Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Hash signature saving in distributed video coding*

Xin-hao CHEN[†1,2], Xing-guo ZHU[1,2], Xiao-lin SHEN[1,2], Lu YU[†‡1,2]

([1]*Institute of Information and Communication Engineering, Zhejiang University, Hangzhou 310027, China*)

([2]*Zhejiang Provincial Key Laboratory of Information Network Technology, Hangzhou 310027, China*)

[†]E-mail: chen2xin1hao4@163.com; yul@zju.edu.cn

Received Jan. 11, 2010; Revision accepted May 31, 2010; Crosschecked Dec. 12, 2010

Abstract: In transform-domain distributed video coding (DVC), the correlation noises (denoted as $\boldsymbol{N}$) between the source block and its temporal predictor can be modeled as Laplacian random variables. In this paper we propose that the noises (denoted as $\boldsymbol{N}'$) between the source block and its co-located block in a reference frame can also be modeled as Laplacian random variables. Furthermore, it is possible to exploit the relationship between $\boldsymbol{N}$ and $\boldsymbol{N}'$ to improve the performance of the DVC system. A practical scheme based on theoretical insights, the hash signature saving scheme, is proposed. Experimental results show that the proposed scheme saves on average 83.2% of hash signatures, 13.3% of bit-rate, and 3.9% of encoding time.

Key words: Distributed video coding (DVC), Hash, Laplacian
doi:10.1631/jzus.C1000008 **Document code:** A **CLC number:** TN919.8

1 Introduction

Distributed video coding (DVC) is an alternative video coding paradigm developed in recent years (Girod *et al.*, 2005; Guillemot *et al.*, 2007; Pereira *et al.*, 2008; Dufaux *et al.*, 2009). Compared with the conventional hybrid video coding adopted by many well known international standards like MPEG-2 Video (ISO/IEC 13818-2:1994), H.263 (Cote *et al.*, 1998), and H.264/AVC (ISO/IEC 14496-10:2003, and also ITU-T Recommendation H.264, 2003; Ostermann *et al.*, 2004), DVC does not rely on motion-compensated prediction techniques that require a temporal predictor at the encoder and that lead to a complicated encoder with a much simpler decoder (Wedi and Musmann, 2003). DVC allows a more flexible allocation of the complexity between the encoder and the decoder with an acceptable coding efficiency, and has an improved robustness against channel noises. These functional features make DVC well suited to many emerging video applications such as resource-constrained wireless video camera, video sensor networks, and low-power surveillance, in which hybrid video coding is problematic.

The foundations of DVC are the Slepian-Wolf theorem for lossless distributed coding (Slepian and Wolf, 1973) and the Wyner-Ziv theorem for lossy compression with side information (Wyner, 1974; Wyner and Ziv, 1976; Yang *et al.*, 2007). These theorems suggest that it is possible for DVC to approach the coding efficiency of hybrid video coding, without the need to access the temporal predictor, namely, the side information available at the decoder.

There is a rate-allocation issue in practical DVC systems: it is difficult for the encoder to estimate the minimum bit-rate of a specific coding block adaptively without access to the side information available at the decoder. This is because the minimum bit-rate is dependent on the correlation noises between the source information and the side information but the statistics of noises are highly variable because of time-varying video content. According to their different solutions to the rate-allocation issue, practical systems can be classified as either feedback-based or feedback-free systems.

---

‡ Corresponding author
* Project supported by the National Basic Research Program (973) of China (No. 2009CB320903) and the Program for New Century Excellent Talents in University

In feedback-based systems (Aaron *et al.*, 2002; 2004a; 2004b; Varodayan *et al.*, 2008), the source information is turbo or low density parity-check (LDPC) codes encoded, stored in buffer, and transmitted in packet according to the requests from the decoder side via the feedback channel. The decoder then attempts to reconstruct the source information using the side information and the bits already received. If the decoding of source information fails, the decoder will send requests for more bits via the feedback channel. Other feedback-based systems have been proposed by Artigas *et al.* (2007), Ascenso and Pereira (2007), Guo *et al.* (2007), Hua and Chen (2008), Kuganeswaran *et al.* (2008), and Taewon *et al.* (2009).

The feedback channel simplifies the rate-allocation issue of DVC and improves the coding efficiency: a loss up to 1.2 dB will occur if the feedback channel is removed (Brites and Pereira, 2007). However, feedback-based systems have several drawbacks: (1) the requirement of feedback channel; (2) the delay introduced by using feedback channel between the decoder and the encoder; (3) the requirement for serious synchronization and interaction between the decoder and the encoder.

Another approach to performing rate-allocation is to estimate approximately the correlation noises at the encoder, as part of a feedback-free system (Puri and Ramchandran, 2003; Puri *et al.*, 2007). The encoder calculates the mean squared error (MSE) between a source block and a co-located block that has the same 2D coordinates as the source block in a reference frame, to estimate the correlation noises between the source block and its temporal predictor in the reference frame. Other feedback-free systems have been proposed (Fowler, 2005; Asif and Soraghan, 2008).

