and encoded into bits by an adaptive arithmetic coder proposed in [14]. These bits will be stored in the buffer and transmitted in small amount upon decoder request.

3.2 Decoding Using CS Reconstruction Algorithm with SI

At the WZ decoder side, the decoder takes previously reconstructed frames to form the SI, $\boldsymbol{Y}$, which is an estimate of original WZ frame $\boldsymbol{X}$. However, there will always be a certain amount of errors in $\boldsymbol{Y}$ since the statistics of video source are dynamically varying in spatial and temporal domain. In our framework, these errors can be corrected by some CS measurements upon decoder request and the process of error recovery is presented in **Fig. 3**. Firstly, we divide the SI $\boldsymbol{Y}$ into blocks with size $B\times B$ and compute error energy E_e of each block $\boldsymbol{B}_{si}$ in SI by the virtual correlation channel which can be modeled using a Laplacian distribution of the difference between original frame and SI. The decoder estimates the Laplacian parameter by observing the statistics from previously decoded frame [15]. Secondly, the decoder is to determine whether error energy E_e of each block is greater than or equal to threshold T. If the current block exceeds the threshold T, then the decoder requires the encoder to transmit K CS measurements to correct these errors and otherwise to skip the current block. Then, in order to control bit rate, the decoder is to detect whether total CS measurements M of the current block is less than the preset upper limit M_{upper} , and reconstructs the current block by using CS recovery algorithm if $M < M_{upper}$ satisfied. There are still some errors existing in the reconstructed block and the next error recovery is performed. Finally, we combine all reconstructed blocks $\boldsymbol{x}_i$ into the decoded WZ frame.

For the CS reconstruction algorithm, we merge the priori knowledge on SI into the reconstructive process in contrast with the traditional CS algorithm using only sparse prior. The optimal model of our CS reconstruction algorithm is as follows,

$$\min_{x_i} \left\| \boldsymbol{\Psi} \boldsymbol{x}_i \right\|_1 ,$$

$$s.t.\ \boldsymbol{y}_i = \boldsymbol{A}\boldsymbol{x}_i = \boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{x}_i, \qquad (2)$$

$$\left\|\boldsymbol{x}_i - \boldsymbol{B}_{si}\right\|_2 \le \varepsilon .$$

In this optimal model, the SI $\boldsymbol{B}_{\mathrm{si}}$ is regarded as a noisy estimation of current block $\boldsymbol{x}_{\mathrm{i}}$, so the solution to (2) exists in the intersection of the l2-ball $\mathbf{P} = \{\boldsymbol{x}_{\mathrm{i}}: \|\boldsymbol{x}_{\mathrm{i}}\text{-}\mathrm{B}_{\mathrm{si}}\|_2 \leqslant \varepsilon\}$ and the hyperplane $\mathbf{H} = \{\boldsymbol{x}_{\mathrm{i}}: \boldsymbol{y}_{\mathrm{i}} = \boldsymbol{A}\boldsymbol{x}_{\mathrm{i}}\}$. However, there are still infinite points in the intersection, and we add a regularization using the priori knowledge that the low and medium DCT coefficients of $\boldsymbol{x}_{\mathrm{i}}$ are sparse to find the optimal solution to (2).

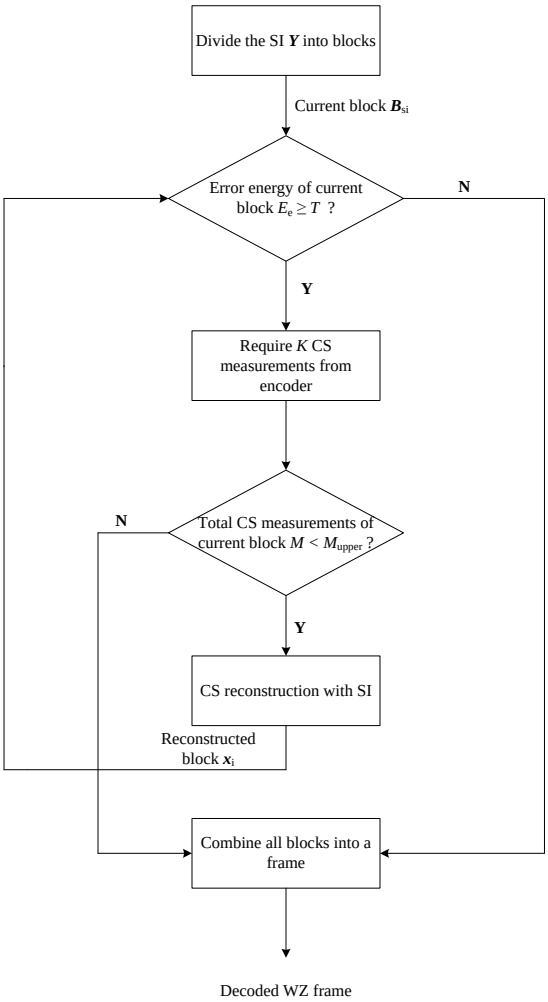


Fig. 3. The flowchart of error recovery for each block in SI

In order to solve the model (2), a variant of projected Landweber (PL) algorithm [16] is proposed as highlighted in **Table 1**. In each iteration, we first compute the distance between the current block at iteration k, $\boldsymbol{x}_{\mathrm{i}}^{(\mathrm{k})}$, and $\boldsymbol{B}_{\mathrm{si}}$ as follows,

$$dist = \left\|\boldsymbol{x}_i^{(k)} - \boldsymbol{B}_{si}\right\|_2 . \qquad (3)$$

