Block boundary matching algorithm for generating side information in distributed video codec

Kwang-Hyun Choi
Jae-Yung Lee
Byeung-Woo Jeon
Jong-Ki Han

Block boundary matching algorithm for generating side information in distributed video codec

Kwang-Hyun Choi
Jae-Yung Lee
Byeung-Woo Jeon
Jong-Ki Han
Sejong University
Department of Information and Communications
    Engineering
Room 911, Chungmu-Building
98 Kunja-Dong, Kwangjin-Ku, Seoul, Republic of
    Korea
E-mail: hjk@sejong.ac.kr

Abstract. Distributed video codec (DVC) has been developed to construct a simple encoder that utilizes information theory for distributed sources in the circumstance of mobile multimedia communication. In the DVC codec, an efficient algorithm to generate side information (SI) is one of the most important techniques to improve the coding performance. We propose a scheme to increase the quality of SI frame, where the proposed scheme consists of three steps. In the first step, SI frame is constructed by motion estimation and motion compensation in the DVC decoder. Then, in the second step, the blocks in the temporary SI frame are classified into reliable or unreliable ones. The unreliable blocks are updated by block boundary matching algorithm in the third step. Simulation results show that the proposed algorithm outperforms the conventional methods significantly. In addition, the proposed scheme can be combined with the conventional schemes generating SI frame to increase the coding performance of the DVC codec. © *The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.OE.52.10.103111]

1 Introduction

In the last few years, a variety of mobile multimedia devices have been developed and become popular. The end-users would like to capture various scenes and to send those through up-link channels. Usually, the capability of the mobile devices that the end-users have is constrained with low-power and low-CPU speed. In this circumstance, the conventional video codecs, such as MPEG-2, H.264/AVC, and MPEG-4, are not appropriate for the portable devices because those codecs need the high power to encode video sequences. Thus, demand for simple and low-power video encoder has been continuously increasing. Distributed video codec (DVC) is one of the solutions to encode video sequence with low complexity on the encoder side. The complexity of DVC encoder is significantly lower than those of the conventional encoders. DVC codec is regarded as an advanced codec that is appropriate for low power applications such as wireless surveillance, sensor network, and mobile camera phones.

DVC had been developed based on Slepian–Wolf[1] and the Wyner–Ziv[2] theorems which had proved that two signals can be decoded with prediction on the decoder side after those signals have been encoded without prediction between them on the encoder side. A theoretical method for lossless coding with side information (SI) was performed in Ref. 1. This method is referred to as Slepian–Wolf coding and is used with a channel-coding scheme. Slepian–Wolf coding was extended to a lossy compression by Wyner and Ziv.[2,3] The Wyner–Ziv codec consists of a quantization module, channel coding module, such as turbo code and low-density parity-check accumulator, punctured matrix, and SI generator. Most DVC codecs have been developed based on the

models[4,5] which Stanford University and University of California Berkeley have provided. The model of UC Berkeley is named power-efficient robust, high-compression, syndrome-based multimedia (PRISM) coding.[5] In Ref. 6, a project, which is called distributed coding for video services (DISCOVER), had been performed to construct the low-power video codec based on Slepian–Wolf and Wyner–Ziv theorems. In DISCOVER codec, key frames were encoded with the conventional intra-coding technique (intra-mode of H.264/AVC). On the other hand, the Wyner–Ziv frames were split to nonoverlapped blocks and then the blocks were transformed and quantized. The quantized coefficients were ordered bit plane by bit plane. Those coefficients were fed into a systematic channel encoder.

Recently, practical coding techniques using Slepian–Wolf and Wyner–Ziv codecs have been further studied.[7–11] Qing et al.[7] proposed a model for the correlation of noise statistics, which was utilized to increase coding performance. Coding performances of various DVC codecs were analyzed in Ref. 8, where motion vectors (MVs) were estimated with subpixel resolution to generate SI. In Refs. 9–11, enhanced SI was generated using various algorithms. Petrazzuoli et al.[9] proposed a method to generate the SI, where more than two intra-decoded frames are used to estimate the position of the current MC block. A bilateral ME-based scheme to generate SI was described in Ref. 10. In Ref. 11, SI was generated using a side matching algorithm. Among the tools to increase the performance of the DVC codec, an efficient scheme to generate the SI frame is one of the most important techniques because it dominantly affects the quality of the picture in the DVC decoder and the encoding rate. In this paper, we propose a scheme to generate the SI frame using block boundary

matching after analyzing the coding blocks in the initial SI frame.

The main functions of the SI generation module are motion estimation (ME) and motion compensation (MC). Thus, it is important to estimate MVs efficiently in DVC decoder. Because this topic is crucial also in a variety of video codecs,[12–17] for several decades, the related researches have been performed by various researchers to increase the accuracy of MVs,[12–14] and to speed up the ME process.[15–17] In this paper, we propose an efficient scheme to estimate MVs while the specific properties of MVs in DVC decoder are considering.

This paper is organized as follows. In Sec. 2, the model for DVC codec is formulated. The proposed scheme is explained in Sec. 3. Simulation results are presented in Sec. 4. Section 5 concludes this paper.

2 Wyner–Ziv Codec Model

Figure 1 shows a video communication system that incorporates the Wyner–Ziv (WZ) codec. In this system, the odd frames $\{X_{2i-1}, i = 1, 2, 3, \ldots\}$ are encoded using the intra-coding mode of the H.264/AVC standard,[18] while the even frames $\{X_{2i}, i = 1, 2, 3, \ldots\}$ are encoded using the WZ codec which consists of a Slepian–Wolf coding module and an outer quantizer-reconstruction pair. The data generated from encoding intra-frame X_{2i-1} and WZ frame X_{2i} are transmitted separately over independent channels. We assume that the channel for intra-frame is robust enough to prevent channel errors. $\{X'_{2i-1}, i = 1, 2, 3, \ldots\}$ and $\{X'_{2i}, i = 1, 2, 3, \ldots\}$ are the decoded frames on the decoder side.