Because motion estimation is no longer performed at the encoder and there are no motion vectors in the bitstream to indicate the coordinates of the temporal predictor in the reference frame, a hash signature of the source information $\boldsymbol{X}$ is often generated at the encoder. The signature is transmitted in the bitstream to help the decoder find the correct side information $\boldsymbol{Y}$ among candidates in the region of the motion search (Puri and Ramchandran, 2003; Aaron *et al.*, 2004b; Ascenso and Pereira, 2007; Taewon *et al.*, 2009).

The correlation noises (denoted as $\boldsymbol{N}$) between the source block and the temporal predictor are often modeled as Laplacian sources. In this paper, we propose that the noises (denoted as $\boldsymbol{N}'$) between the source block and the co-located block in the reference frame can also be modeled as Laplacian sources. Furthermore, the statistical relationship between $\boldsymbol{N}$ and $\boldsymbol{N}'$ can be exploited to improve the system performance. A practical scheme guided by theoretical insights, the hash signature saving scheme, is proposed for improving the coding efficiency and saving encoding time.

2 Description of the original feedback-free and hash-based system

We adopted the PRISM framework, which is characterized by a feedback-free architecture with a hash-based motion search at the decoder (Puri and Ramchandran, 2003; Puri *et al.*, 2007). The diagram of our implementation is illustrated in Fig. 1.

At the encoder, the source information is 64 quantized discrete cosine transform (DCT) coefficients of each 8×8 block in the source frame, and is denoted as a vector of random variables $\boldsymbol{X}=(X_0, X_1, \ldots, X_{63})$,

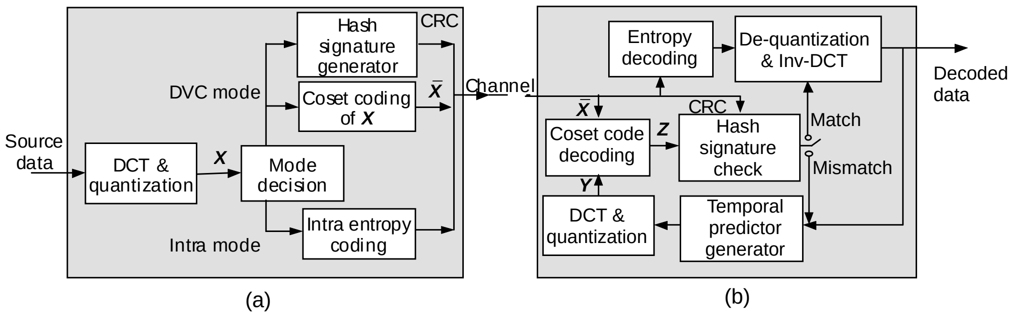

Fig. 1 Diagram of the original codec of feedback-free and hash-based distributed video coding (DVC) system
(a) Encoder; (b) Decoder

where X_0 is the DC coefficient. Given $\boldsymbol{X}$, the mode decision module selects the coding mode for each block from three modes, skip, DVC, and intra, according to the MSE between the source block and its co-located block in the reference frame. Two thresholds, T_1 and T_2, are predetermined by statistical experiment. If the MSE>T_1, a rapid motion is detected and thus the intra mode will be adopted, and $\boldsymbol{X}$ is coded losslessly in the intra entropy coding module; if the MSE<T_2, the skip mode is adopted; otherwise, the DVC mode is adopted. The intra mode and the quantization adopt the H.263 tools (Cote *et al.*, 1998). The quantization parameter (QP) is transmitted via channel for each block, except the blocks coded in the skip mode.

For each block that is coded by DVC mode, the coset indices of X_i are computed as

$$\overline{X_i} = X_i - \left\lfloor \frac{X_i}{2C_i} \right\rfloor \cdot 2C_i, \tag{1}$$

where $0 \le i < 64$ and the operator $\lfloor \cdot \rfloor$ denotes the floor of a variable. $\boldsymbol{C}=(C_1, C_2, \ldots, C_{63})$ is a constant parameter vector calculated from the probability of correct decoding (denoted as α). The encoder and the decoder will use the same $\boldsymbol{C}$ stored in memory. The formulated relationship between $\boldsymbol{C}$ and α will be further analyzed in this work.

$\boldsymbol{X}$ is arranged as a binary sequence in the hash signature generator module, and then a 16-bit cyclic redundancy check (CRC) of this binary sequence is computed as the hash signature of $\boldsymbol{X}$, and transmitted in the bitstream. As a result, for each block coded in the DVC mode, the bitstream is composed of the coset indices and the hash code of $\boldsymbol{X}$.

At the decoder, the coding mode of each block is first decoded from the bitstream. For the blocks coded in the skip mode, the reconstructed information is copied from the co-located block in the reference frame; for the blocks coded in the intra mode, $\boldsymbol{X}$ is losslessly decoded in the entropy decoding module.