To project $\boldsymbol{x}_i^{(k)}$ onto the l2-ball **P**, we simply apply a thresholding operation,

$$\tilde{\boldsymbol{x}}_i^{(k)} = \begin{cases} \boldsymbol{x}_i^{(k)} & dist \le \varepsilon \\ \boldsymbol{B}_{si} + \dfrac{\varepsilon^2}{dist} \boldsymbol{x}_i^{(k)} & dist > \varepsilon \end{cases}. \tag{4}$$

Then, the POCS (Projection onto Convex Set) [17] is performed to find the closest vector $\hat{\boldsymbol{x}}_i^{(k)}$ on **H** for $\tilde{\boldsymbol{x}}_i^{(k)}$,

$$\hat{\boldsymbol{x}}_i^{(k)} = \tilde{\boldsymbol{x}}_i^{(k)} + \boldsymbol{A}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}(\boldsymbol{y}_i - \boldsymbol{A}\tilde{\boldsymbol{x}}_i^{(k)}). \tag{5}$$

Since the low and medium frequency DCT coefficients of the current block are sparse, the hard thresholding [18] is used to enforce their sparsity, and the process of hard thresholding is presented as follows,

$$\boldsymbol{\alpha}_i^{(k)} = \boldsymbol{\Psi}\hat{\boldsymbol{x}}_i^{(k)}, \tag{6}$$

$$\hat{\boldsymbol{\alpha}}_i^{(k)} = \begin{cases} \boldsymbol{\alpha}_i^{(k)} & \left|\boldsymbol{\alpha}_i^{(k)}\right| \le \lambda\sigma^{(k)}\sqrt{2\log L} \\ 0 & else \end{cases}, \tag{7}$$

$$\bar{\boldsymbol{x}}_i^{(k)} = \boldsymbol{\Gamma}\hat{\boldsymbol{\alpha}}_i^{(k)}, \tag{8}$$

where λ is a constant control factor to mange convergence, L is the number of the transform coefficients, and $\sigma^{(k)}$ is estimated using robust median estimator,

$$\sigma^{(k)} = \frac{median(\left|\boldsymbol{\alpha}_i^{(k)}\right|)}{0.6745}. \tag{9}$$

The function of matrix $\boldsymbol{\Gamma}$ is to reposition the shrunk low and medium frequency DCT coefficients and transform DCT coefficients into pixel domain. Finally, we do the POCS again to ensure that the estimated block at iteration k+1, $\boldsymbol{x}_i^{(k+1)}$ exists in the hyperplane **H**. In addition, we initialize with $\boldsymbol{x}_i^{(0)} = \boldsymbol{B}_{si}$ and terminate when $|D^{(k+1)}\text{-}D^{(k)}| < Tol$, where

$$D^{(k)} = \frac{1}{B}\left\|\hat{\boldsymbol{x}}_i^{(k-1)} - \boldsymbol{x}_i^{(k)}\right\|_2. \tag{10}$$

We analyze the complexity of the proposed CS algorithm as follows. At each iteration, the thresholding operation which projects $\boldsymbol{x}_i^{(k)}$ onto the l2-ball **P** only requires $O(B^4)$ operations. The computational cost of POCS is essentially the cost of applying $\boldsymbol{A}$ and its transpose $\boldsymbol{A}^{\mathrm{T}}$**.** (If necessary, $(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}})^{-1}$ can be computed beforehand), and applying the $M \times B^2$ measurement matrix $\boldsymbol{A}$ will require $O(MB^2)$ operations in general. For the process of hard thresholding, the major amount of computation is the cost of DCT and IDCT transform which require $O(LB^2)$ operations, if we use fast transformation of DCT and IDCT, the complexity can be reduced to $O(L)$.

Table 1. Summary of the proposed CS reconstruction algorithm with SI

Algorithm model : $\min_{x_i} \|\boldsymbol{\Psi x}_i\|_1, \ s.t. \ \boldsymbol{y}_i = \boldsymbol{A x}_i = \boldsymbol{\Phi \Psi x}_i, \ \|\boldsymbol{x}_i - \boldsymbol{B}_{si}\|_2 \le \varepsilon$

Input: Initial solution $\boldsymbol{x}_i^{(0)} = \boldsymbol{B}_{si}$; CS measurements $\boldsymbol{y}_i$; Measurement matrix $\boldsymbol{\Phi}$; Low and medium frequency DCT matrix $\boldsymbol{\Psi}$; Error tolerance ε; Termination threshold *Tol*; Control factor λ; Maximum number of iterations s_{max}.

Output: Reconstructed block $\boldsymbol{x}_i$.

for $k = 1$ to s_{max} **do**

Compute *dist* according to Eq (3);

if $dist \le \varepsilon$

$\tilde{\boldsymbol{x}}_i^{(k)} \Leftarrow \boldsymbol{x}_i^{(k)}$;

else

$\tilde{\boldsymbol{x}}_i^{(k)} \Leftarrow \boldsymbol{B}_{si} + \frac{\varepsilon^2}{dist} \boldsymbol{x}_i^{(k)}$;

end if

Compute $\hat{\boldsymbol{x}}_i^{(k)}$ according to Eq (5);

Compute $\overline{\boldsymbol{x}}_i^{(k)}$ by doing hard thresholding according to Eq (6)-(9);

Compute $\boldsymbol{x}_i^{(k+1)}$ and $D^{(k+1)}$ according to Eq (5) and Eq (10);

if $|D^{(k+1)}-D^{(k)}| < Tol$

break;

end if

end for

4. SI Extrapolation Using MAAR Model

The SI extrapolation based on AR model proposed by [8] is able to achieve SI with much higher accuracy by using three assumptions which are as follows:

1) All the pixels within each block in SI share the same AR coefficients on account of the piecewise stationary characteristics of the frame;

2) Each block in SI and its co-located block in previous reconstructed frame obey the same motion trends;

3) The same AR coefficients are used to interpolate the block in SI and its co-located block in previous reconstructed frame along the motion trajectory within adjacent frames.

