Adaptive Correlation Noise Model for DC Coefficients in Wyner-Ziv Video Coding

Hao Qin, Bin Song, Yue Zhao, and Haihua Liu

An adaptive correlation noise model (CNM) construction algorithm is proposed in this paper to increase the efficiency of parity bits for correcting errors of the side information in transform domain Wyner-Ziv (WZ) video coding. The proposed algorithm introduces two techniques to improve the accuracy of the CNM. First, it calculates the mean of direct current (DC) coefficients of the original WZ frame at the encoder and uses it to assist the decoder to calculate the CNM parameters. Second, by considering the statistical property of the transform domain correlation noise and the motion characteristic of the frame, the algorithm adaptively models the DC coefficients of the correlation noise with the Gaussian distribution for the low motion frames and the Laplacian distribution for the high motion frames, respectively. With these techniques, the proposed algorithm is able to make a more accurate approximation to the real distribution of the correlation noise at the expense of a very slight increment to the coding complexity. The simulation results show that the proposed algorithm can improve the average peak signal-to-noise ratio of the decoded WZ frames by 0.5 dB to 1.5 dB.

Keywords: Wyner-Ziv video coding, correlation noise model, transform domain, Gaussian distribution, Laplacian distribution.

Manuscript received Apr. 29, 2011; revised Sept. 19, 2011; accepted Oct. 21, 2011.

This work was supported by the National Natural Science Foundation of China (No. 60802032), the 111 Project (B08038), the Fundamental Research Funds for the Central Universities (K50511010020), and also supported by ISN State Key Laboratory.

Hao Qin (phone: +86 29 8820 4409, hqin@mail.xidian.edu.cn), Bin Song (corresponding author, bsong@mail.xidian.edu.cn), Yue Zhao (yueliang1019@yeah.net), and Haihua Liu (haihualiu1985@163.com) are with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, China.

http://dx.doi.org/10.4218/etrij.12.0111.0273

I. Introduction

Traditional video compressing standards, for example, MPEG or H.26x, employ a complex motion estimation framework in their encoders to analyze the correlation between the temporally adjacent frames, so they usually include a high-complexity encoder and a low-complexity decoder. This architecture is well-suited for applications that the video contents are coded only once but decoded many times, and for streaming scenarios where the encoder has much more computational resources available than the decoder. However, more and more new applications which have emerged in recent years, such as wireless low-power surveillance and multimedia sensor networks, wireless cameras, and mobile camera phones, require light encoders. To meet this requirement, a new video coding scheme, distributed video coding (DVC), was proposed to enable exploration of the video statistics, partially or totally, in the decoder [1], [2] and lower the encoder's complexity. DVC lays its theoretical foundation on the distributed source coding principles stated by the Slepian-Wolf and the Wyner-Ziv (WZ) theorems [3], [4]. Two DVC solutions, the WZ video coding system [5], [6] and the PRISM [7], proposed by Bernd Girod's group at Stanford University and Ramchandran's group at the University of Berkeley, California, respectively, have received most of the attention because of their practicality. Compared with the PRISM, the WZ video coding system has a lighter encoder and can better exploit the temporal and spatial correlation. Therefore, most of the practical DVC implementations are based on the WZ video coding architecture.

To utilize the parity bits to correct errors of the side information in transform domain efficiently, the decoder in a WZ video coding system needs to have reliable knowledge

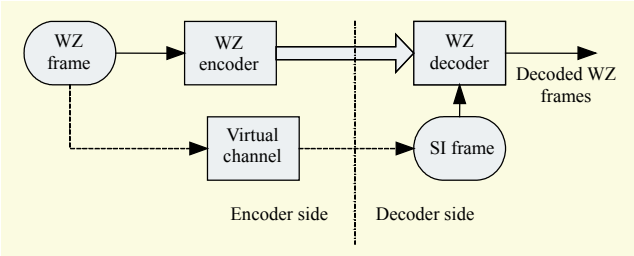

Fig. 1. Relation between WZ frame and SI frame.

about how to model the correlation noise between the original WZ frame and its estimation frame in the decoder, which is also known as the side information (SI) frame [8], [9]. Because the SI could be considered a "corrupted" version of the original information, the correlation noise (*WZ–SI*) can also be interpreted as a virtual channel with an error pattern that satisfies a certain statistical distribution. The relationship of the original WZ frame and its corresponding SI frame can be further illustrated with Fig. 1. As we know, the more accurately the distribution describes (*WZ–SI*) in the transform domain, the higher the rate distortion (RD) performance [10] is. However, it is very difficult in practice to precisely model the correlation noise in the decoder since the original WZ frame is unavailable in the decoder [11].

The Laplacian distribution with location parameter zero is a correlation noise model (CNM) widely used in the WZ video coding literatures [11]. The CNM can be constructed in two ways, online and offline. The online CNM uses the motion-compensated residual in the decoder to estimate the residual between the original WZ frame and its corresponding SI frame. Contrarily, the offline CNM uses the original WZ frame and the SI frame to achieve the residual. Thus, the performance of offline CNM provides insights on the maximum or "ideal" estimation performance that can be achieved, but offline CNM is impracticable because it requires either the SI frame be available in the encoder or the original WZ frame be available in the decoder. On the other hand, the online CNM also has its own drawbacks. First, its accuracy is limited because it can only provide the estimation of the CNM parameter in the decoder. Second, the online CNM presented in [11] does not take the statistical property of the transform domain correlation noise and the video content scene change into consideration.