For the blocks coded in the DVC mode, the hash signature and the coset indices of $\boldsymbol{X}$ are decoded. A specified coset index $\overline{X_i}$ indicates that X_i is one of the elements in the coset $\Lambda^{\overline{X_i}} = \{\overline{X_i} + j \cdot 2C_i \mid j \in \mathbb{Z}\}$. In this coset, each element shares the same coset index $\overline{X_i}$.

To select one element of this coset as the reconstructed information Z_i, a temporal predictor of $\boldsymbol{X}$ in the transform-domain, namely the side information $\boldsymbol{Y}$, is yielded. In this study, the side information is 64 quantized DCT coefficients of an 8×8 temporal predictor of $\boldsymbol{X}$ yielded from reference frames, and is denoted as a vector of random variables $\boldsymbol{Y}=(Y_0, Y_1, \ldots, Y_{63})$, where Y_0 is the DC coefficient. The reference frames are decoded previously and stored in memory. Given a candidate of $\boldsymbol{Y}$, the reconstructed information $\boldsymbol{Z}=(Z_0, Z_1, \ldots, Z_{63})$ is computed as follows:

$$Z_i = \arg\min_{z \in \Lambda^{\overline{X_i}}} (|z - Y_i|), \ \ 0 \le i < 64. \tag{2}$$

Since the distance between two adjacent elements in the coset $\Lambda^{\overline{X_i}}$ is $2C_i$, $Z_i=X_i$ if and only if $|X_i-Y_i|<C_i$. We denote the correlation noises between $\boldsymbol{X}$ and $\boldsymbol{Y}$ as a vector of random variables $\boldsymbol{N}=(N_0, N_1, \ldots, N_{63})$, where $N_i=X_i-Y_i$, $0 \le i < 64$.

The hash signature of $\boldsymbol{Z}$ is then computed and compared with the hash signature of $\boldsymbol{X}$ decoded from the bitstream. If the two signatures match, $\boldsymbol{Z}$ is treated as the correct reconstructed information and the motion search stops; otherwise, the decoder will try the next motion vector in the region of the motion search to yield a new candidate of $\boldsymbol{Y}$, and then the coset code decoding and the hash check procedure are repeated. The initial motion vector is zero; i.e., the first $\boldsymbol{Y}$ is yielded from the co-located block of $\boldsymbol{X}$ in the reference frame. In the case that none of $\boldsymbol{Z}$ in the region of the motion search match with $\boldsymbol{X}$, the decoder applies the error resilience technique. If the length of the hash signature is sufficient, the probability that the decoder will find at least one correct $\boldsymbol{Z}$ is equal to α.

3 Modeling and statistical analysis

In this section, we address a special case in which the co-located information can be used as side information to yield the correct reconstructed information at the decoder. We denote the probability of the special case as $\alpha_{\text{co-located}}$. The formulated relationship between α and $\alpha_{\text{co-located}}$ will be presented.

We denote the quantized DCT coefficients of the co-located block in the reference frame as a vector of random variables $\boldsymbol{Y}'=(Y_0', Y_1', \ldots, Y_{63}')$, where Y_0' is the DC coefficient. We denote the noises between $\boldsymbol{X}$

and $\boldsymbol{Y}'$ as a vector of random variables $\boldsymbol{N}'=(N_0', N_1', \ldots, N_{63}')$, where $N_i'=X_i-Y_i'$, $0\leq i<64$. The reconstructed information $\boldsymbol{Z}'$ that uses $\boldsymbol{Y}'$ as the side information in Eq. (2) can then be computed as follows:

$$Z'_i = \arg\min_{z\in\Lambda^{X_i}} (|z-Y_i'|),\ \ 0\leq i<64. \tag{3}$$

Since $Z_i=X_i$ if and only if $|X_i-Y_i|<C_i$ in Eq. (2), we can obtain

$$\alpha_i = P(Z_i = X_i) = P(|N_i| < C_i), \tag{4}$$

$$\alpha'_i = P(Z_i' = X_i) = P(|N_i'| < C_i), \tag{5}$$

where α_i represents the probability of correct decoding of each coefficient, given the optimal side information $\boldsymbol{Y}$. Similarly, α_i' represents the probability of correct decoding of each coefficient, given the co-located information $\boldsymbol{Y}'$ in Eq. (3). The experiment to obtain the statistics of N_i and N_i' is carried out. During the off-line statistical experiments, $\boldsymbol{Y}$ is yielded by the full-search motion estimation and sub-pixel motion-compensated prediction. According to the result of the experiment, $\boldsymbol{N}$ can be well modeled as a vector of independent Laplacian random variables, as well as $\boldsymbol{N}'$. An example of N_0 is shown in Fig. 2 and also presented by Aaron *et al.* (2002), Puri *et al.* (2007), Kuganeswaran *et al.* (2008), and Mukherjee (2009). The probability density function of N_i is given by

$$f_{N_i}(n) = \frac{b_i}{2}\exp(-b_i|n|), \tag{6}$$

where b_i is a scale parameter computed from the variance of N_i:

$$\sigma_{N_i}^2 = 2/b_i^2. \tag{7}$$

By substituting Eq. (6) into Eq. (4), the relationship between α_i and C_i can be formulated as Eq. (8) and illustrated as in Fig. 3.