Owing to the fact that the computation of AR coefficients depends on a high similarity along the motion trajectory, it is important to estimate accurate motion field of SI. However, the second assumption is obviously unreasonable since some differences exist between

motion fields of two continuous frames, especially for the video sequences with high motions. Therefore, we should predict reasonably the motion field of SI. In this section, we will introduce MAAR model to overcome the unreasonable assumption on motion trajectory in AR model. Besides, in order to obtain the more accurate AR coefficients, the LMS algorithm with smooth constrain and the interpolation using overlapped block are proposed to reduce block effects and over-fitting.

4.1 MAAR Model Description

In MAAR model, the SI of each pixel within the current WZ frame t is generated as a weighted summation of the pixels within a particular window in the previous reconstructed WZ/K frame t-1 as shown in **Fig. 4**. Let $\boldsymbol{X}_t$ be the current WZ frame at t, and $\boldsymbol{Y}_t$ be the SI of $\boldsymbol{X}_t$. For each pixel in $\boldsymbol{X}_t$, the window, indicated by the circles and the red arrow at frame $\boldsymbol{X}_{t\text{-}1}$, is determined by the integer-pixel accuracy motion field which is estimated as follows:

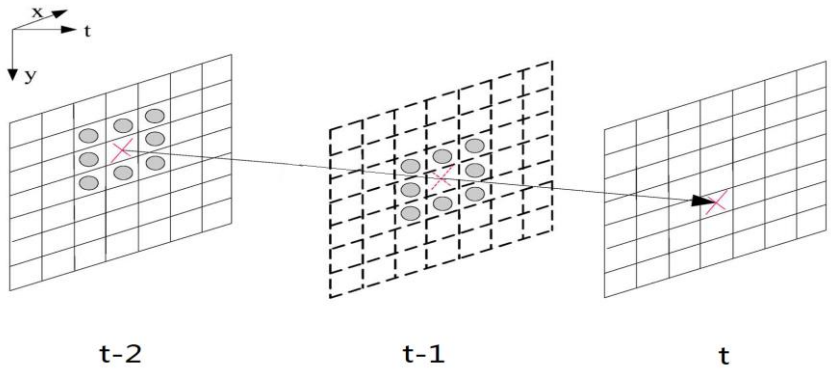


Fig. 4. The MAAR model using forward derivation

1) Motion vectors are estimated for each block in the reconstructed frame $\boldsymbol{X}_{t\text{-}1}$ by taking the reconstructed frame $\boldsymbol{X}_{t\text{-}2}$ as reference. To overcome the problem that a small block size is likely to obtain inaccurate motion vectors if common block-matching algorithm (BMA) is used, we use overlapped block motion estimation (OBME). For each $b\times b$ block in $\boldsymbol{X}_t$, we firstly enlarge the block size to $3b/2\times 3b/2$, and then use the enlarged block to find the best matching block in search window with size $b\times b$ and give the computed motion vector to the $b\times b$ block.

2) For each block, a new motion vector is calculated by the weighted median vector filter [19]. This leads to a smoothed motion vector field where true motion is capture.

3) The pixel from the reconstructed frame $\boldsymbol{X}_{t\text{-}1}$ are projected to the next time instant using the motion field obtained above assuming that the motion is linear and that, therefore, the warping of frame $\boldsymbol{X}_{t\text{-}2}$ into $\boldsymbol{X}_{t\text{-}1}$ will linearly continue from frame $\boldsymbol{X}_{t\text{-}1}$ to frame $\boldsymbol{X}_t$.

4) Since the motion vectors obtained do not necessarily intercept the frame $\boldsymbol{X}_t$ at the center of each non-overlapped block in the frame $\boldsymbol{X}_t$, it is possible for the frame $\boldsymbol{X}_t$ to have no or multiple motion trajectories on some blocks. In order to assign a single motion vector for each non-overlapped block in the frame $\boldsymbol{X}_t$, the motion vector nearest to the center of the

non-overlapped block is selected from the available candidate vectors obtained in the previous step.

After the determination of the window, the weighted summation is performed as

$$\boldsymbol{Y}_t(m,n) = \sum_{-R\le(i,j)\le R} \boldsymbol{X}_{t-1}(\tilde{m}+i,\tilde{n}+j)\cdot\alpha(i,j)\,. \tag{11}$$

Here $\boldsymbol{Y}_t(m,n)$ represents the SI of the pixel located at (m,n), $(\tilde{m},\tilde{n})$ represents the corresponding integer-pixel position in $\boldsymbol{X}_{t-1}$ determined by the motion vector of $\boldsymbol{X}_t(m,n)$, and $\alpha(i,j)$ is the forward MAAR coefficient from frame $\boldsymbol{X}_{t-1}$ to frame $\boldsymbol{X}_t$. R is defined to be the radius of the window, the size of which is $(2R+1)\times(2R+1)$. Due to the high similarity along the motion trajectory, we assume that all pixels in motion trajectory from frame $\boldsymbol{X}_{t-2}$ to frame $\boldsymbol{X}_t$ are computed using the same MAAR coefficients, that is,

$$\boldsymbol{X}_{t-1}(\tilde{m},\tilde{n}) = \sum_{-R\le(i,j)\le R} \boldsymbol{X}_{t-2}(\bar{m}+i,\bar{n}+j)\cdot\alpha(i,j)\,, \tag{12}$$

where $(\bar{m},\bar{n})$ represents the corresponding integer-pixel position in $\boldsymbol{X}_{t-2}$ determined by the motion vector of $\boldsymbol{X}_{t-1}(\tilde{m},\tilde{n})$ which is also estimated using OBME.