To improve the accuracy of the online CNM, an adaptive CNM construction algorithm in transform domain WZ video coding is proposed in this paper. The proposed algorithm utilizes two methods to address the accuracy issue. First, the encoder supplies the mean of the DC coefficients of the original WZ frame to the decoder to assist constructing the CNM. Second, the decoder adaptively chooses different CNM distributions for DC coefficients based on the video content scene change.

This paper is organized as follows. Section II gives a brief overview to the WZ video coding architecture. The statistical property of the transform domain correlation noise is discussed in section III. Section IV presents the adaptive CNM construction algorithm in detail. The proposed algorithm is simulated and the obtained numeric results are given in section V. Finally, the conclusion is provided in section VI.

II. Transform Domain WZ Video Coding System

Figure 2 illustrates the architecture of a transform domain WZ video coding system without the feedback channel. As shown in the figure, the video frames are categorized into key frames and WZ frames. The key frames are coded using a traditional intra-frame video encoder such as H.264/AVC, while the WZ frames are intraframe encoded but interframe decoded.

In the encoder, integer 4×4 discrete cosine transform (DCT) is firstly applied to each WZ frame block by block. The obtained DCT coefficients are further grouped into 16 DCT coefficient bands $X_1, X_2, \ldots, X_{16}$, where band X_k consists of all the coefficients that occupy the same position k within the 4×4 blocks. Then, uniform quantization with 2^{M_k} levels (where the number of levels 2^{M_k} depends on the DCT coefficient band) is performed over each DCT coefficient band to get the quantized coefficients q_k. Finally, the bitplanes, extracted from q_k, are fed into a low density parity check (LDPC) encoder to generate the parity bits for transmission.

The decoder utilizes a motion-compensated frame interpolation framework to produce the SI frame for each WZ frame based on its forward and backward adjacent key frames. Next, the same processes applied to the original WZ frame in the encoder, including DCT, quantization, and bitplane extraction, are also performed on the SI frame in the same order. Then the decoder calculates the residual frame Z as an approximation to (*WZ–SI*) in the pixel domain using (1).

$$Z(x,y) = \frac{f_{\mathrm{f}}(x+dx_{\mathrm{f}}, y+dy_{\mathrm{f}}) - f_{\mathrm{b}}(x+dx_{\mathrm{b}}, y+dy_{\mathrm{b}})}{2}, \quad (1)$$

where $f_{\mathrm{f}}(x+dx_{\mathrm{f}}, y+dy_{\mathrm{f}})$ and $f_{\mathrm{b}}(x+dx_{\mathrm{b}}, y+dy_{\mathrm{b}})$ represent the forward and backward motion-compensated frames, respectively, and dx_{f}, dy_{f} and dx_{b}, dy_{b} denote the motion vectors for frames f_{f} and f_{b}, respectively. On the basis of residual frame Z, the CNM (for example, the online CNM or the proposed adaptive CNM) can be built for each DCT coefficient. By feeding the received parity bits from the encoder, SI frame, and parameters of CNM into the LDPC decoder, the bitplanes of the WZ frame are recovered. After combining the recovered

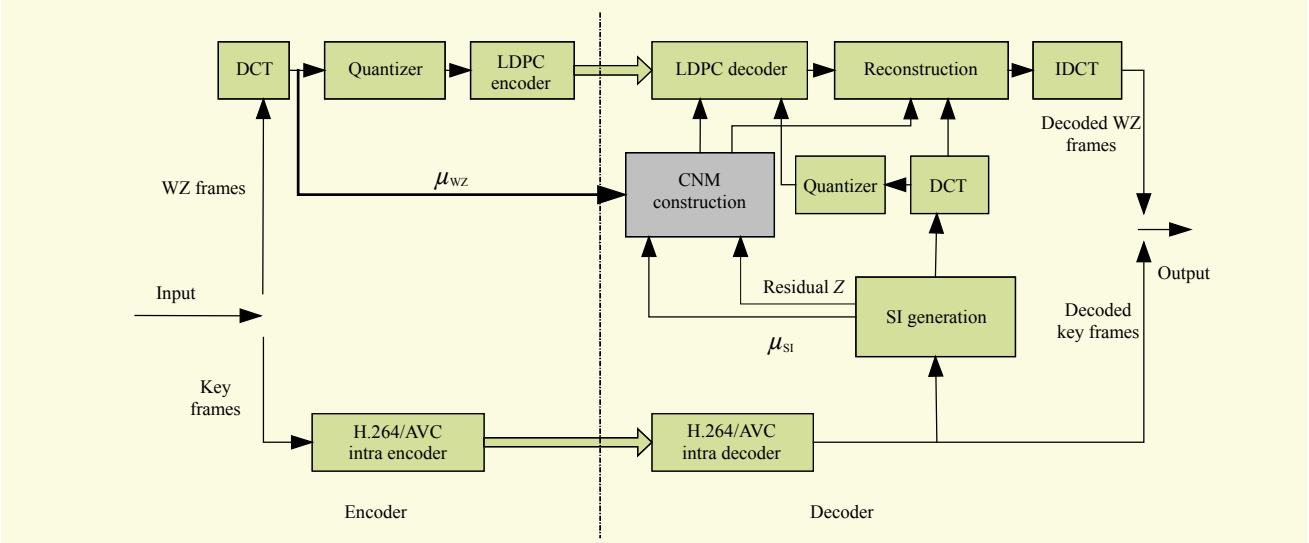

Fig. 2. Transform domain unidirectional WZ video coding system.

bitplanes into quantized coefficients, the WZ frame is decoded by dequantization, reconstruction and IDCT. In Fig. 2, μ_{WZ} and μ_{SI} represent the mean of the DC coefficients in the original WZ frame and the SI frame, respectively, and these two values are critical to the construction of the proposed adaptive CNM.