When a WZ frame is encoded on the encoder side, the frame is split to nonoverlapped 8×8 blocks and the blocks are transformed by discrete cosine transform (DCT). The transformed frame is denoted by T_{2i}. The DCT transformed frame is quantized by a scalar quantizer. The quantized frame is denoted by Q_{2i}, where a quantized datum is represented by a binary index. The binary indexes are encoded using a turbo

coder to protect it from channel error. The turbo encoder changes each binary index to another binary vector which consists of information and redundant bits. Among the resulted bits, only the redundant bits are transmitted to the decoder, whereas the information bits are not sent. The set of binary vectors which turbo encoder generates is denoted by P_{2i}.

On the decoder side, while the odd frames $\{X'_{2i-1}, i = 1, 2, 3, \ldots\}$ are reconstructed by the H.264/AVC decoder with the intra-mode, the even frames $\{X'_{2i}, i = 1, 2, 3, \ldots\}$ are made by WZ decoder. In decoding WZ frames, when the quantized frame Q'_{2i} is generated by the turbo decoder both redundant and information bits are needed. But because the information bits have not been sent from the encoder, the turbo decoder should use the prediction values for the information bits. In Fig. 1, $QT(Y_{2i})$ is used instead of the information bits, where Y_{2i} is SI frame. Y_{2i} is constructed using ME and MC for the previous and next intra-frames, X'_{2i-1} and X'_{2i+1}, that have been reconstructed by H.264/AVC decoder. $QT(Y_{2i})$ is resulted from quantizing after DCT transforming Y_{2i}. After the turbo decoder has generated Q'_{2i}, the decoded WZ frame X'_{2i} is reconstructed by applying inverse DCT after inverse quantizing the Q'_{2i}.

Because the X'_{2i} is made using $QT(Y_{2i})$, the quality of X'_{2i} depends on the quality of Y_{2i}. Therefore, various research[9–11] has been performed to increase the quality of the SI frame. Petrazzuoli et al.[9] proposed a method to generate the SI where more than two intra-decoded frames are used to estimate the position of the current MC block. In the first step of Ref. 9, temporary MVs are estimated from X'_{2i+1} to X'_{2i-1}. Then, the half-sized vectors of the temporary MVs are used as MVs for the corresponding blocks in the SI frame. This step provides a temporary SI frame. In the next step, new forward and backward MVs are estimated from Y_{2i} to X'_{2i+1} and X'_{2i-1}, respectively. Based on the forward and backward MVs, the SI frame is refined. In Ref. 10, a bilateral ME-based scheme to generate SI frame had been
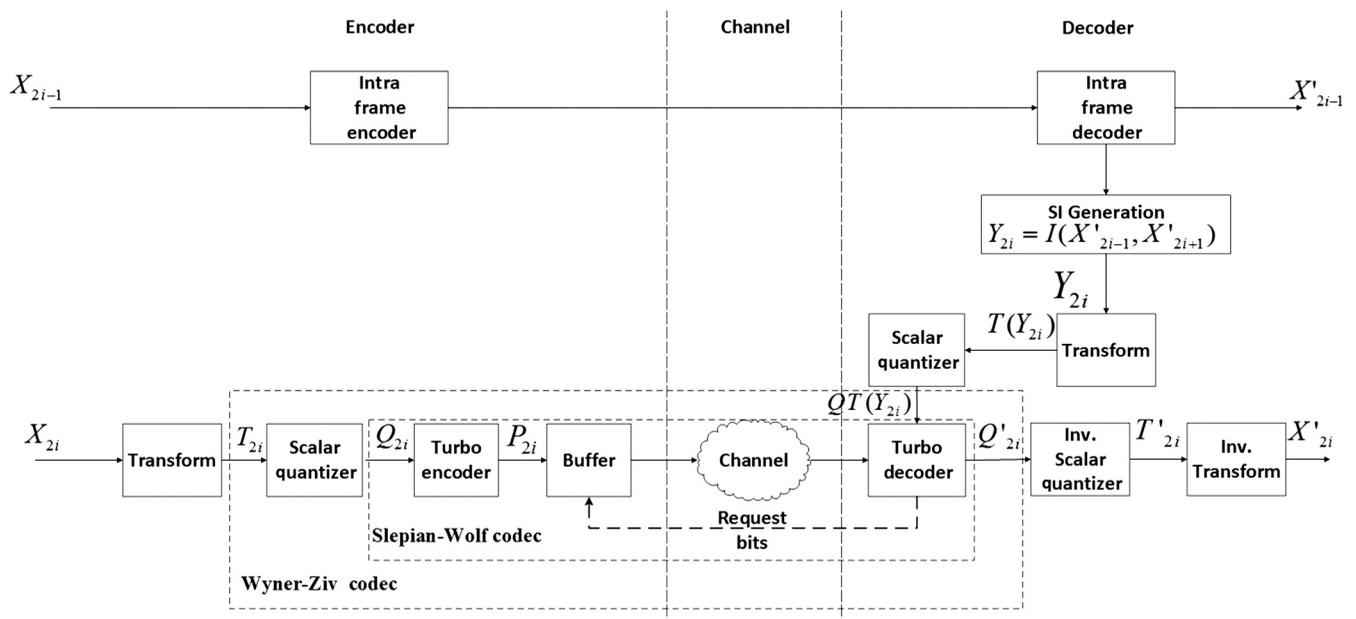
Fig. 1 The video communication system using DVC codec.

proposed, where side matching distortion was used in the ME process. In the initial step of Ref. 10, seed blocks were selected to increase the performance of ME process, which increases the quality of SI frame. Ko et al.[11] had proposed an algorithm to generate SI frame using a side matching algorithm, where blocks in SI frame were refined considering the mismatch error between the template of the current blocks and the corresponding pixels in the intra-frames X'_{2i+1} and X'_{2i-1}.

In this paper, we make an initial SI frame by using ME and MC processes where these procedures are not constrained by a specific algorithm. One of the conventional ME and MC schemes can be used in this step. Because the qualities of some blocks in the initial SI frame may be very low, the blocks are refined according to the reliabilities of the blocks. The reliability of each block is calculated using the concentration ratio of MVs of the neighbor blocks.

