

# Optimal entropy-constrained non-uniform scalar quantizer design for low bit-rate pixel domain DVC

Bo Wu · Nan Zhang · Siwei Ma · Debin Zhao · Wen Gao

Published online: 24 August 2012  
© Springer Science+Business Media, LLC 2012

**Abstract** In this paper, an optimal entropy-constrained non-uniform scalar quantizer is proposed for the pixel domain DVC. The uniform quantizer is efficient for the hybrid video coding since the residual signals conforming to a single-variance Laplacian distribution. However, the uniform quantizer is not optimal for pixel domain distributed video coding (DVC). This is because the uniform quantizer is not adaptive to the joint distribution of the source and the SI, especially for low level quantization. The signal distribution of pixel domain DVC conforms to the mixture model with multi-variance. The optimal non-uniform quantizer is designed according to the joint distribution, the error between the source and the SI can be decreased. As a result, the bit rate can be saved and the video quality won't sacrifice too much. Accordingly, a better R-D trade-off can be achieved. First, the quantization level is fixed and the optimal RD trade-off is achieved by using a Lagrangian function  $J(Q)$ . The rate and distortion components is designed based on  $P(Y|Q)$ . The conditional probability density function of SI  $Y$  depend on quantization partitions  $Q$ ,  $P(Y|Q)$ , is approximated by a Guassian mixture model at encoder. Since the SI can not be accessed at encoder, an estimation of  $P(Y|Q)$  based on the distribution of the source is proposed. Next,  $J(Q)$  is optimized by an iterative Lloyd-Max algorithm with a novel quantization partition

---

B. Wu · N. Zhang (✉)  
School of Biomedical Engineering, Capital Medical University, No.10 Xitoutiao, You An Men, Beijing,  
Peoples's Republic of China  
e-mail: zhangnan@ccmu.edu.cn

S. Ma · W. Gao  
School of Electronic Engineering and Computer Science, Peking University, Beijing, China

S. Ma  
e-mail: swma@pku.edu.cn

W. Gao  
e-mail: wgao@pku.edu.cn

D. Zhao  
Harbin Institute of Technology, Harbin, China  
e-mail: dbzhao@jdl.ac.cn

updating algorithm. To guarantee the convergence of  $J(Q)$ , the monotonicity of the interval in which the endpoints of the quantizer lie must be satisfied. Then, a quantizer partition updating algorithm which considers the extreme points of the histogram of the source is proposed. Consequently, the entropy-constrained optimal non-uniform quantization partitions are derived and a better RD trade-off is achieved by applying them. Experiment results show that the proposed scheme can improve the performance by 0.5 dB averagely compared to the uniform scalar quantization.

**Keywords** DVC · Non-uniform scalar quantizer · Optimal quantizer

## 1 Introduction

Recently, the new applications such as wireless low power video surveillance and wireless sensor network are emerging. In these applications, a light encoder is required because the computation and memory resources on sensors are scarce. Accordingly, a DVC scheme which has a much low computation complexity video encoder has been proposed [1–3, 7]. With the development of the advanced channel codes, the practical DVC schemes are emerging, such as the practical Wyner-Ziv codec developed by Stanford [1, 2], the DISCOVER system [3, 7] proposed by DISCOVER DVC organization etc. The DVC coders can be classified into pixel-domain and transform-domain coders. Transform domain coders have greater coding efficiency than the pixel-domain ones. However, pixel domain DVC encoders are less complex than their transform domain counterparts [19]. Therefore, the pixel domain DVC is still a current research focus [4, 13, 25, 30]. In this paper, we present an optimal entropy-constrained non-uniform scalar quantizer for the low bit-rate pixel domain DVC.

Scalar quantizers have been applied in diverse lossy source coding schemes, such as H.264/AVC, due to their efficiency and simplicity. However, the quantizer design for the practical DVC is just emerging [9, 18, 20, 23, 28]. In this paper, we focus on the optimal non-uniform scalar quantizer design for practical pixel domain DVC. The source signals distribution of pixel domain DVC can be model by a mixture Gaussian with multiple variances, which is different from the conventional hybrid video coding. The uniform quantizer is not appropriate to this model in case of low bit-rate coding. This causes the bit error rate (BER) between the source and SI increases and lowered the coding efficiency. Therefore, we adopt a non-uniform quantizer for the pixel domain DVC. In the proposed quantizer design, the quantization level is fixed first and the conditional probability density function of SI  $Y$  depend on quantization partitions  $Q$ ,  $P(Y|Q)$ , is approximated by a Gaussian mixture model at encoder. Next, the coding rate and distortion model of pixel domain DVC is designed based on  $P(Y|Q)$ . Then, a rate-distortion Lagrangian function is established and the function is optimized by an iterative Lloyd-Max algorithm with a novel quantization partition updating algorithm. Consequently, the entropy-constrained optimal non-uniform quantization partitions are derived and a better RD trade-off is achieved by applying them. From the simulation results, the BER between the source and SI in one quantization partition is effectively decreased and the RD performance for pixel domain DVC is effectively improved.