Because of the statistical correlation between the original WZ frame and the SI frame, the accuracy of the CNM is essential to the RD performance in WZ video coding, especially in unidirectional WZ video coding.

III. Adaptive CNM of Transform Domain WZ Video Coding

The existing CNM chooses the Laplacian distribution with location parameter zero to model the correlation noise for all kinds of video content scene changes. Although it can provide an acceptable performance, the Laplacian distribution may not be preferable for certain motion sequences. In this section, some simulation results will be presented and summarized to demonstrate how the statistical distributions of actual correlation noise (ACN) differ for different motion sequences. Note that the GOP size is set to 2 in these simulations. Thus, the GOP structure of "key frame, WZ frame, key frame, WZ frame …" was adopted.

The DC and AC1 coefficient bands of the 71st and 119th WZ frame in the quarter common intermediate format (QCIF) sequence "Foreman" are plotted in Fig. 3. Several observations can be made based on the figure. First, the AC1 component of actual correlation noise (ACN_AC1) is distributed around zero while the DC component of actual correlation noise (ACN_DC) is distributed around a positive value. Second, according to Figs. 3(a) and 3(b), the statistical distribution of ACN_DC fits better to the Laplacian distribution and the Gaussian distribution for the 71st and 119th WZ frames, respectively. This phenomenon indicates that the Laplacian distribution with location parameter zero is insufficient to approximate the distribution of ACN_DC.

One of the differences between the 71st and 119th WZ frame in "Foreman" is that the former is a high-motion frame while the latter is a smooth one when compared with its adjacent frames. Our further observations on the other sequences also imply that the motion characteristic of the video frame might have a great influence on the statistical distribution of ACN_DC.

To verify the above perceptions, we use the curve fitting tool of MATLAB to test the goodness-of-fit (GOF) of ACN_DC to the Laplacian distribution or the Gaussian distribution. The GOF is defined as

$$R = 1 - \frac{\sum_{i=1}^{n} (f(x_i) - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}, \quad (2)$$

where x_i denotes the i-th observed ACN value, $f(x_i)$ is the probability density of the Laplacian distribution or the Gaussian distribution at x_i, y_i is the actual probability density at x_i, and $\bar{y}$ is the average of y_i. It should be noted that the closer R approaches to 1, the better the goodness.

The GOF test results of DC coefficient band for different QCIF sequences are summarized in Table 1, where R_L and R_G is the GOF of ACN_DC with the Laplacian distribution and the Gaussian distribution, respectively. From Table 1, it can be seen that some sequences favor the Gaussian distribution while