3 Proposed Algorithm

The proposed scheme consists of three steps to generate the SI frame. The initial SI frame is constructed by using forward, backward, and bidirectional ME and MC between X'_{2i-1} and X'_{2i+1} in the first step. Then, the blocks in SI frame are analyzed in the second step. Finally, in the third step, the SI frame is refined using the information generated in the second step.

3.1 First Step: Generating the Initial SI

In the first step, a temporary SI frame is constructed by ME and MC for X'_{2i-1} and X'_{2i+1}. After temporary backward and forward MVs are estimated from X'_{2i+1} to X'_{2i-1} and X'_{2i-1} to X'_{2i+1}, respectively, the first temporary SI frame is constructed by MC based on the temporary MVs. Then, new forward and backward MVs are estimated from the first temporary SI frame to X'_{2i+1} and X'_{2i-1}, respectively. By using the MC based on the new forward and backward MVs, the second temporary SI frame is constructed.

Figure 2 shows the relationship between the two key frames (X'_{2i-1}, X'_{2i+1}) and the second temporary SI frame (Y_{2i}). In Y_{2i}, a block $B_{s,t}$ whose size is $N \times N$ is constructed by ME and MC. s and t denote the horizontal and vertical indexes of the block in Y_{2i}. $MV^0_{s,t}$ and $MV^1_{s,t}$ are the MVs estimated for X'_{2i-1} and X'_{2i+1}, respectively, to reconstruct $B_{s,t}$. The superscript 0 and 1 imply the backward and forward data, respectively. The horizontal and vertical components of

$MV^l_{s,t}$ are denoted by $MV^l_{s,t}(x)$ and $MV^l_{s,t}(y)$, respectively, where the superscript l is 0 or 1. As can be seen from other researches,[9–11] it is difficult for the temporary SI frame to demonstrate high quality because the ME and MC modules generate some poor quality blocks.

3.2 Second Step: Evaluating the Reliability

In the second step, the reliability of each $B_{s,t}$ is evaluated. If a $B_{s,t}$ has high quality, then the MVs of the $B_{s,t}$ are highly correlated to those of the neighboring blocks. To describe the neighbor blocks, we define two sets related to the neighboring blocks with $\Omega = \{B_{s+m,t+n}, (m,n) = (-1,-1), (-1,0), (-1,1), (0,-1), (0,1), (1,-1), (1,0), (1,1)\}$ and $\Phi = \Omega \cup \{B_{s,t}\}$. Ω is a set of the neighbor blocks of the current block $B_{s,t}$. Adding the current block $B_{s,t}$ to Ω results in the extended set Φ. Note that the number of blocks in Ω and Φ are 8 and 9, respectively. If the variance of MVs of blocks in Φ is smaller than that of blocks in Ω, it means that adding MVs of $B_{s,t}$ to the set of MVs of blocks in Ω increases the concentration of MVs. This case implies the quality of $B_{s,t}$ is high and the block is reliable.

The variance of MVs can be evaluated by using eigenvalues of the covariance matrix of the MVs. The covariance matrixes of MVs related to Ω and Φ are

$$C^l_\Omega = \begin{bmatrix} R^l_\Omega(xx), & R^l_\Omega(xy) \\ R^l_\Omega(yx), & R^l_\Omega(yy) \end{bmatrix}, \tag{1}$$

$$C^l_\Phi = \begin{bmatrix} R^l_\Phi(xx), & R^l_\Phi(xy) \\ R^l_\Phi(yx), & R^l_\Phi(yy) \end{bmatrix}, \tag{2}$$

where

$$R^l_\Omega(xx) = \frac{\sum_{MV^l_{s,t} \in \Upsilon^l_\Omega}[MV^l_{s,t}(x) - m^l_\Omega(x)]^2}{8}, \tag{3}$$

$$R^l_\Omega(xy) = R^l_\Omega(yx)$$
$$= \frac{\sum_{MV^l_{s,t} \in \Upsilon^l_\Omega}\{[MV^l_{s,t}(x) - m^l_\Omega(x)] \times [MV^l_{s,t}(y) - m^l_\Omega(y)]\}}{8}, \tag{4}$$

$$R^l_\Omega(yy) = \frac{\sum_{MV^l_{s,t} \in \Upsilon^l_\Omega}[MV^l_{s,t}(y) - m^l_\Omega(y)]^2}{8}, \tag{5}$$

$$R^l_\Phi(xx) = \frac{\sum_{MV^l_{s,t} \in \Upsilon^l_\Phi}[MV^l_{s,t}(x) - m^l_\Phi(x)]^2}{9}, \tag{6}$$

$$R^l_\Phi(xy) = R^l_\Phi(yx)$$
$$= \frac{\sum_{MV^l_{s,t} \in \Upsilon^l_\Phi}\{[MV^l_{s,t}(x) - m^l_\Phi(x)] \times [MV^l_{s,t}(y) - m^l_\Phi(y)]\}}{9}, \tag{7}$$

Fig. 2 The initial side information (SI) frame is generated through forward, backward, and bidirectional ME and MC.

$$R_\Phi^l(yy) = \frac{\sum_{MV_{s,t}^l \in \Upsilon_\Phi^l} [MV_{s,t}^l(y) - m_\Phi^l(y)]^2}{9}. \tag{8}$$