$$\alpha = \int_{-C_i}^{C_i} \frac{b_i}{2}\exp(-b_i|x|)\mathrm{d}x = 1-\exp(-b_iC_i). \tag{8}$$

Similarly, we can obtain

$$\alpha'_i = \int_{-C_i}^{C_i} \frac{b_i'}{2}\exp(-b_i'|x|)\,\mathrm{d}x = 1-\exp(-b_i'C_i), \tag{9}$$

where b' is the scale parameter of N_i', and can be computed as in Eq. (7). Assuming Z_i is statistically independent, we can obtain

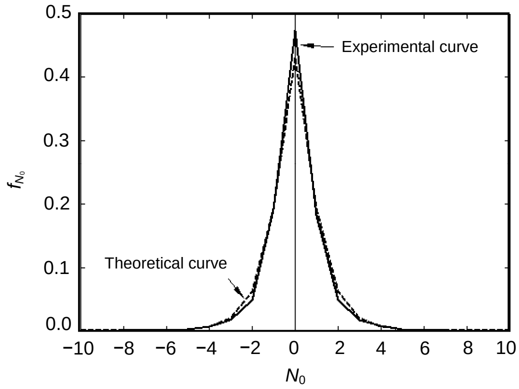

Fig. 2 Experimental and theoretical distribution of N_0, container_cif

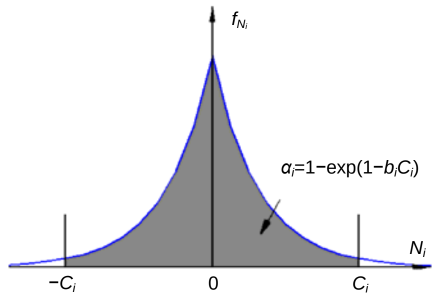

Fig. 3 The relationship between α_i and C_i

$$\alpha = \prod_i \alpha_i = \prod_i [1-\exp(-b_iC_i)]. \tag{10}$$

Similarly, we can obtain

$$\alpha_{\text{co-located}} = \prod_i \alpha_i' = \prod_i [1-\exp(-b_i'C_i)]. \tag{11}$$

Furthermore, by letting the probabilities of correct decoding for each coefficient be equal to each other, we can write

$$\alpha_i = \sqrt[64]{\alpha}\ . \tag{12}$$

From Eqs. (10) and (12), the parameter vector $\boldsymbol{C}$ can be computed if α is given:

$$C_i = -\frac{1}{b_i}\ln(1-\sqrt[64]{\alpha}). \tag{13}$$

Substituting Eq. (13) into Eq. (9) and Eq. (11) respectively, we can obtain

$$\alpha_i' = 1-(1-\sqrt[64]{\alpha})^{b_i'/b_i}, \tag{14}$$

$$\alpha_{\text{co-located}} = \prod_i [1-(1-\sqrt[64]{\alpha})^{b_i'/b_i}], \tag{15}$$

where b_i' and b can be estimated from the statistical experiment. $\alpha_{\text{co-located}}$ is always smaller than α, because $\boldsymbol{Y}$ is the temporal predictor for the source block with minimum MSE and b_i' is always smaller than b_i. Eq. (14) is illustrated in Fig. 4. Table 1 shows the theoretical results of $\alpha_{\text{co-located}}$ under four QPs, given the optimal α in the codec of Section 2. Furthermore, letting $\beta=\alpha_{\text{co-located}}/\alpha$, Fig. 5 shows the theoretical curves of β versus α for hall_monitor, CIF, and four different QPs. Note that $\beta\geq0.8$ for this video sequence, given that $0.8\leq\alpha\leq1$.

Because the decoder can locate the co-located block without motion search, $\alpha_{\text{co-located}}$ also represents the probability that each block can be decoded correctly without hash signature if the parameter vector $\boldsymbol{C}$ is unchanged; i.e., the hash signature is unnecessary for 90.5% of blocks on average, according to the results in Table 1. This result can be used to improve the system performance.