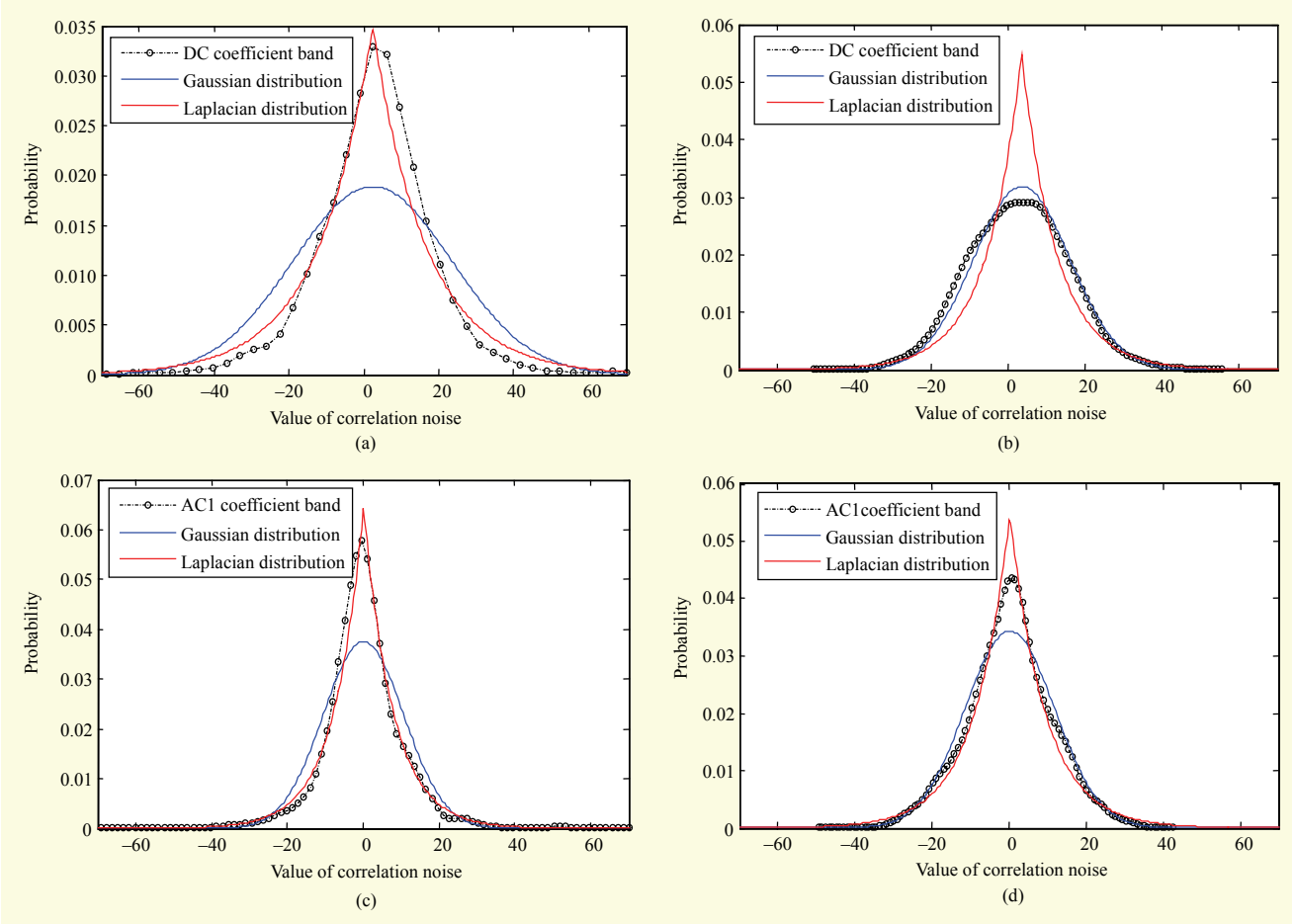

Fig. 3. Distribution of actual correlation noise: (a) 71st (WZ-SI) of Foreman, (b) 119th (WZ-SI) of Foreman, (c) 71st (WZ-SI) of Foreman, and (d) 119th (WZ-SI) of Foreman.

Table 1. GOF of DC coefficient band.

Sequence, total WZ frames	Laplacian better than Gaussian distribution			Gaussian better than Laplacian distribution		
	Frame count	$\overline{R_L}$	$\overline{R_G}$	Frame count	$\overline{R_L}$	$\overline{R_G}$
Foreman, 149	56	0.8728	0.72048	93	0.7806	0.9328
Stefan, 149	135	0.8412	0.6779	14	0.8853	0.9207
Coastguard, 149	2	0.9063	0.7778	147	0.7434	0.9405
Hall, 149	34	0.8938	0.8467	105	0.8928	0.8232
Bus, 74	45	0.8810	0.7759	29	0.8539	0.9074
Suzie, 74	20	0.9243	0.7761	54	0.9478	0.8565
Mother-Daughter, 149	38	0.8891	0.8080	111	0.8855	0.8207
Miss America, 74	27	0.9411	0.9060	47	0.9207	0.877
News, 149	93	0.9078	0.7872	56	0.9200	0.8857

others favor the Laplacian distribution. For example, 93 of the 149 WZ frames in the "Foreman" sequence have a GOF of ACN_DC closer to 1 via the Gaussian distribution rather than the Laplacian distribution ($\overline{R_L}$ =0.7806; $\overline{R_G}$ =0.9328); in other words, the distribution of ACN_DC in the majority of the frames is closer to the Gaussian distribution.

Table 2. GOF of AC1 coefficient band.

Sequence, total WZ frames	Laplacian better than Gaussian distribution			Gaussian better than Laplacian distribution		
	Frame count	$\overline{R_L}$	$\overline{R_G}$	Frame count	$\overline{R_L}$	$\overline{R_G}$
Foreman, 149	116	0.9252	0.7664	33	0.8472	0.9301
Stefan, 149	149	0.6839	0.4533	0	–	–
Coastguard, 149	139	0.9569	0.8716	10	0.9102	0.9263
Hall, 149	149	0.8521	0.6350	0	—	—
Bus, 74	74	0.8998	0.6877	0	–	–
Suzie, 74	74	0.9505	0.7967	0	–	–
Mother-Daughter, 149	149	0.9103	0.6813	0	–	–
Miss America, 74	74	0.8826	0.6299	0	–	–
News, 149	149	0.8567	0.5883	0	–	–

Table 1 shows that for the low- to medium-motion sequences, such as "Foreman," "Hall," and "Coastguard," more frames prefer to use the Gaussian distribution to model ACN_DC. However, for high-motion sequences, such as "Stefan" and "Bus," a frame's ACN_DC is more likely modeled using the Laplacian distribution.

The GOF of ACN_AC1 distribution was also analyzed, and the results are given in Table 2. It is clear that the ACN_AC1 of all frames fits closer to the Laplacian distribution. Further analyses for other AC bands also demonstrate similar GOF results, which are not included in the paper due to limited space.

Based on the above observations, it can be concluded that all the AC bands of ACN should be modeled by the Laplacian distribution with location parameter zero while ACN_DC should be modeled by the Laplacian distribution or the Gaussian distribution depending on the motion characteristics of the WZ frame. Therefore, a new adaptive transform domain CNM is proposed in (3).

$$f(n)=\begin{cases}\dfrac{1}{2b}e^{-\frac{|n|}{b}}, & \text{for all the AC bands of each frame,}\\ \dfrac{1}{2b}e^{-\frac{|n-\mu|}{b}}, & \text{for the DC band of a high-motion frame,}\\ \dfrac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{(n-\mu)^2}{2\sigma^2}}, & \text{for the DC band of a low-motion frame,}\end{cases} \tag{3}$$

where $f(n)$ is the probability density function of the correlation noise, μ and σ^2 denote the mean and the variance of the CNM, respectively, and b represents the scale parameter of the Laplacian distribution that can be directly calculated with $b=\sqrt{\sigma^2/2}$.