In the above equations, Υ_Ω^l and Υ_Φ^l are the sets of $MV_{s,t}^l$'s of blocks in Ω and Φ, respectively. $m_\Omega^l(x), m_\Omega^l(y), m_\Phi^l(x)$, and $m_\Phi^l(y)$ are the mean values of $MV_{s,t}^l(x)$ and $MV_{s,t}^l(y)$ in Υ_Ω^l and Υ_Φ^l, respectively. If the eigenvalues of C_Ω^l and C_Φ^l are denoted by $\lambda_1(\Upsilon_\Omega^l)$, $\lambda_2(\Upsilon_\Omega^l)$, $\lambda_1(\Upsilon_\Phi^l)$, and $\lambda_2(\Upsilon_\Phi^l)$, respectively, the reliability $J_{s,t}$ of the block $B_{s,t}$ is defined as follows:

$$J_{s,t} = \frac{1}{2} \times \frac{|\lambda_1(\Upsilon_\Omega^0)| + |\lambda_2(\Upsilon_\Omega^0)|}{|\lambda_1(\Upsilon_\Phi^0)| + |\lambda_2(\Upsilon_\Phi^0)|} + \frac{1}{2} \times \frac{|\lambda_1(\Upsilon_\Omega^1)| + |\lambda_2(\Upsilon_\Omega^1)|}{|\lambda_1(\Upsilon_\Phi^1)| + |\lambda_2(\Upsilon_\Phi^1)|}. \tag{9}$$

If the block $B_{s,t}$ is made using the ME and MC from only the previous intra-coded frame X'_{2i-1} or only the next frame X'_{2i+1}, then the reliability of Eq. (9) becomes

$$J_{s,t} = \frac{|\lambda_1(\Upsilon_\Omega^0)| + |\lambda_2(\Upsilon_\Omega^0)|}{|\lambda_1(\Upsilon_\Phi^0)| + |\lambda_2(\Upsilon_\Phi^0)|} \tag{10}$$

or

$$J_{s,t} = \frac{|\lambda_1(\Upsilon_\Omega^1)| + |\lambda_2(\Upsilon_\Omega^1)|}{|\lambda_1(\Upsilon_\Phi^1)| + |\lambda_2(\Upsilon_\Phi^1)|}, \tag{11}$$

respectively. In Eqs. (9)–(11), as the value of $|\lambda_1(\Upsilon_\Phi^l)| + |\lambda_2(\Upsilon_\Phi^l)|$ becomes smaller, the MVs in Υ_Φ^l become more locally concentrated. Therefore, the state of $J_{s,t} > 1$ refers to the case in which adding $MV_{s,t}^l$ to Υ_Ω^l results in an increased concentration of Υ_Φ^l. This implies that if the $J_{s,t}$ of a block $B_{s,t}$ is larger than those of the other blocks, the quality of the block $B_{s,t}$ is higher than others.

3.3 Third Step: Refinement of Blocks

In this step, if $J_{s,t}$'s of some blocks are <1, then the blocks are classified as unreliable and remade. On the other hand, if $J_{s,t}$'s of some blocks are not <1, then the blocks are classified as reliable. After the unreliable blocks are sorted in a decreasing order of their $J_{s,t}$'s, those are remade in the order by ME/MC procedure using block boundary matching (BBM). Figure 3 shows that an SI frame consists of reliable and unreliable blocks. Because the neighbor reliable blocks can be used to remake the unreliable block $B_{s,t}$, the MV of $B_{s,t}$ is re-estimated with templates $T_U(s,t)$, $T_B(s,t)$, $T_L(s,t)$, and $T_R(s,t)$ which are the upper, bottom, left, and right templates, respectively. Note that those templates consist of pixels in neighbor reliable blocks. If some neighbor blocks around the current unreliable block $B_{s,t}$ are unreliable, then the corresponding templates are not used. The MV of $B_{s,t}$ is estimated by minimizing the following cost function:

$$\begin{aligned} COST(MV_{s,t}) = \frac{1}{Z} \sum_{l=0,1} &\{w_U \times D[T_U(s,t), R_U^l(MV_{s,t}^l)] \\ &+ w_B \times D[T_B(s,t), R_B^l(MV_{s,t}^l)] \\ &+ w_L \times D[T_L(s,t), R_L^l(MV_{s,t}^l)] \\ &+ w_R \times D[T_R(s,t), R_R^l(MV_{s,t}^l)]\}, \end{aligned} \tag{12}$$
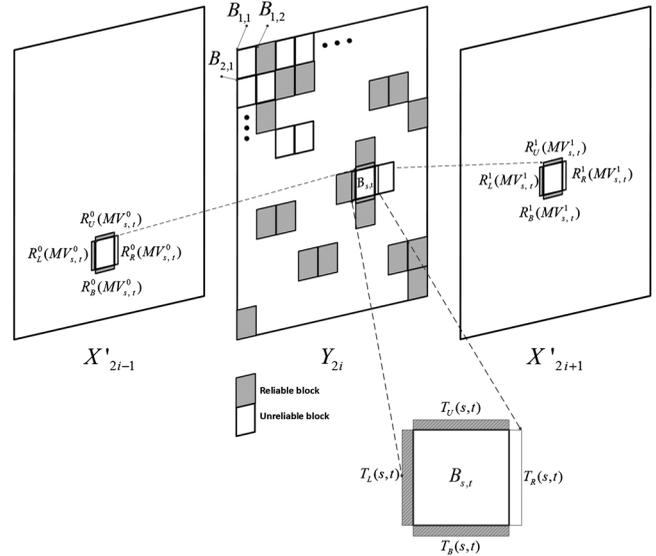
Fig. 3 The SI frame is updated through ME and MC using block boundary matching.

where superscripts "$l = 0$" and "$l = 1$" imply that the data are related to X'_{2i-1} and X'_{2i+1}, respectively. $R_U^l(MV_{s,t}^l)$, $R_B^l(MV_{s,t}^l)$, $R_L^l(MV_{s,t}^l)$, and $R_R^l(MV_{s,t}^l)$ are sets of pixels in reference frames whose pixels are overlapped with templates $T_U(s,t)$, $T_B(s,t)$, $T_L(s,t)$, and $T_R(s,t)$ displaced by $MV_{s,t}$, respectively. $D[T_q(s,t), R_q^l(MV_{s,t}^l)]$ denotes the mean-squared difference between $T_q(s,t)$ and $R_q^l(MV_{s,t}^l)$, where q is one of the $\{U, B, L, R\}$. If $T_q(s,t)$ is used because it is included in the reliable neighbor block, then w_q is set to 1, otherwise, it is set to 0. If ME/MC is performed with forward or backward only, then the corresponding one in $\{l = 0, 1\}$ is used in Eq. (12). The variable Z is the number of pixels used to calculate the cost function of Eq. (12).