The remainder of this paper is structured as follows. Section 2 reviews related work in optimal scalar quantizer design and its extension to Wyner-Ziv video coding (WZVC) scenario. Section 3 provides a detailed description to the optimal entropy-constrained quantizer design for low bit rate practical pixel domain DVC. In Section 4, the proposed optimum non-uniform quantizer is applied to the pixel domain DVC and the experiment results are discussed. In Section 5, the conclusions and directions for future work are presented.

## 2 Related work

Scalar quantizers have been studied for decades. The Lloyd-Max scalar quantization [5, 10, 14, 15] is to find a set of quanta  $\{y_1, \dots, y_n\}$  and a set of partition endpoints  $\{x_1, \dots, x_{n-1}\}$  that minimize the mean squared error distortion for a given probability distribution  $p(x)$ . The necessary optimal condition of the Lloyd-Max process shows that each quantum is the centroid of the area of  $p(x)$  between two adjacent endpoints, whereas each endpoint is at the midpoint of two adjacent quanta. Because there are not a closed form equations for the centroid and midpoint, a trial-and-error iterative process was proposed to successively adjust  $y_i$  value until a stationary solution is obtained. The obtained optimal quantizer is not uniform. Since the rate of the Lloyd-Max scalar quantizer is only determined by the quantization level, this kind of quantizer is named fixed-rate quantizer. If the entropy coding is considered in optimal scalar quantizer design, a better rate-distortion tradeoff can be achieved. Thus the fixed-rate quantizer design is generalized to entropy-constrained quantizer design [5]. In [5], Berger provided the rate-quantization ( $R-Q$ ) model and distortion-quantization ( $D-Q$ ) model which are correlated to the quantization parameter  $Q$ . Then, the Lagrangian function was used, which is a linear combination of the rate  $R$  and distortion  $D$  as the optimal function for quantizer design. It is well known that the Lloyd-Max design techniques may produce only locally optimal quantizer. The globally optimal fixed-rate scalar quantizer design has been explored in [8, 26, 27]. Bruce [8] showed that a dynamic programming technique can be used to compute the globally optimal  $K: N$  quantizer in polynomial-time for general error measure. Then, Wu [26] reduced the time complexity of a dynamic programming algorithm for the mean square error measure. Later, Wu and Zhang [27] generalized the previous algorithms to a considerably wider class of error measure. In [16], the dynamic programming strategy is generalized to the optimal entropy constrained scalar quantizer design. They developed a globally optimal quantizer which gives the optimal tradeoff between rate and distortion. And this algorithm is suitable for a family of single-source multiple-receiver and side-information source coding application including the Wyner-Ziv coding scenario. However, it was pointed out that the globally optimal is not guaranteed for Wyner-Ziv quantizer even with respect to the convex codecell constraint [16].

The optimal scalar quantizer design is extended to Wyner-Ziv coding scenario recently. In general, a Wyner-Ziv codec can be considered to consist of a quantizer followed by a Slepian-Wolf encoder (SWC) [21, 29]. The Slepian-Wolf encoder exploits the correlation between the source and SI over a virtual channel. A good Slepian-Wolf codec is capable of approaching the joint entropy between the source and the SI, so it can be regarded as a lossless entropy encoder. The entropy-constrained quantizer design for Wyner-Ziv coding was proposed in [18, 23]. Since the Lloyd-Max algorithm is easy to implement and has low complexity, many authors extended it to generate the optimal entropy-constrained scalar quantizers in Wyner-Ziv coding. In [18], the authors studied the design of quantizer for distributed lossless source coding in terms of the trade-off between distortion and rate. The contribution is that the ideal coding rate equals to the joint conditional entropy of the quantization indices given the SI. Then, an extended Lloyd-Max algorithm which generates the optimal quantizer with disconnected quantization regions was raised for the distributed source coding case. The authors also prove that that uniform quantization is optimal at high rates for distributed source coding. In [23], the authors first designed an optimal quantizer with a fixed number of partitions and the optimal function is a linear combination of the rate and distortion. Then, the necessary conditions were given and they were only correlated with the endpoints of the quantizer. At last, an iterative algorithm was proposed to find the optimal partitions to minimize the optimal equation. The iterative algorithm is an extended

Lloyd-Max type I algorithm, in which each endpoint is bracketed in a convex interval first and then a bisection search is used to locate a local minimum trade-off serving as the new endpoint. This updating procedure guarantees the cost is non-increasing and the optimization function converges to a limit. However, the above optimal quantizer design algorithm is not suitable for practical DVC scenario, since the real signal is not continuous and the joint conditional entropy is difficult to obtain. The reason is that the source and the SI cannot be simultaneously accessed at the encoder or decoder. Therefore, it is necessary to exploit a method to generate optimal quantizer for practical DVC.