According to (3), the motion characteristic of a WZ frame must be identified before the CNM for that frame can be determined. Let *SC* denote the motion characteristic of the original WZ frame compared with its two adjacent decoded key frames. The larger the *SC* of a WZ frame is, the more likely it is that the frame has high-motion contents. Because the original WZ frame is not available in the decoder, it is common that the decoder uses the corresponding SI frame as an approximation of the WZ frame. Therefore, the residual frame Z is actually an indicator as to the similarity of two blocks (motion compensated) in the reference frames and can be used to approximate *SC* with (4).

$$SC=\sum_x\sum_y\left[WZ(x,y)-SI(x,y)\right]^2\approx\sum_x\sum_y Z^2(x,y). \tag{4}$$

Given the *SC* of a WZ frame, it is possible to tell whether the motion of the frame is high or low depending on a threshold N. The frame contains high-motion contents if $SC>N$; otherwise, the frame has a smooth motion. The accuracy of the frame motion estimation can be greatly improved if we can update threshold N adaptively with (5) whenever the current WZ frame is regarded as a high-motion frame.

$$N=\varepsilon\cdot N+(1-\varepsilon)\cdot SC, \tag{5}$$

where $\varepsilon\in[0,1]$ is an empirical parameter.

First, N is initialized to the *SC* value of the first WZ frame. Then N is updated recursively with (5) whenever a WZ frame is regarded as a high-motion frame. It is particularly important to carefully choose ε as it has a direct effect on the coding performance. The larger ε is, the less impact the current WZ frame motion will have on N. In other words, N tends to keep stable for large ε, and the proposed model may even fail to reflect the video motion trends if N is set too large. On the contrary, a small ε may lead to dramatic fluctuations of N and make the algorithm unstable. Based on the above discussions, ε

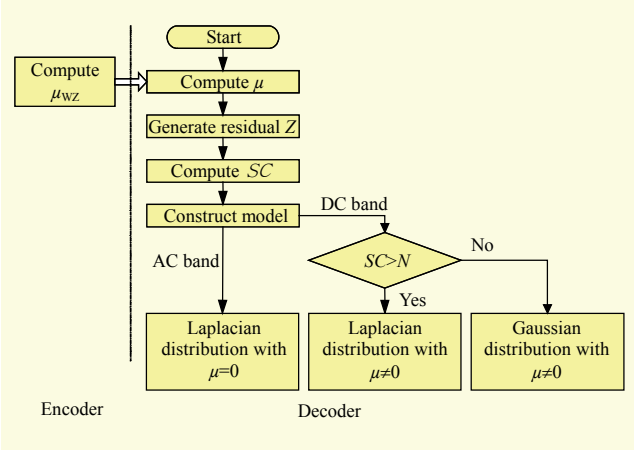

Fig. 4. Construction algorithm of adaptive CNM.

should be empirically chosen to achieve a proper tradeoff between stability and adaptability.

IV. Adaptive CNM Construction Algorithm

Based on the adaptive CNM given above, an adaptive CNM construction algorithm in transform domain WZ video coding is proposed and its basic flow is diagrammed in Fig. 4. The algorithm employs an approach similar to the one proposed in section VII of [11]. Its basic principle is to estimate the distribution parameter of ACN at two different granularity levels: the DCT band level and the coefficient level. In the DCT band level, one distribution parameter is estimated for each DCT band. Those parameters are updated for each frame and thus vary along the video sequence. Next, the distribution parameter of each DCT band is adjusted further for each coefficient within that DCT band, leading to a finer adaptation of the CNM at the coefficient level. What follows is a detailed discussion of how the adaptive CNM is constructed.

As shown in Fig. 4, μ_{WZ}, the mean of DC coefficients of the original WZ frame, is calculated in the encoder and transmitted to the decoder. In the decoder, μ_{SI}, the average of DC coefficients of the corresponding SI frame, is calculated before $\mu = \mu_{WZ} - \mu_{SI}$, the mean of ACN_DC, is determined. Then the decoder will calculate the residual frame Z and SC of that frame using (1) and (4), respectively. With the obtained μ, Z, and SC, the decoder will be able to calculate the mean μ_i and the variance σ_i^2 and construct the DCT band level CNM using (3) accordingly for each DCT band, where $i = 1, 2, ..., 16$. The above DCT band level CNMs still need to be further refined at the coefficient level with the following steps.

Step 1. Adjust the variance for each coefficient according to

$$\Delta_i(m) = |T_i(m)| - \mu_i, \tag{6}$$

$$\sigma_i(m) = \begin{cases} \sigma_i, & [\Delta_i(m)]^2 \le \sigma_i^2, \\ \Delta_i(m), & [\Delta_i(m)]^2 > \sigma_i^2, \end{cases} \tag{7}$$

where $|T_i(m)|$ and $\sigma_i^2(m)$ are the absolute value and the variance of the m-th coefficient in the i-th DCT band of the residual frame Z, respectively, and $\Delta_i(m)$ is the distance between $|T_i(m)|$ and μ_i.

Step 2. The Laplacian distribution with location parameter zero and scaling parameter $b_i(m) = \sqrt{\sigma_i^2(m)/2}$ is used to model the coefficients of all the AC bands. The motion characteristic of a frame, determined based on SC and N, is used for modeling the DC coefficients. Specifically, the Laplacian distribution with scaling parameter $b_i(m)$ and location parameter μ will be chosen to model each coefficient of ACN_DC for each high-motion frame; otherwise, the Gussian distribution with mean μ and variance $\sigma_i^2(m)$ will be used.