4 Simulation Results

In this section, the gain of the proposed scheme is represented by Bjøntegaard delta (BD) rate reduction.[19] The intra-frames are encoded with the intra-mode of JM18.0. The number of frames to be encoded is 300. The GOP structure is "IWIWI. . . ," where "I" and "W" denote the intra- and WZ frames, respectively. When the intra- and WZ frames are encoded, the QPs are set to {27, 31, 35, 39} which are used in the calculation of BD rate reduction.[19] Note that the quantization module of H.264/AVC is used for both intra- and WZ frames. In the decoders of DVC codecs, the MVs were estimated with 1/4 pixel resolution.

In Figs. 4 and 5, and Tables 1 and 2, the DVC codec incorporating the proposed scheme (BBM) is compared with DVC codecs using the conventional algorithms,[9–11] where the size $N \times N$ of $B_{s,t}$ is set to 8×8 or 16×16. Figures 4 and 5 show that the rate-distortion (RD) curves of the DVC codecs using the proposed scheme are higher than those incorporating the conventional methods. It implies that the proposed scheme outperforms the conventional schemes in the viewpoint of RD. In Tables 1 and 2, the proposed algorithm has gains of −5.9%, −5.3%, −8.1% and −6.3%, −5.7%, −8.1% on average BD rates against the conventional schemes. Note that the negative number of BD rate implies that the proposed scheme reduces the total number of
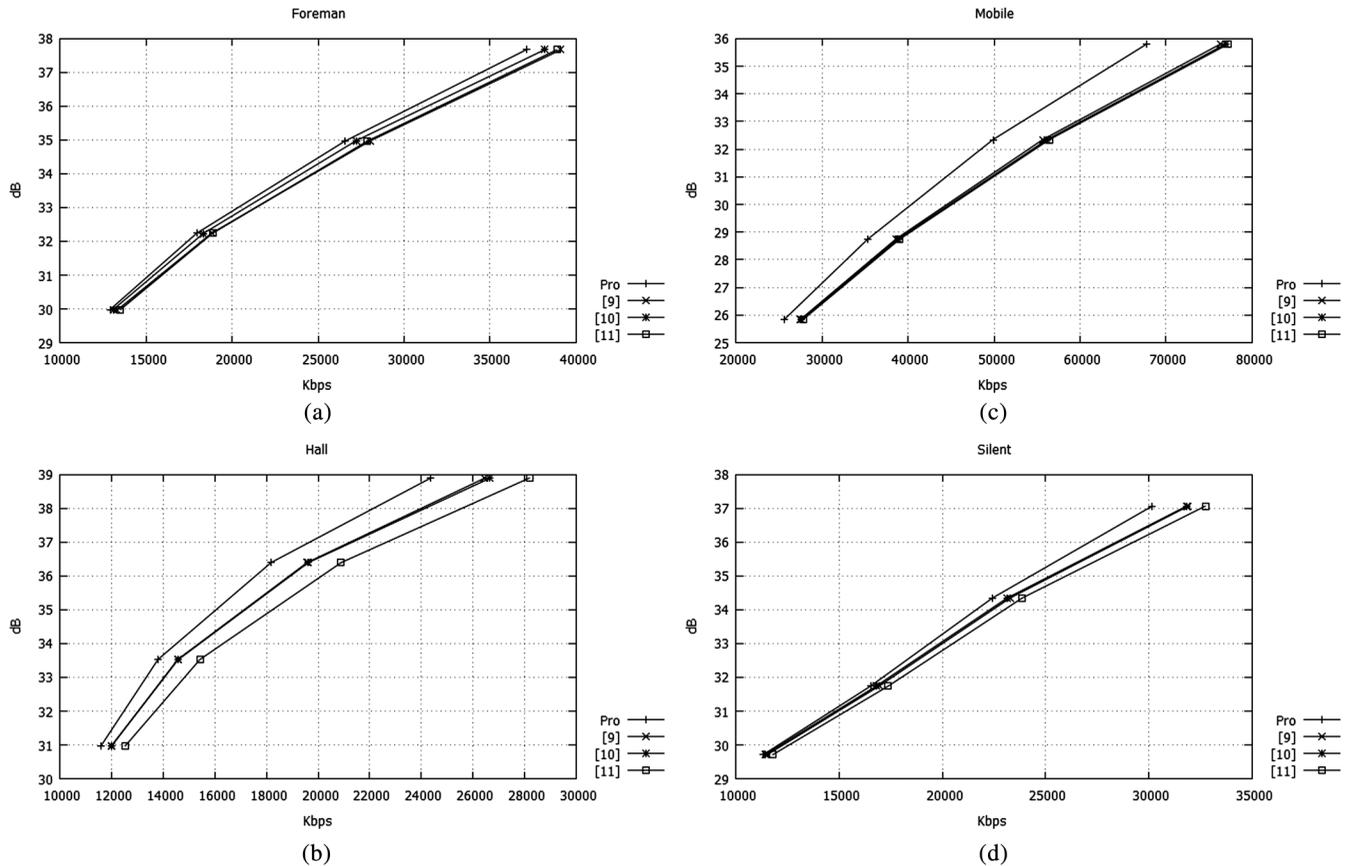
Fig. 4 Comparison between rate-distortion (RD) curves of the conventional and the proposed algorithms, where the BBM is performed with $N \times N = 8 \times 8$. (a) Foreman, (b) mobile, (c) hall, (d) silent.

the bits generated from encoding the video sequence while the image quality resulted from the proposed scheme is equal to those of the conventional schemes. The gains of the case of $N \times N = 16 \times 16$ are larger than that of $N \times N = 8 \times 8$ because the template of $N \times N = 16 \times 16$ block contains the more useful information to construct the SI frame than that of $N \times N = 8 \times 8$ block.