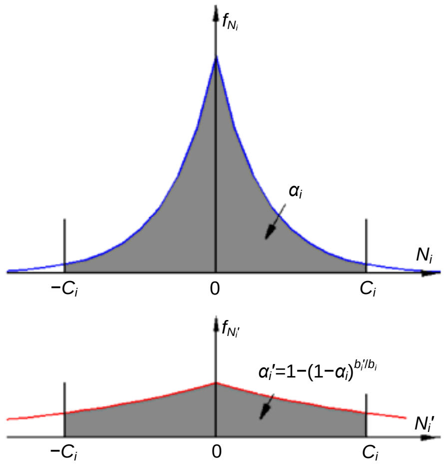

Fig. 4 The relationship between α_i' and α_i

Table 1 Theoretical values of $\alpha_{\text{co-located}}$ under four QPs

Sequence	$\alpha_{\text{co-located}}$ (%)			
	QP=14	QP=20	QP=25	QP=30
container_cif	98.4	98.4	97.8	98.1
hall_monitor_cif	95.0	96.5	97.4	97.2
hall_qcif	97.9	98.4	98.4	97.6
mother_daughter_cif	92.7	94.5	95.3	96.0
foreman_cif	48.1	58.3	62.9	65.4
news_cif	82.6	88.2	90.6	93.0
paris_cif	84.4	91.1	93.7	95.6
bridge-close_cif	98.3	98.4	97.7	96.6
Average	87.2	90.5	91.7	92.4

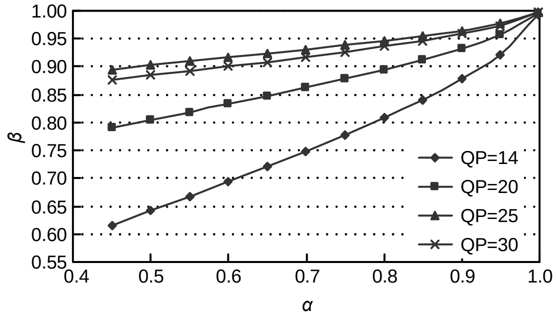

Fig. 5 Theoretical values of β over α for hall_monitor, CIF, and four different QPs

4 Proposed hash signature saving scheme

This section presents a practical scheme, named hash signature saving. A diagram of the proposed codec is shown in Fig. 6. Define a signal vector $\boldsymbol{S}=(S_0, S_1, \dots, S_{63})$:

$$S_i=\begin{cases}0, & |N_i'|<C_i,\\ 1, & \text{otherwise,}\end{cases}\quad 0\leq i<64. \tag{16}$$

If $\boldsymbol{S}$ is a zero vector, then the decoder will yield a correctly decoded block with $\boldsymbol{Y}'$, and the hash signature of $\boldsymbol{X}$ is unnecessary for this block; otherwise, the hash signature of $\boldsymbol{X}$ is still transmitted to the decoder side. As a result, the encoder has to indicate the presence of the hash signature in the bitstream by a 1-bit syntax element hash_flag, which is given by

$$\text{hash\_flag}=\begin{cases}0, & \boldsymbol{S}=\boldsymbol{0},\\ 1, & \text{otherwise.}\end{cases} \tag{17}$$

Compared with the original codec, the proposed encoder computes $\boldsymbol{N}'$ and hash_flag, and then sends hash_flag into the bitstream additionally; if hash_flag =1, it will compute the hash signature of $\boldsymbol{X}$ and send it into the bitstream; otherwise, no hash signature is generated and transmitted. Since the computing of the hash signature is more complex than the computing of $\boldsymbol{N}'$ and hash_flag, it is possible that the proposed scheme saves encoding computation.

The proposed decoder will detect the syntax element hash_flag in the bitstream for each block of the DVC mode. If hash_flag=1 is detected, then the decoder will parse the bitstream to obtain the hash signature of $\boldsymbol{X}$ and decode the block as described in

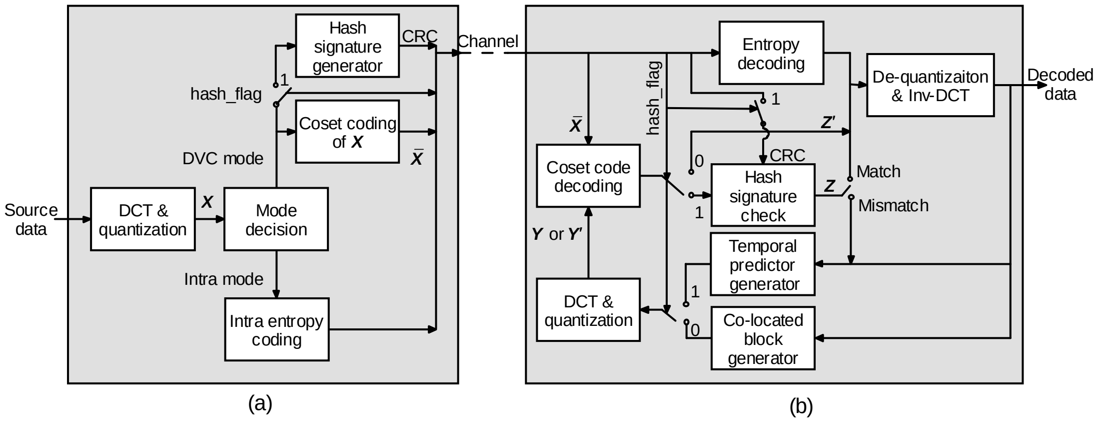

Fig. 6 The proposed codec: (a) encoder; (b) decoder

Section 2; otherwise, the co-located information $\boldsymbol{Y}'$ is adopted to decode the current block by Eq. (3) without the motion search or hash check.