4.2 Computation of MAAR Coefficients and Overlapped Block Interpolation

Obviously, the coefficient estimation plays a critical role for the quality of SI generated by the proposed MAAR model. Since there is no access to the actual pixel in the current WZ frame $\boldsymbol{X}_t$ at the decoder, the previous two continuous frames $\boldsymbol{X}_{t-1}$ and $\boldsymbol{X}_{t-2}$ are fully utilized to estimate MAAR coefficients. A forward derivation is proposed as shown in **Fig. 5**. According to the estimated motion field of $\boldsymbol{X}_t$, for each block $\boldsymbol{B}_t(k,l)$ located at position (k,l) within $\boldsymbol{X}_t$, we find its matching block $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$ in the previous reconstructed WZ/key frame $\boldsymbol{X}_{t-1}$. Then, we find the best matching block $\boldsymbol{B}_{t-2}(\bar{k},\bar{l})$ for $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$ in the reconstructed WZ/key frame $\boldsymbol{X}_{t-2}$. Depending on the high similarity along the motion trajectory, the forward MAAR coefficients for interpolating $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$ as the linear combination of pixels in $\boldsymbol{B}_{t-2}(\bar{k},\bar{l})$ are same with those for interpolating $\boldsymbol{B}_t(k,l)$ as the linear combination of pixels in $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$. Therefore, the forward MAAR coefficients can be calculated by using known $\boldsymbol{B}_{t-2}(\bar{k},\bar{l})$ and $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$. Due to the piecewise stationary characteristics of the frame, all the pixel within block $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$ can be predicted using the same MAAR coefficients as follows,

$$\boldsymbol{B}_{t-1} = \boldsymbol{C}_{t-2}\boldsymbol{\alpha} + \boldsymbol{n} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1N} \\ c_{21} & c_{22} & \cdots & c_{2N} \\ \vdots & \vdots & \cdots & \vdots \\ c_{S1} & c_{S2} & \cdots & c_{SN} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_N \end{bmatrix}. \tag{13}$$

Here, $\boldsymbol{B}_{t-1}$ represents the vectorized signal of $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$ through raster scanning, S denotes the number of pixels within $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$, N is the number of the MAAR coefficients, that is, $N = (2R+1)\times(2R+1)$, $\boldsymbol{n}$ is the additive white Gaussian noise, and the $(2R+1)\times(2R+1)$ window of each pixel within $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$ is packed into a $1\times N$ row vector, then a matrix $\boldsymbol{C}_{t-2}$ sized $S\times N$ is obtained. The best coefficients can be computed by LMS criterion, which can be described as

$$\boldsymbol{\alpha} = \arg\min_{\boldsymbol{\alpha}}\{\| \boldsymbol{B}_{t-1} - \boldsymbol{C}_{t-2}\boldsymbol{\alpha} \|_2^2\}. \tag{14}$$

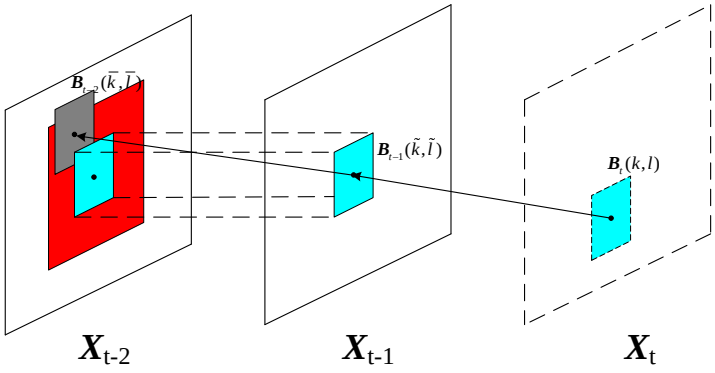


Fig. 5. Coefficient approximation illustration

However, without prior knowledge of the MAAR coefficients $\boldsymbol{\alpha}$, the model (14) often produces over-fitting. To reducing the bad effects caused by over-fitting, the most common approach is to regularize the LMS model using Tikhonov regularization which imposes an l2 penalty on the norm of $\boldsymbol{\alpha}$, that is,

$$\boldsymbol{\alpha} = \arg\min_{\boldsymbol{\alpha}}\{\| \boldsymbol{B}_{t-1} - \boldsymbol{C}_{t-2}\boldsymbol{\alpha} \|_2^2 + \mu\|\boldsymbol{\Lambda}\boldsymbol{\alpha}\|_2^2\}, \tag{15}$$

where $\boldsymbol{\Lambda}$ is known as the Tikhonov matrix [20]. The $\boldsymbol{\Lambda}$ term allows the imposition of prior knowledge on the solution $\boldsymbol{\alpha}$. In our case, we can exploit the approach that the pixels within the window determined by the motion vector of block $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$ which are the most similar from the target pixel should be given larger weight than ones which are the most dissimilar. Therefore, we propose the diagonal $\boldsymbol{\Lambda}$ in the form of

$$\boldsymbol{\Lambda} = diag(\|\boldsymbol{B}_{t-1} - \boldsymbol{c}_1\|_2, \|\boldsymbol{B}_{t-1} - \boldsymbol{c}_2\|_2, \cdots, \|\boldsymbol{B}_{t-1} - \boldsymbol{c}_N\|_2), \tag{16}$$

where $\boldsymbol{c}_i$ ($i = 1 \cdots N$) is column vector of matrix $\boldsymbol{C}_{t-2}$. With this structure, $\boldsymbol{\Lambda}$ penalizes weights of large magnitude assigned to the pixels which have a significant distance from the target pixel. For the block $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$, then, the MAAR coefficients $\boldsymbol{\alpha}$ can be calculated directly by the usual Tikhonov solution,

$$\boldsymbol{\alpha} = (\boldsymbol{C}_{t-2}^{\mathrm{T}}\boldsymbol{C}_{t-2} + \mu\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\Lambda})^{-1}\boldsymbol{C}_{t-2}^{\mathrm{T}}\boldsymbol{B}_{t-1}, \tag{17}$$

where we use $\mu = 1.5$ from this point on.