Step 3. Update N with (5), accordingly.

The computational complexity issue may be briefly analyzed, as below. The proposed algorithm requires the encoder to calculate the mean of DC coefficients for each original WZ frame. Because all the DC coefficients are always available in the transform domain, the complexity incurred in the encoder by the proposed algorithm may be ignored. In the decoder, the Gaussian distribution introduced in this paper can be fully modeled with only two parameters, μ and σ^2. These parameters are also used to determine the Laplacian distribution because the scaling parameter b of the distribution can be achieved with $b = \sqrt{\sigma^2/2}$. As a result, although the Gaussian distribution is used in the proposed CNM, there is no extra complexity overhead incurred in the decoder.

V. Experiment Results

To demonstrate the performance of the proposed adaptive CNM and the corresponding construction algorithm, several QCIF (176×144) and CIF (352×288) video sequences, including "Foreman," "Coastguard," "Suzie," "Miss America," "Hall," and "Mother-Daughter," were tested in the WZ video coding system given in Fig. 2. The GOP structure adopted in the experiments is "KWKW…," where K and W denote the key frame and WZ frame, respectively. The key frames are encoded with the intramode of JM14.0 [12]. Note that ε=0.5 is empirically chosen in (5) when the proposed CNM is applied.

In the experiment, the sum square error (SSE) between the distribution of ACN and the CNM is used as the criteria to measure the accuracy of the CNM, and SSE is formulated as

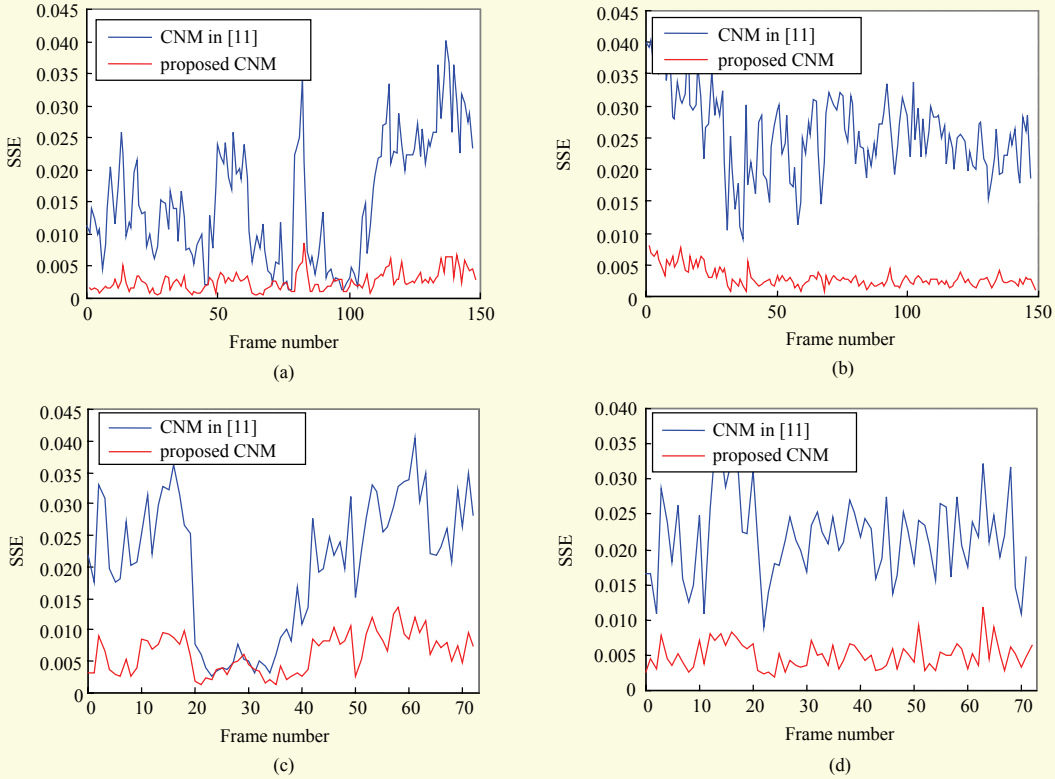

Fig. 5. Comparison of online CNM and adaptive CNM for different QCIF sequences: (a) Foreman @30 Hz, (b) Coastguard @30 Hz, (c) Suzie @30 Hz, and (d) Miss America @30 Hz.