To analyze the effectiveness of the proposed scheme, let B_{proposed} denote the average bits of hash_flag and the hash signature for each block: $B_{\text{proposed}}=17-16\alpha_{\text{co-located}}$. The necessary condition for the improvements of coding efficiency is then: $B_{\text{proposed}}<16$, i.e., $\alpha_{\text{co-located}}>6.25\%$. According to the theoretical results of $\alpha_{\text{co-located}}$ listed in Table 1, the coding efficiency can be improved by the proposed scheme when applied to most video sequences.

5 Experimental results

Experiments were carried out to evaluate the performance of the proposed scheme. Each input video sequence was divided into separate groups of pictures (GOP). The first frame (namely the key frame) in each GOP was coded using the intra mode as mentioned in Section 2. The other frames in the GOP, denoted as Wyner-Ziv (WZ) frames, were coded using all available modes. The decoding procedure of WZ frames requires the previously decoded frames as the reference frame at the decoder. In these experiments, only one reference frame was used, although there could be more reference frames for each WZ frame.

The first experiment evaluated the bit-rate and encoding time reduction. The results are shown in Table 2, showing that the proposed scheme saved on average 83.2% of hash signatures, 13.3% of bit-rate, and 3.9% of encoding time.

Table 2 Performance comparison between the proposed system and the original system*

Sequence	Overhead (%)	ΔHash (%)	ΔBit-rate (%)	ΔTime (%)
mother_daughter_cif	17.9	−91.3	−16.5	0.4
bridge-close_cif	18.2	−86.2	−15.4	−6.8
hall_monitor_cif	19.7	−80.6	−16.7	−7.6
paris_cif	11.4	−85.9	−9.7	−8.4
news_cif	12.0	−80.4	−9.8	−0.5
container_cif	15.2	−80.1	−12.4	−2.4
hall_qcif	15.6	−78.1	−12.8	−1.8
Average	15.7	−83.2	−13.3	−3.9

* Seven sequences of different resolutions were used to evaluate the performance of the codec, 280 frames of each sequence were encoded, and GOP=4. The frame rate was 10 Hz. Four QPs were chosen: 14, 20, 25, and 30. Overhead: the overhead of hash signatures (i.e., the percentage of bits) in the original bitstream; ΔHash: the percentage of hash signature saved by the proposed scheme, also the experimental result of $\beta=\alpha_{\text{co-located}}/\alpha$, ΔBit-rate: the change of bit-rate, calculated according to Bjontegaard (2001); ΔTime: the reduction in time consumed by the encoding procedure

Rate-distortion curve comparisons are presented in Figs. 7a–7d. The results show that the proposed algorithm provides on average a 0.5–0.75 dB gain of decoded image quality on these video sequences, compared to the original scheme.

Furthermore, we made comparisons between our proposed scheme and reference systems reported in the literature. In Fig. 7a, the reference system is a feedback-based scheme with LDPC codes and unsupervised motion vector learning from Varodayan *et al.* (2008). Our proposed scheme (GOP=8) outperforms this scheme by about 1–2 dB. Our proposed scheme (GOP=8) also outperforms the feedback-free scheme with an adaptive puncturing rate of Asif and Soraghan

(2008) by about 1–3 dB (Fig. 7b), the feedback-based system with the hash technique of Taewon *et al.* (2009) by about 0–2 dB at low bit-rate, and the DISCOVER system of Artigas *et al.* (2007) by about 0–2 dB at low bit-rate (Fig. 7c). Fig. 7d shows that our proposed scheme (only the first frame is coded as the key frame) outperforms the PRISM system of Puri and Ramchandran (2003) by about 1 dB, although in the literature the unsuccessfully decoded blocks were not taken into account during the calculation of the peak signal-to-noise ratio (PSNR).