After computing MAAR coefficients of each block in WZ frame $\boldsymbol{X}_t$, the interpolation is performed block by block. Whereas block edges may not always be consistent with the heterogeneous object edges, and thus block effects are usually perceived in regions where

one block has a significantly different motion compared with its neighbors. In order to reduce block effects, the overlapped block is introduced to perform interpolation. As illustrated in **Fig. 6**, for each $b\times b$ block in $\boldsymbol{X}_t$, we enlarge the block size to $2b\times 2b$ and compute MAAR coefficients of the enlarged block. When interpolating the SI of $\boldsymbol{X}_t$, the four region A, B,C and D in each enlarged block overlap the neighboring blocks, e.g. the region A overlaps the top left four neighboring blocks V1, V2,V3 and V4. Therefore, each pixel in the enlarged block has the four candidate estimates, and we get the final pixel value by uniformly averaging the four candidate estimates. Another advantage of using the overlapped block is that it can reduce the bad effect of over-fitting due to increasing the number of training pixel samples when computing MAAR coefficients.

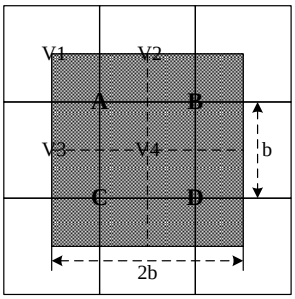


Fig. 6. Illustration of overlapped block

In addition to using forward derivation to calculate the MAAR coefficients, a backward derivation can also be performed to compute the MAAR coefficients. The backward derivation comprises the following steps. Firstly, the optimal backward MAAR coefficients $\boldsymbol{\beta}$ for interpolate $\boldsymbol{B}_{t-2}(\bar{k},\bar{l})$ as the linear combination pixels in $\boldsymbol{B}_{t-1}(\tilde{k},\tilde{l})$ can be derived by the way similar to the forward derivation case. Secondly, we exploit the centrosymmetric property between the forward derivation and backward derivation proposed in [8] to derive another approximated forward MAAR coefficient set to predict $\boldsymbol{B}_t(k,l)$, that is,

$$\boldsymbol{\beta}'(i,j)=\boldsymbol{\beta}(-i,-j). \tag{18}$$

Here $\boldsymbol{\beta}'(i,j)$ is the corresponding rearranged forward coefficient from frame $\boldsymbol{X}_{t-2}$ to frame $\boldsymbol{X}_{t-1}$. Finally, replacing $\boldsymbol{\alpha}(i,j)$ with $\boldsymbol{\beta}'(i,j)$ in Eq. (11), we can get another $\boldsymbol{Y}_t(m,n)$. Note that the overlapped block is still used to interpolate the SI. In order to fully capture the different properties of forward and backward derivation based on MAAR model, the final SI is obtained by uniformly averaging the two results of forward and backward derivation.

5. Experimental Results and Analysis

We have conducted various experiments in this section to evaluate the performance of the proposed CS based WZ codec using MAAR model based SI generation. The two CIF@30Hz video sequences *Foreman* and *Bus* are selected as the test sequences and the key frames of the test sequences are encoded by the intra-frame encoder in H.264/AVC reference software version JM 12.4. In the first subsection, the MAAR model based SI extrapolation

performance is compared with other existing methods. In the second subsection, we evaluate the correction capability of the proposed CS based WZ codec. In the last subsection, we study the rate-distortion performance of the proposed low-delay DVC systems.

5.1 Evaluation of SI

In order to evaluate our MAAR model based SI generation (MAAR_avg), we use the two continuous key frames to interpolate the SI and the QPs of the key frames are set to be 28. The comparison group includes two other SI extrapolation methods: conventional motion-compensated extrapolation (MCE) proposed by [7] and AR model with the probability based fusion (AR_Fusion) proposed by [8]. Note that AR_Fusion represents the fusion results by applying the fusion method on the interpolation by forward derivation, the interpolation by backward derivation and the conventional MCE, and the block size b of the above three methods is set to be 8, and R represents the radius of AR and MAAR models used to generate the SI for the corresponding sequence.

(a) (b) (c) (d)

Fig.7. Results of SI extrapolation for *Foreman* (13th frame). (a) Original; (b) MCE, PSNR = 27.44 dB; (c) AR_Fusion, R = 2, PSNR = 29.34 dB; (d) MAAR_avg, R = 2, PSNR = 29.97 dB.

(a) (b) (c) (d)

Fig.8. Results of SI extrapolation for *Bus* (18th frame). (a) Original; (b) MCE, PSNR = 20.19 dB; (c) AR_Fusion, R = 1, PSNR = 24.58 dB; (d) MAAR_avg, R = 1, PSNR = 25.83 dB.