To understand the tendency of gains according to the block size, we show the relationship between the BD rate gain and the block size $N \times N$ in Fig. 6, where the DVC codec incorporating BBM is compared with the DVC codec using the conventional algorithm.[9] Although the performance depends on the test sequences, the overall gains show that the case of $N \times N = 16 \times 16$ provides the best performance among three cases of the block size. The boundary template of $N \times N = 8 \times 8$ is less useful to update the unreliable blocks than that of $N \times N = 16 \times 16$, because the number of pixels in the template of $N \times N = 8 \times 8$ is smaller than that of $N \times N = 16 \times 16$. On the other hand, because in the case of $N \times N = 32 \times 32$, the correlation between pixels in a block and templates of the block is lower than that of the case of $N \times N = 16 \times 16$, the case of $N \times N = 16 \times 16$ gives more gain than that of $N \times N = 32 \times 32$.

In Tables 1 and 2, the complexity of the DVC codec using the proposed scheme is between 97% and 105% of those using the conventional schemes. Note that 100% implies the complexity of the proposed scheme is equal to that of the conventional scheme. The complexities represented in the tables show that the complexity of the proposed

scheme is approximately equal to those of the conventional schemes.

Validity of the reconstructed SI has been checked in Table 3 where the averaged PSNRs of SI reconstructed in DVC decoder were measured. Even though it is not allowed in DVC scenario, it is helpful to check the accuracy of SI generation algorithms. As we can see from this table, the PSNRs of SI generated by the proposed BBM are higher than those resulting from the conventional algorithms. It implies that the proposed scheme outperforms the conventional algorithms in increasing the quality of the generated SI frame.

To analyze the performances of the steps of the proposed algorithm, the percentage of the blocks regarded as unreliable in the proposed scheme is shown in Table 4, where the percentages of the unreliable blocks are <3%. Although the portion is small, the gains resulted by updating the unreliable blocks are significant. In this table, we measure BD rate gains of the third step against the first step of the proposed scheme. This table shows that the third step of the proposed algorithm increase the coding efficiency significantly.

The second and third steps of the proposed scheme described in Sec. 3 can be used to enhance the performance of the conventional algorithms,[9–11] because the steps increase the quality of the SI frame by updating those that have been made by the conventional algorithms. To demonstrate it, the performance comparisons between DVC codecs using the combined techniques (conventional schemes + second and
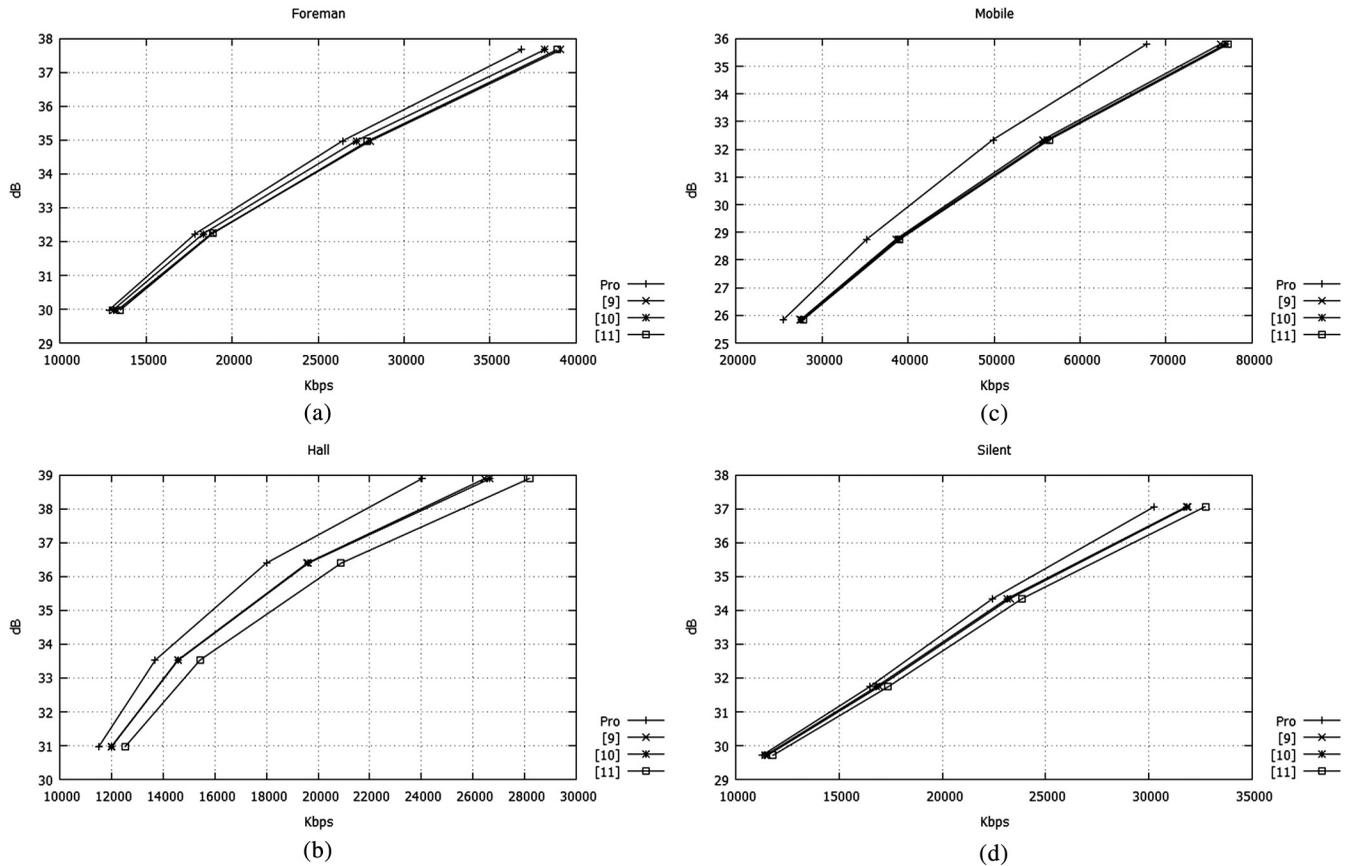
Fig. 5 Comparison between RD curves of the conventional and the proposed algorithms, where the BBM is performed with $N \times N = 16 \times 16$. (a) Foreman, (b) mobile, (c) hall, (d) silent.