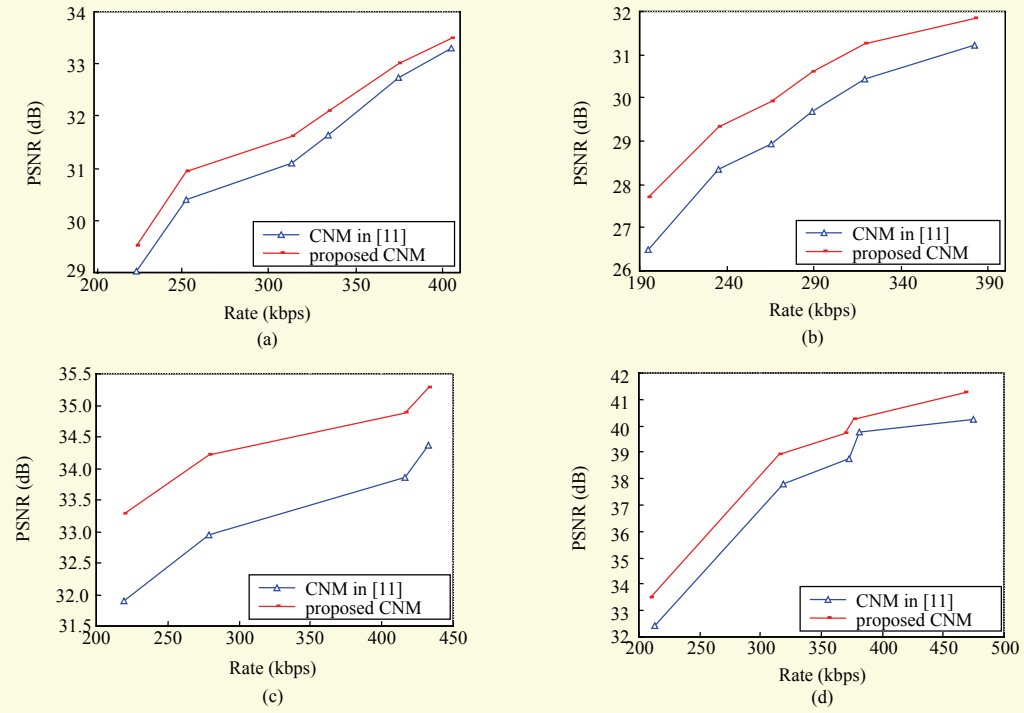

Fig. 6. RD performance for different QCIF sequences: (a) Foreman @30 Hz, (b) Coastguard @30 Hz, (c) Suzie @30 Hz, and (d) Miss America @30 Hz.

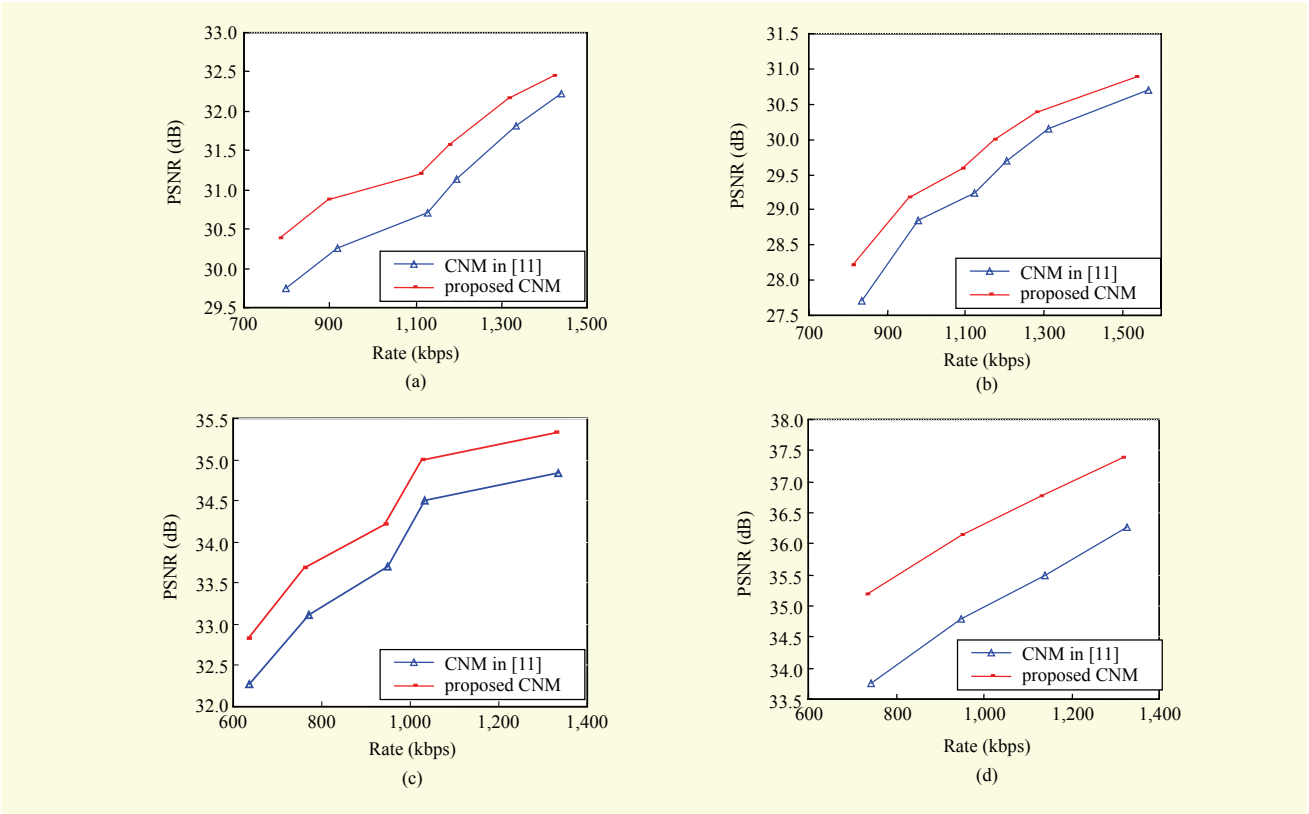

Fig. 7. RD performance for different CIF sequences: (a) Foreman @30 Hz, (b) Coastguard @30 Hz, (c) Hall @30 Hz, and (d) Mother-Daughter @30 Hz.

$$SSE = \sum_{i=1}^{n} (f(x_i) - y_i)^2, \quad (8)$$

where x_i, $f(x_i)$, and y_i represent the i-th sample of ACN_DC, the probability density at x_i given by the CNM and the actual probability density of ACN_DC at x_i, respectively. An SSE comparison between the online CNM in [11] and the proposed adaptive CNM was conducted, and the results are plotted in Fig. 5. It is clear, from Fig. 5, that the proposed adaptive CNM describes the actual correlation noise better than the online CNM in [11].