The SI quality comparison is shown in **Figs. 7-8** for the CIF sequence *Foreman* and *Bus*. Firstly, we observe 13th frame of *Foreman* as shown in **Fig. 7**. It can be seen from **Fig. 7(b)**

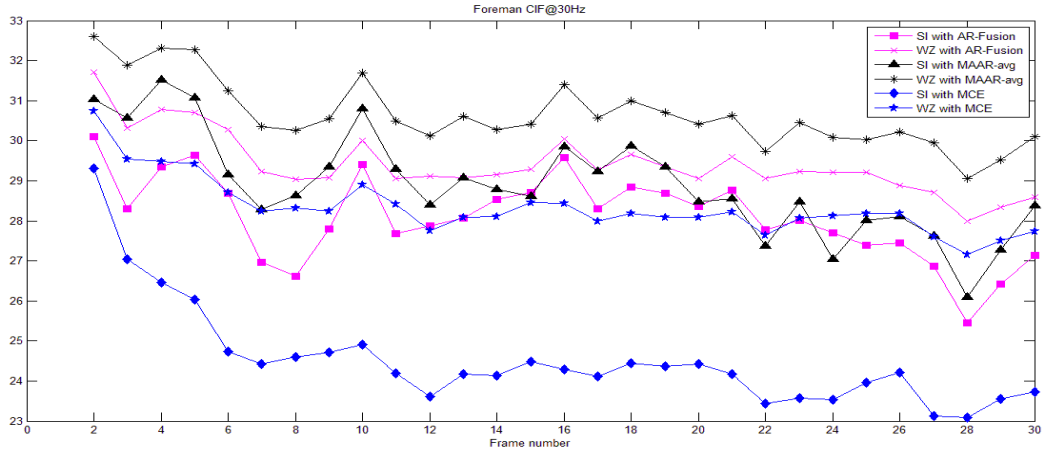
that the SI estimated by MCE contains many cracks. This is main reason that the pixels in uncovered areas are predicted by local spatial interpolation. In **Fig. 7(c)**, the block effects (highlighted in all red circles) and spoiled points caused by over-fitting (highlighted in bottom left and right red circles) appear in edges of object, while the proposed MAAR_avg method recovers them gracefully. Besides, when MAAR_avg is used, the PSNR gain can be up to 2.53 dB and 0.63 dB compared to the MCE and AR_Fusion. Therefore, depending on the superiority that the reasonable regularization on MAAR coefficients and overlapped block interpolation, our MAAR model based SI extrapolation can effectively reduce block effects and bad effects caused by over-fitting. Secondly, for 18th frame of *Bus* as shown in **Fig. 8**, the apparent errors in the "barrier" and cracks can be observed in the SI predicted by MCE as shown in **Fig. 8(b)**. For the SI interpolated by AR_Fusion in **Fig. 8(c)**, the inaccurate motion field leads to object displacement (highlighted in all red circles), however, the proposed MAAR_avg effectively overcomes this problem and acquires a better subjective quality. In terms of objective evaluation, the MAAR_avg also gains up to 5.64 dB and 1.25 dB over MCE and AR_Fusion. These can verity that our MAAR model can obtain the more accurate motion field than MCE and AR model.

5.2 Evaluation of WZ Codec

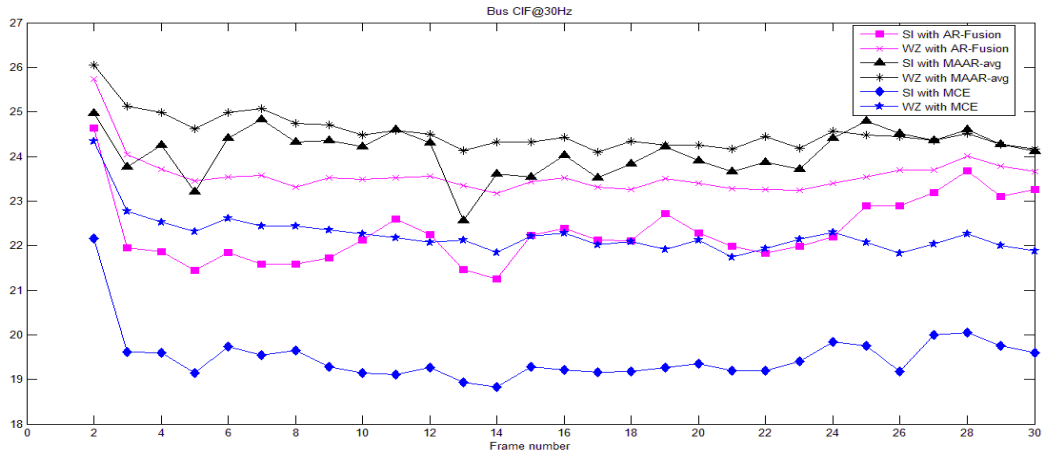
To better illustrate the performance of our proposed CS based WZ codec, we present the PSNR of each decoded WZ frame using the above-mentioned SI generations in **Fig. 9**, where the QPs of the key frames are set to be 28. The WZ frame frequency is set to one WZ frame every two frames and the first WZ frame is encoded by H.264/AVC intra-frame encoder in order to derive the motion information of the second WZ frame for its SI generation. It is noted that the block size b of SI generations is set to be 8, CS measurements K required once and the upper limit M_{upper} of each block is 5 and 35, and control factor λ, error tolerance ε, termination threshold Tol and maximum number of iterations $s_{\max}$ are respectively set to be 0.6, 0.5, 0.001 and 200 for CS reconstruction. It can be seen that when our proposed WZ codec is applied, the PSNR of each SI gets improved. Among the three SI generation methods, the MCE obtains the most significant PSNR gain, since our WZ codec has high-efficient correction capability and can substantially raise the quality of SI generated by the poorer method. With the improved quality of SI, although the PSNR gain has tapered off, the bit rate is decreased, as shown in **Table 2**. This is the main reason that our WZ codec can adaptively control the bit rate according to the varying SI quality. If the SI has lots of errors, our WZ codec will transmit the more bits to correct them, otherwise it reduce the bit rate to alleviate the burden of channel.

Table 2 also summaries average PSNR of decoded WZ frames when the different QPs of key frames are used. We can observe that regardless of SI generations and QPs, the proposed CS based WZ codec can effectively improve the quality of SI, e.g. when AR_Fusion is used and QP is 26, the average PSNR gains can be up to 1.29 dB compared to the SI for *Foreman*. In addition, the CS based WZ codec associated with MAAR_avg also

acquire the highest average PSNR and the lowest bit rate for the decoded WZ frames. Therefore, the low-delay DVC system composed by our WZ codec and SI generation MAAR_avg can get the best performance, since it elegantly integrates the better error-correction capability and high quality SI by flexibly controlling bit rate and efficiently improving prediction accuracy of SI generation.