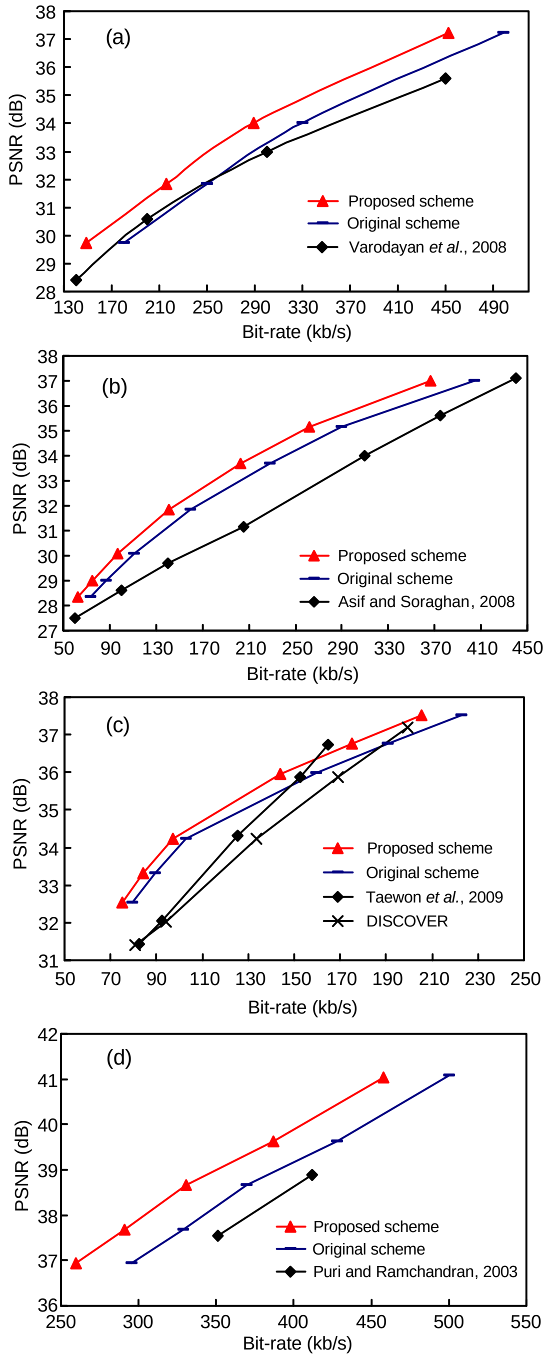

Fig. 7 Rate-distortion curves for forman_qcif at 15 Hz (a), salesman_qcif at 30 Hz (b), hall_qcif at 15 Hz (c), and carphone_qcif at 15 Hz (d)

6 Conclusions

In this paper, the noises between the source block and the co-located block in the reference frame are modeled as Laplacian random variables, like the modeling of the correlation noises between the source block and its temporal predictor. The probability of correct decoding with the co-located block is presented as a function of the probability of correct decoding with the temporal predictor. Based on theoretical insights and the results of further statistical experiments, a practical scheme, the hash signature saving scheme, is proposed for improving the performance of feedback-free DVC systems. The experimental results show that the proposed scheme saves on average 83.2% of hash signatures, 13.3% of bit-rate, and 3.9% of encoding time.

References

Aaron, A., Zhang, R., Girod, B., 2002. Wyner-Ziv Coding of Motion Video. Asilomar Conf. on Signals, Systems and Computers, **1**:240-244. [doi:10.1109/ACSSC.2002.1197184]

Aaron, A., Rane, S., Setton, E., Girod, B., 2004a. Transform-domain Wyner-Ziv codec for video. *SPIE*, **5308**:520. [doi:10.1117/12.527204]

Aaron, A., Rane, S., Girod, B., 2004b. Wyner-Ziv Video Coding with Hash-Based Motion Compensation at the Receiver. Int. Conf. on Image Processing, p.3097-3100. [doi:10.1109/ICIP.2004.1421768]

Artigas, X., Ascenso, J., Dalai, M., Klomp, S., Kubasov, D., Ouaret, M., 2007. The DISCOVER Codec: Architecture, Techniques and Evaluation. Picture Coding Symp., p.1-4.

Ascenso, J., Pereira, F., 2007. Adaptive Hash-Based Side Information Exploitation for Efficient Wyner-Ziv Video Coding. Int. Conf. on Image Processing, p.1157-1160. [doi:10.1109/ICIP.2007.4379238]

Asif, M., Soraghan, J.J., 2008. Wyner Ziv Codec Design for Surveillance System Using Adaptive Puncturing Rate. 3rd Int. Symp. on Communications, Control and Signal Processing, p.1454-1459. [doi:10.1109/ISCCSP.2008.4537456]

Bjontegaard, G., 2001. Calculation of Average PSNR Differences Between RD-Curves. VCEG-M33, 13th Meeting.