The average peak signal-to-noise ratio (PSNR) of the decoded WZ frames, when the same sequences are applied, was also studied for both approaches. The results of the QCIF sequences given in Fig. 6 show that the proposed adaptive CNM offers an average PSNR improvement by up to 0.5 dB for "Foreman," about 1 dB for "Coastguard" and "Miss America," and more than 1 dB for "Suzie." For CIF sequences, Fig. 7 shows that the improvements can be up to 0.5 dB for "Foreman," "Coastguard," and "Hall," and more than 1 dB for "Mother-Daughter."

Based on the observations made above, it can be concluded that the proposed adaptive CNM not only has a clear advantage over the online CNM [11] in terms of the accuracy of the CNM, but also greatly improves the video quality of the decoded WZ frames.

VI. Conclusion

This paper proposed an adaptive CNM construction algorithm for a transform domain WZ video coding system. By utilizing the assistant information transmitted from the encoder, the statistical property of the transform domain correlation noise, and the motion characteristic of the video content, the proposed CNM can be adaptively constructed with Gaussian or Laplacian distribution. Numerical results demonstrate that the proposed adaptive CNM can match the actual distribution of the correlation noise more closely and thus greatly improve the quality of decoded WZ frame although the proposed algorithm might introduce a minor increment on the coding complexity.

References

[1] B. Girod et al, "Distributed Video Coding," *Proc. IEEE*, vol. 93, no. 1, Jan. 2005, pp. 71-83.

[2] R. Puri et al, "Distributed Video Coding in Wireless Sensor Networks," *IEEE Signal Proc. Mag.*, vol. 23, no. 4, July 2006, pp. 94-106.
[3] D. Slepian and J. Wolf, "Noiseless Coding of Correlated Information Sources," *IEEE Trans. Info. Theory*, vol. 19, no. 4, July 1973, pp. 471-480.
[4] A. Wyner and J. Ziv, "The Rate-Distortion Function for Source Coding with Side Information at the Decoder," *IEEE Trans. Info. Theory*, vol. 22, no. 1, Jan. 1976, pp. 1-10.
[5] A. Aaron, R. Zhang, and B. Girod, "Wyner-Ziv Coding of Motion Video," *Proc. Asilomar Conf. Signals, Syst. Comput.*, 2002, pp. 240-244.
[6] A. Aaron, S. Rane, and B. Girod, "Transform-Domain Wyner-Ziv Codec for Video," *Proc. SPIE Visual Commun. Image Process. Conf.*, 2004, pp. 520-528.
[7] R. Puri, A. Majumdar, and K. Ramchandran, "PRISM: A Video Coding Paradigm With Motion Estimation at the Decoder," *IEEE Trans. Image Process.*, vol. 16, no. 10, Oct. 2007, pp. 2436-2448.
[8] C. Brites, J. Ascenso, and F. Pereira, "Studying Temporal Correlation Noise Modeling for Pixel Based Wyner-Ziv Video Coding," *IEEE Int. Conf. Image Process.*, 2006, pp. 273-276.
[9] X.P. Fan, O.C. Au, and N.M. Cheung, "Transform-Domain Adaption Correlation Estimation (TRACE) for Wyner-Ziv Video Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 11, Nov. 2010, pp. 1423-1436.
[10] X. Huang and S. Forchhammer, "Improved Virtual Channel Noise Model for Transform Domain Wyner-Ziv Video Coding," *IEEE Int. Conf. Acoustics Speech Signal Process.*, Apr. 2009, pp. 921-924.
[11] C. Brites and F. Pereira, "Correlation Noise Modeling for Efficient Pixel and Transform Domain Wyner-Ziv Video Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 9, Sept. 2008, pp. 1177-1190.
[12] H.264/AVC Software Coordination, "JM Reference Software," accessed Dec. 2010. Available: http://iphome.hhi.de/suehring/tml/download/ old_jm/jm14.0.zip

Hao Qin received his BS, MS, and PhD in communication and information systems from Xidian University, Xi'an, China, in 1996, 1999, and 2003, respectively. He is currently an associate professor of the School of Telecommunications Engineering, Xidian University. In addition to video compression theories and technologies, his research interests and areas of publication include broadband communications.

Bin Song received his BS, MS, and PhD in communication and information systems from Xidian University, Xi'an, China, in 1996, 1999, and 2002, respectively. In 2002, he joined the School of Telecommunications Engineering, Xidian University, where he is currently a professor in communications and information systems. He has authored over 30 journal papers and conference papers. His research interests and areas of publication include video compression and transmission technologies, video transcoding, error- and packet-loss-resilient video coding, distributed video coding, and video signal processing based on compressed sensing, and multimedia communications.

Yue Zhao received her BS in communication and information systems from Xidian University, Xi'an, China, in 2009. She is currently working toward an MS in communication and information systems at Xidian University. Her research interests include video compression and distributed video coding.

Haihua Liu received her BS in communication engineering from Shandong Normal University, Jinan, China, in 2008, and her MS in communication and information systems from Xidian University, Xi'an, China, in 2011. Her research interests include video compression and distributed video coding.