(a)



(b)

Fig. 9. PSNR of each decoded WZ frame by the proposed CS based WZ codec with different SI generations. (a) *Foreman* (R =2); (b) *Bus* (R = 1).

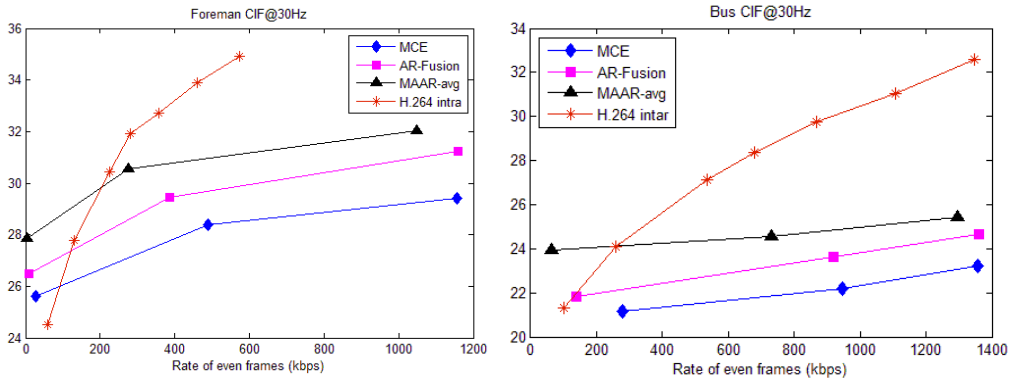
Table 2. Evaluation of decoded WZ frames when various SI generations are used and the key frames are encoded under different QPs by H.264/AVC intra encoder.

Foreman (R=2)				*Bus* (R=1)			
SI generations	PSNR of SI (dB)	PSNR of WZ (dB)	Bit Rate (kbps)	SI generations	PSNR of SI (dB)	PSNR of WZ (dB)	Bit Rate (kbps)
QP = 26				QP = 26			
MCE	24.69	28.37	489.18	MCE	19.49	22.21	945.91
AR_Fusion	28.16	29.45	387.18	AR_Fusion	22.36	23.62	918.26

MAAR_avg	**28.92**	**30.55**	**276.13**	MAAR_avg	**24.13**	**24.56**	**726.79**
QP = 28				QP = 28			
MCE	24.51	28.33	492.80	MCE	19.50	22.25	952.04
AR_Fusion	28.08	29.41	390.08	AR_Fusion	22.32	23.58	929.60
MAAR_avg	**28.90**	**30.65**	**268.74**	MAAR_avg	**24.09**	**24.54**	**739.76**
QP = 30				QP = 30			
MCE	24.47	28.39	498.82	MCE	19.53	22.21	961.92
AR_Fusion	28.03	29.39	393.95	AR_Fusion	22.22	23.57	956.27
MAAR_avg	**28.75**	**30.62**	**284.47**	MAAR_avg	**24.00**	**24.52**	**759.51**

5.3 Rate-distortion Performances

Fig. 10 compares the rate-distortion performances of the proposed CS based WZ codec with various SI generations and the intra coded results of H.264/AVC reference software version JM 12.4. It is noted that we only consider the even frames. From **Fig. 10**, it is observed that our MAAR model obtains the higher performance than the two other SI generations when using CS based WZ codec. In addition, although the CS based WZ codec have inferior performance compared with H.264/AVC intra encoder at high bit rates, the proposed MAAR model is able to reduce the gap between them and even superior to H.264/AVC intra encoder at low bit rates. This is because the propose MAAR interpolation has the superior ability of predicting the future data based on its accurate motion field and reasonable regularization on the overlapped block interpolation coefficients. Therefore, the DVC system composed of CS based WZ codec and MAAR based SI generation is desirable for low latency cases.



(a) (b)

Fig. 10. Rate-distortion curves for H.264/AVC intra and the proposed CS based WZ codec with various SI generations. (a) *Foreman*; (b) *Bus*.

6. Conclusions

In this paper, we propose CS based WZ codec and the MAAR based SI generation in low-delay DVC. In the proposed WZ codec, different from the conventional WZ codec, the

channel coding is replaced by block based CS to correct the errors existing in the SI. At the WZ encoder side, the WZ frame is divided into small blocks and their low and medium frequency DCT coefficients are sampled with the same Gaussian measurement matrix. In order to effectively improve the quality of SI, a scheme for correcting errors in the SI is proposed at decoder. Firstly, depending on the estimated virtual correlation channel, when some blocks containing a certain amount of errors in SI are detected and a requirement is send to the encoder through a feedback channel in the decoder, and corresponding CS measurements are transmitted from the encoder buffer to the decoder to correct these errors. Then, a CS reconstruction algorithm is proposed to recover errors by using CS measurements and crucial SI. To obtain high quality SI, we improve the AR model based SI generation and introduce MAAR model to refine the inaccurate motion field appearing in the AR model. Besides, in order to reduce block effects and bad effects of over-fitting in the AR model, Tikhonov regularization using the priori knowledge on similarity between the target pixel and the corresponding training pixel samples and the interpolation using the overlapped block are performed in our MAAR model. Simulation experiments show that our MAAR based SI generation achieves better results compared to other SI extrapolation methods in terms of both subjective and objective performance, and the proposed CS based WZ codec can effectively improve the quality of SI and obtain a good error correcting capability.