Brites, C., Pereira, F., 2007. Encoder Rate Control for Transform Domain Wyner-Ziv Video Coding. IEEE Int. Conf.

on Image Processing, p.569-572. [doi:10.1109/ICIP.2007.4379078]
Cote, G., Erol, B., Gallant, M., Kossentini, F., 1998. H.263+: Video coding at low bit rates. *IEEE Trans. Circ. Syst. Video Technol.*, **8**(7):849-866. [doi:10.1109/76.735381]
Dufaux, F., Gao, W., Tubaro, S., Vetro, A., 2009. Distributed video coding: trends and perspectives. *EURASIP J. Image Video Process*. [doi:10.1155/2009/508167]
Fowler, J.E., 2005. An Implementation of PRISM Using QccPack. Technical Report MSSU-COE-ERC-05-01, Mississippi State University. Available from http://www.ece.msstate.edu/~fowler/Publications/Papers/Fow2005.pdf
Girod, B., Aaron, A., Rane, S., Rebollo-Monedero, D., 2005. Distributed video coding. *Proc. IEEE*, **93**(1):71-83. [doi:10.1109/JPROC.2004.839619]
Guillemot, C., Pereira, F., Torres, L., Ebrahimi, T., Leonardi, R., Ostermann, J., 2007. Distributed monoview and multi-view video coding. *IEEE Signal Process. Mag.*, **24**(5):67-76. [doi:10.1109/MSP.2007.904808]
Guo, M., Lu, Y., Wu, F., Li, S.P., Gao, W., 2007. Distributed Video Coding with Spatial Correlation Exploited Only at the Decoder. IEEE Int. Symp. on Circuits and Systems, p.41-44. [doi:10.1109/ISCAS.2007.378177]
Hua, G., Chen, C.W., 2008. Distributed Video Coding with Zero Motion Skip and Efficient DCT Coefficient Encoding. IEEE Int. Conf. on Multimedia and Expo, p.777-780. [doi:10.1109/ICME.2008.4607550]
ISO/IEC 13818-2:1994. Generic Coding of Moving Pictures and Associated Audio Information—Part 2: Video.
ISO/IEC 14496-10:2003. Coding of Audiovisual Objects—Part 10: Advanced Video Coding.
ITU-T Recommendation H.264, 2003. Advanced Video Coding for Generic Audiovisual Services.
Kuganeswaran, T., Fernando, X., Guan, L., 2008. Distributed Video Coding and Transmission over Wireless Fading Channel. Canadian Conf. on Electrical and Computer Engineering, p.1513-1516. [doi:10.1109/CCECE.2008.4564794]
Mukherjee, D., 2009. Parameter selection for Wyner-Ziv coding of Laplacian sources with additive Laplacian or Gaussian innovation. *IEEE Trans. Signal Process.*, **57**(8):3208-3225. [doi:10.1109/TSP.2009.2018617]
Ostermann, J., Bormans, J., List, P., Marpe, D., Narroschke, M., Pereira, F., Stockhammer, T., Wedi, T., 2004. Video coding with H.264/AVC: tools, performance, and complexity. *IEEE Circ. Syst. Mag.*, **4**(1):7-28. [doi:10.1109/MCAS.2004.1286980]
Pereira, F., Torres, L., Guillemot, C., Ebrahimi, T., Leonardi, R., Klomp, S., 2008. Distributed video coding: selecting the most promising application scenarios. *Signal Process. Image Commun.*, **23**(5):339-352. [doi:10.1016/j.image.2008.04.002]
Puri, R., Ramchandran, K., 2003. PRISM: a “Reversed” Multimedia Coding Paradigm. IEEE Int. Conf. on Image Processing, p.617-620. [doi:10.1109/ICIP.2003.1247037]
Puri, R., Majumdar, A., Ramchandran, K., 2007. PRISM: a video coding paradigm with motion estimation at the decoder. *IEEE Trans. Image Process.*, **16**(10):2436-2448. [doi:10.1109/TIP.2007.904949]
Slepian, J.D., Wolf, J.K., 1973. Noiseless coding of correlated information sources. *IEEE Trans. Inform. Theory*, **19**(4):471-480. [doi:10.1109/TIT.1973.1055037]
Taewon, D., Hiuk, J.S., Byeungwoo, J., 2009. Motion Linearity Based Skip Decision for Wyner-Ziv Coding. 2nd IEEE Int. Conf. on Computer Science and Information Technology, p.410-413. [doi:10.1109/ICCSIT.2009.5234792]
Varodayan, D., Chen, D., Flierl, M., Girod, B., 2008. Wyner-Ziv coding of video with unsupervised motion vector learning. *Signal Process. Image Commun.*, **23**(5):369-378. [doi:10.1016/j.image.2008.04.009]
Wedi, T., Musmann, H.G., 2003. Motion- and aliasing-compensated prediction for hybrid video coding. *IEEE Trans. Circ. Syst. Video Technol.*, **13**(7):577-587. [doi:10.1109/TCSVT.2003.815171]
Wyner, A.D., 1974. Recent results in the Shannon theory. *IEEE Trans. Inform. Theory*, **20**(1):2-10. [doi:10.1109/TIT.1974.1055171]
Wyner, A.D., Ziv, J., 1976. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. Inform. Theory*, **22**(1):1-10. [doi:10.1109/TIT.1976.1055508]
Yang, S.T., Zhao, M.J., Qiu, P.L., 2007. On Wyner-Ziv problem for general sources with average distortion criterion. *J. Zhejiang Univ.-Sci. A*, **8**(8):1263-1270. [doi:10.1631/jzus.2007.A1263]