Table 1 Bjøntegaard delta (BD) rate gains and the relative complexity of the proposed DVC codec. $N \times N = 8 \times 8$.

Sequences		Ref. 9 versus BBM	Ref. 10 versus BBM	Ref. 11 versus BBM
Foreman (CIF)	BD rate	−4.8%	−2.2%	−4.7%
	Complexity	101%	91%	107%
Mobile (CIF)	BD rate	−9.4%	−10.1%	−10.6%
	Complexity	96%	96%	99%
Hall (CIF)	BD rate	−6.1%	−6.3	−11.5%
	Complexity	101%	101%	103%
Silent (CIF)	BD rate	−3.3	−2.7	−5.7%
	Complexity	101%	102%	106%
Average	BD rate	−5.9%	−5.3%	−8.1%
	Complexity	100%	97%	105%

Table 2 BD rate gains and the relative complexity of the proposed DVC codec. $N \times N = 16 \times 16$.

Sequences		Ref. 9 versus BBM	Ref. 10 versus BBM	Ref. 11 versus BBM
Foreman (CIF)	BD rate	−5.3%	−2.7%	−5.2%
	Complexity	99%	100%	104%
Mobile (CIF)	BD rate	−7.0%	−7.2%	−12.3%
	Complexity	109%	91%	90%
Hall (CIF)	BD rate	−9.6%	−10.3%	−10.8%
	Complexity	86%	81%	83%
Silent (CIF)	BD rate	−3.4%	−2.8%	−5.8%
	Complexity	104%	118%	112%
Average	BD rate	−6.3%	−5.7%	−8.1%
	Complexity	104%	98%	97%
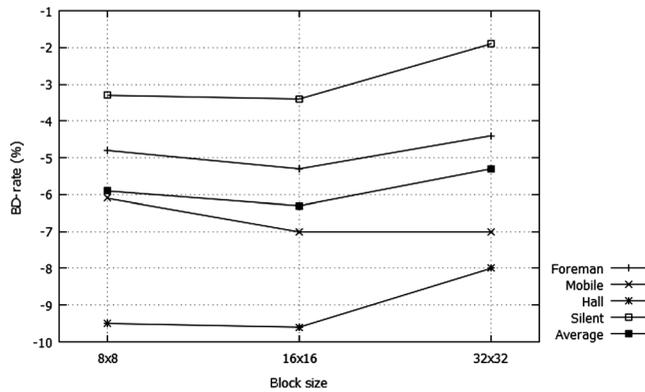
Fig. 6 Bjøntegaard delta rate gains of the DVC codec incorporating BBM according to the block sizes against the DVC codec using the conventional algorithm (Ref. 9).

Table 3 Averaged PSNRs of the SI's reconstructed in the DVC decoder. $N \times N = 16 \times 16$.

Sequences	Ref. 9	Ref. 10	Ref. 11	BBM
Foreman (CIF)	27.59	27.90	27.16	28.48
Mobile (CIF)	22.69	22.65	22.48	23.86
Hall (CIF)	32.10	31.94	30.75	33.74
Silent (CIF)	29.90	29.86	28.72	30.23
Average	28.07	28.09	27.28	29.08

third steps of BBM) and the conventional schemes are represented in Table 5, where $N \times N = 16 \times 16$. In the DVC codec using the combined techniques, instead of the first step of the proposed algorithm, one of the conventional schemes is used to make the temporary SI frame. The temporary SI frame is updated by the second and third steps of the proposed algorithm. The DVC codecs using the combined techniques are more efficient than codecs using the conventional schemes. As for the results related to Ref. 9

Table 4 Percentage of the unreliable blocks in the proposed scheme. BD rate gains of the third step against the first step in the proposed scheme. $N \times N = 16 \times 16$.

Sequences	Percentage of the unreliable blocks	BD rate gains of the third step against the first step in the proposed schemes
Foreman (CIF)	2.33%	−0.5%
Mobile (CIF)	1.91%	−1.1%
Hall (CIF)	1.25%	−0.9%
Silent (CIF)	1.45%	−0.1%
Average	1.55	−0.65

Table 5 BD rate gains and the relative complexity of the DVC codec using the combined techniques. $N \times N = 16 \times 16$.

Sequences		Ref. 9 versus Ref. 9 + BBM	Ref. 10 versus Ref. 10 + BBM	Ref. 11 versus Ref. 11 + BBM
Foreman (CIF)	BD rate	−0.2%	−3%	−1.5%
	Complexity	110%	106%	109%
Mobile (CIF)	BD rate	−0.1%	−4.2%	−2.9%
	Complexity	82%	104%	86%
Hall (CIF)	BD rate	−0.2%	−3.7%	−1.1%
	Complexity	100%	97%	96%
Silent (CIF)	BD rate	−0.1%	−1.8%	−2.2%
	Complexity	87%	101%	100%
Average	BD rate	−0.2%	−4.4%	−2.0%
	Complexity	95%	102%	98%

in Table 5, the gains of the combined method are insignificant. Because SI frame constructed by Petrazzuoli et al.[9] includes a lot of unreliable blocks, the number of reliable neighbor blocks that the BBM scheme can utilize is small. Note that BBM is useful when a lot of reliable neighbor blocks are included in the temporary SI frame.

In DVC codec, the complexity of the channel decoder (turbo decoder) decreases as the quality of the SI frame increases, because the number of operations sending more parity bits which are requested by the decoder is reduced as the quality of the SI frame increases. Therefore, in Table 5, because the quality of SI frame generated from the combined techniques is higher than those resulting from the conventional schemes, the complexity of the DVC codec using the combined techniques may be smaller than the conventional codecs. This table shows that the proposed scheme (BBM) can be used to increase the performances of the conventional schemes.