Reference

[1] Bernd Girod, Anne Aaron, Shantanu Rane and David Rebollo-Monedero, "Distributed video coding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 71-83, January, 2005. Article (CrossRef Link)

[2] J. Slepian and J.Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Inform. Theory*, vol. 19, no. 4, pp. 471-480, July, 1973. Article (CrossRef Link)

[3] A. Wyner and J. Ziv, '"The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Inform. Theory*, vol. 22, no. 1, pp. 1-10, January, 1976. Article (CrossRef Link)

[4] A. Araon, R. Zhang and B. Girod, "Wyner-Ziv coding of motion video," in *Proc. of the Asilomar Conference on Signals and Systems*, pp. 240-244, November 3-6, 2002. Article (CrossRef Link)

[5] A. Aaron, D. Varodayan and B. Girod, "Wyner-Ziv residual coding of video," in *Proc. of the International Picture Coding Symposium*, pp. 1-5, April, 2006. Article (CrossRef Link)

[6] A. B. B. Adikari, W. A. C. Fernando, H. Kodikara Arachchi and W. A. R. J. Weerakkody, "Sequential motion estimation using luminance and chrominance informance for distributed video coding of Wyner-Ziv frames," *Electron. Lett.*, vol. 42, no. 7, pp. 398-399, 2006. Article (CrossRef Link)

[7] L. Natrio, C. Brites, J. Ascenso and F. Pereira, "Extrapolating side information for low-delay pixel-domain distributed video coding," in *Proc. of International Workshop on Very Low Bitrate Video Coding*, pp.16-21, September, 2005. Article (CrossRef Link)

[8] Yongbing Zhang, Debin Zhao, Hongbin Liu, Yongpeng Li, Siwei Ma and Wen Gao, "Side information generation with auto regressive model for low-delay distributed video coding," *J. Vis. Commun. Image R.*, vol.23, no. 1, pp. 229-236, January, 2012. Article (CrossRef Link)

[9] T.T. Do, Y. Chen, D. T. Nguyen, N. Nguyen, L. Gan and T. D. Tran, "Distributed compressed video sensing," *in Proc. of the International Conference on Image Processing*, pp. 1393-1396, November, 2009. Article (CrossRef Link)

[10] Do T, Gan L, Nguyen N and Tran T, "Fast and efficient compressive sensing using structurally random matrices," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 139-154, January, 2012. Article (CrossRef Link)

[11] Zhaorui Liu, A. Y. Elezzabi and H. Vicky Zhao, "Maximum frame rate video acquisition using adaptive compressed sensing," *IEEE Trans. on Cir. And Syst. Video Technol*, vol. 21, no. 11, pp. 1704-1717, November, 2011. Article (CrossRef Link)

[12] Yousuf Baig, Edmund M-K. Lai and Amal Punchihewa, "Distributed video coding based on compressed sensing," *in Proc. of 2012 IEEE International Conference on Multimedia and Expo Workshops*, pp. 325-330, July 9-13, 2012. Article (CrossRef Link)

[13] Yongbing Zhang, Debin Zhao, Siwei Ma, Ronggang Wang and Wen Gao, "A Motion-aligned auto-regressive model for frame rate up conversion," *IEEE Trans. on Image Processing*, vol. 19, no. 5, pp. 1248-1258, May, 2010. Article (CrossRef Link)

[14] Skretting K, "Arithmetic Coding and Huffman Coding in MATLAB," 2001, [Online]. Available: www.ux.uis.no/~karlsk/proj99/index.html Article (CrossRef Link)

[15] Catarina Brites and Frenando Pereira, "Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding," *IEEE Trans. on Cir. And Syst. Video Technol*, vol. 18, no. 9, pp. 1177-1190, September, 2008. Article (CrossRef Link)

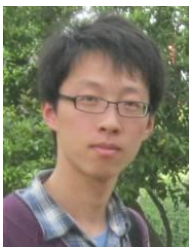
[16] Sungkwang Mun and J. E. Fower, "Block compressed sensing of images using directional transforms," in *Proc. of the International Conference on Image Processing*, pp. 3021-3024, November, 2009. Article (CrossRef Link)

[17] Lu Gan, "Block compressed sensing of natural images," in *Proc. of the International Conference on Digital Signal Processing*, pp. 403-406, July, 2007. Article (CrossRef Link)

[18] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. on Information Theory*, vol. 41, no. 3, pp. 613-627, May, 1995. Article (CrossRef Link)

[19] L. Alparone, M. Barni, F. Bartolini and V. Cappellini, "Adaptively weighted vector-median filters for motion fields smoothing," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 2267-2270, May, 1996. Article (CrossRef Link)

[20] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, V. H. Winston & Sons, Washington, D. C.. 1977. Article (CrossRef Link)



Ran Li is currently a Ph.D. candidate at Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests include compressed sensing, distributed video coding and image communications.

Zongliang Gan received his M.S. degree and the Ph.D. degree in signal and information processing from Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu, China, in 2004 and 2007 respectively. He is currently a lecturer in the Nanjing University of Posts and Telecommunications. His research interests include distributed video coding, image super-resolution reconstruction and image communications.

Ziguan Cui received his Ph.D. degree in signal and information processing from Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu, China, in 2011. He is currently a lecturer in the Nanjing University of Posts and Telecommunications. His research interests include video coding and image communications.

Minghu Wu is currently a Ph.D. candidate at Nanjing University of Posts and Telecommunications, Nanjing, China. He is also an associate professor in the Hubei University of Technology. His research interests include compressed sensing and distributed video coding.

Xiuchang Zhu received his B.S. and M.S. degrees from Nanjing University of Posts and Communications in 1982 and 1987, respectively. He has been working in Nanjing University of Posts and Communications since 1987. At present, he is a Professor and the direct of Jiangsu Key Library of Image Processing and Image Communications. His current research interests focus on multimedia information, especially on the collection, processing, transmission and display of image and video.