5 Conclusions

This paper proposes an efficient method to reconstruct the SI frame in a DVC decoder. In the proposed algorithm, the blocks in the SI frame are classified to reliable and unreliable blocks, and then the unreliable blocks are remade using a BBM scheme. Simulation results show that the proposed scheme outperforms the conventional methods. The proposed scheme can be combined with the conventional schemes to increase the coding performance further.

References

1. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory* **19**(4), 471–480 (1973).
2. A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory* **22**(1), 1–10 (1976).
3. A. D. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Inf. Theory* **21**(3), 2941–300 (1975).
4. B. Girod et al., "Distributed video coding," *Proc. IEEE* **93**(1), 71–83 (2005).
5. R. Puri and K. Ramchandran, "PRISM: a "reversed" multimedia coding paradigm," in *Proc. IEEE Int. Conf. on Image Processing*, Barcelona, Spain, Vol. 1, pp. I-617–I-620 (2003).
6. X. Artigas et al., "The DISCOVER codec: architecture, techniques and evaluation," in *Picture Coding Symposium*, Lisbon, Spain (2007).
7. L.-B. Qing, X.-H. He, and R. Lv, "Modeling non-stationary correlation noise statistics for Wyner–Ziv video coding," in *International Conference on Wavelet Analysis and Pattern Recognition '07*, Vol. **1**, pp. 316–320 (2007).
8. Z. Li, L. Liu, and E. J. Delp, "Rate-distortion analysis of motion side estimation in Wyner–Ziv video coding," *IEEE Trans. Image Process.* **16**(1), 98–113 (2007).
9. G. Petrazzuoli, M. Cagnazzo, and B. P. Popescu, "High order motion interpolation for side information improvement in DVC," in *Proc. IEEE Int. Conf. on Acoustics Speech and Signal Processing 2010*, Dallas, Texas, pp. 2342–2345 (2010).
10. D. Y. Kim, D. S. Jun, and H.W. Park, "An efficient side information generation using seed blocks for distributed video coding," in *Picture Coding Symposium*, Nagoya, Japan, pp. 86–89 (2010).
11. B. Ko, H. Shim, and B. Jeon, "Wyner–Ziv video coding with side matching for improved side information," in *Pacific-Rim Symposium on Image and Video Technology 2007*, Santiago, Chile, pp. 816–825 (2007).
12. S.-C. Tai et al., "A multi-pass true motion estimation scheme with motion vector propagation for frame rate up-conversion applications," *J. Display Technol.* **4**(2), 188–197 (2008).
13. S. Nikitidis, S. Zafeiriou, and I. Pitas, "Camera motion estimation using a novel online vector field model in particle filters," *IEEE Trans. Circuits Syst. Video Technol.* **18**(8), 1028–1039 (2008).
14. K. R. Namuduri, "Motion estimation using spatio-temporal contextual information," *IEEE Trans. Circuits Syst. Video Technol.* **14**(8), 1111–1115 (2004).
15. P. C. Shenolikar and S. P. Narote, "Different approaches for motion estimation," in *Int. Conf. on Control, Automation, Communication and Energy Conservation*, Perundurai, India, pp. 1–4 (2009).
16. D. Zhang, G. Cao, and X. Gu, "Improved motion estimation based on motion region identification," in *Int. Conf. on Systems and Informatics (ICSAI)*, Yantai, China , pp. 2034–2037 (2012).
17. D. Vu, Y. Yang, and L. Bhuyan, "An efficient dynamic multiple-candidate motion vector approach for GPU-based hierarchical motion estimation," in *Int. Performance Computing and Communications Conference (IPCCC)*, Austin, Texas, pp. 342–351 (2012).
18. T. Wiegand, G. J. Sulivan, and A. Luthra, "Draft ITU-T Recommendation H.264 and Final Draft International Standard 14496-10 AVC," *JVT of ISO/IEC JTC1/SC29/WG11 and ITU-TSG16/Q.6, Doc. JVT-G050r1*, Geneva, Switzerland (2003).
19. G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," in *ITU-T SG16 Q.6, VCEG-M33*, Austin, Texas, pp. 1–5 (2001).

Jae-Yung Lee received the BS and MS degrees from the Department of Information and Communication Engineering, Sejong University, Seoul, Republic of Korea, in 2011 and 2013, respectively, and is currently pursuing the PhD degree from Sejong University. His research interests include video coding, signal processing, high level syntax incorporated in high efficiency video coding (HEVC), scalable extension of HEVC (SHVC), and 3-D-HEVC.

Byeung-Woo Jeon received the BS degree (magna cum laude) in 1985, the MS degree in 1987 in electronics engineering from Seoul National University, Seoul, Republic of Korea, and the PhD degree in electrical engineering from Purdue University, West Lafayette, Indiana, in 1992. From 1993 to 1997, he was in the Signal Processing Laboratory, Samsung Electronics, Republic of Korea, where he conducted research and development into video compression algorithms, the design of digital broadcasting satellite receivers, and other MPEG-related research for multimedia applications. Since September 1997, he has been with the Faculty of the School of Information and Communication Engineering, Sungkyunkwan University, Korea, where he is currently a professor.

Jong-Ki Han received the BS, MS, and PhD degrees in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST), Taejon, Korea, in 1992, 1994, and 1999, respectively. From 1999 to 2001, he was a member of technical staff with the Corporate R&D Center, Samsung Electronics Company, Suwon, South Korea. He is currently a professor with the Department of Information and Communications Engineering, Sejong University, Seoul, Republic of Korea. His research interests include image and video coding, transcoding, and signal processing for video sequence. He has participated in standardization of high efficiency video coding (HEVC) since 2010.

Kwang-Hyun Choi received the BS degree from the Department of Information and Communication Engineering, Sejong University, Seoul, Republic of Korea, in 2012 and currently pursuing the MS degree from Sejong University. His research interests include video coding, signal processing, high efficiency video coding and three-dimensional high efficiency video coding.