Our approach differs from these existing works. First, the conditional distribution of SI given the quantization partition bins  $P(Y|Q)$  is approximated at encoder. The mixture Gaussian model is used to model the conditional distribution  $P(Y|Q)$ . Since the SI  $Y$  can not be accessed at encoder, a coarse estimation of  $P(Y|Q)$  based on the distribution of the source  $X$  is proposed. Second, the optimal RD tradeoff is achieved by using a Lagrangian function  $J(Q)$  which is a linear combination of rate and distortion. Both of the rate and distortion in  $J(Q)$  are modeled based on the conditional distribution  $P(Y|Q)$ . Third, to guarantee the convergence of the optimization function  $J(Q)$  in the Lloyd-Max algorithm, the monotonicity of the interval in which the endpoints of the quantizer lie must be satisfied. Therefore, a quantizer partition updating algorithm which considers the extreme points of the histogram of the input source is proposed. At last, experimental results reveal that the proposed non-uniform scalar quantization can improve the coding performance by 0.5 dB averagely compared to the uniform scalar quantization at low bit rate.

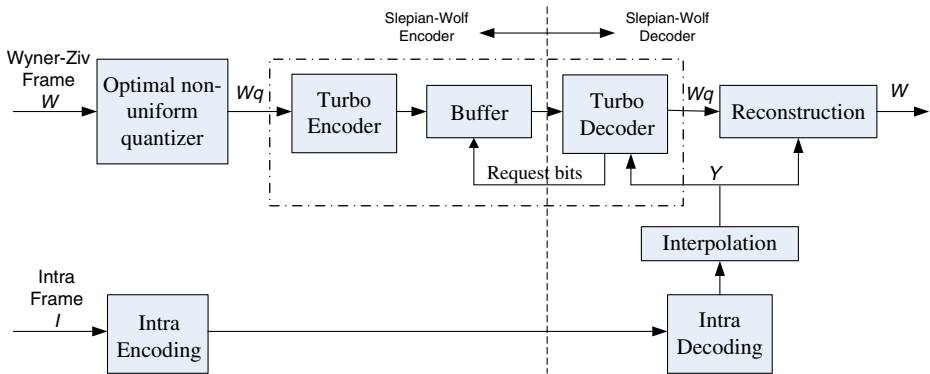
### 3 Optimal non-uniform scalar quantizer design

In pixel domain DVC, in case of low bit-rate, non-uniform scalar quantizer may achieve a better RD trade-off than the uniform scalar quantizer. This is due to the distribution of the source signals and the characteristics of the RD curve of the pixel domain DVC. In this section, we derive the optimal non-uniform quantizers using a modified Lloyd-Max algorithm and apply them to the practical pixel domain DVC. First, the pixel domain DVC scheme with optimal non-uniform quantizer is introduced. Second, the relationship between the rate and distortion in pixel domain DVC is analyzed. Particularly, the capability of reaching a RD trade-off is pointed out. Third, we propose the rate and distortion models which are correlated with quantization bins. At last, a modified Lloyd-Max algorithm is described in detail. It is designed to iteratively derive the optimal quantizer. In this algorithm, a novel quantization partition updating algorithm is adopted.

#### 3.1 Pixel domain DVC scheme

The coding scheme of the employed pixel domain DVC system is similar to the scheme proposed in [1]. The frames of the input video sequence are classified into intra frames ( $I$  frame) and Wyner-Ziv frames ( $W$  frame).  $I$  frames, i.e. odd frames, are coded with the traditional DCT based intra coding method.  $W$  frames, i.e. even frames, are encoded with the turbo codes and decoded using log-MAP algorithm together with SI. The coding scheme is shown in Fig. 1.

At encoder, each pixel of  $W$  frame is quantized using the optimal non-uniform scalar quantizer with  $2^M$  levels. According to the compression rate of turbo codes, several quantizer indexes are organized together to form a symbol. For example, if 4-level quantizer is used, two indexes (each index has 2 bits) form a symbol for the rate 4/5 turbo codes. All these



**Fig. 1** Framework of DVC scheme based on the optimal non-uniform quantizer

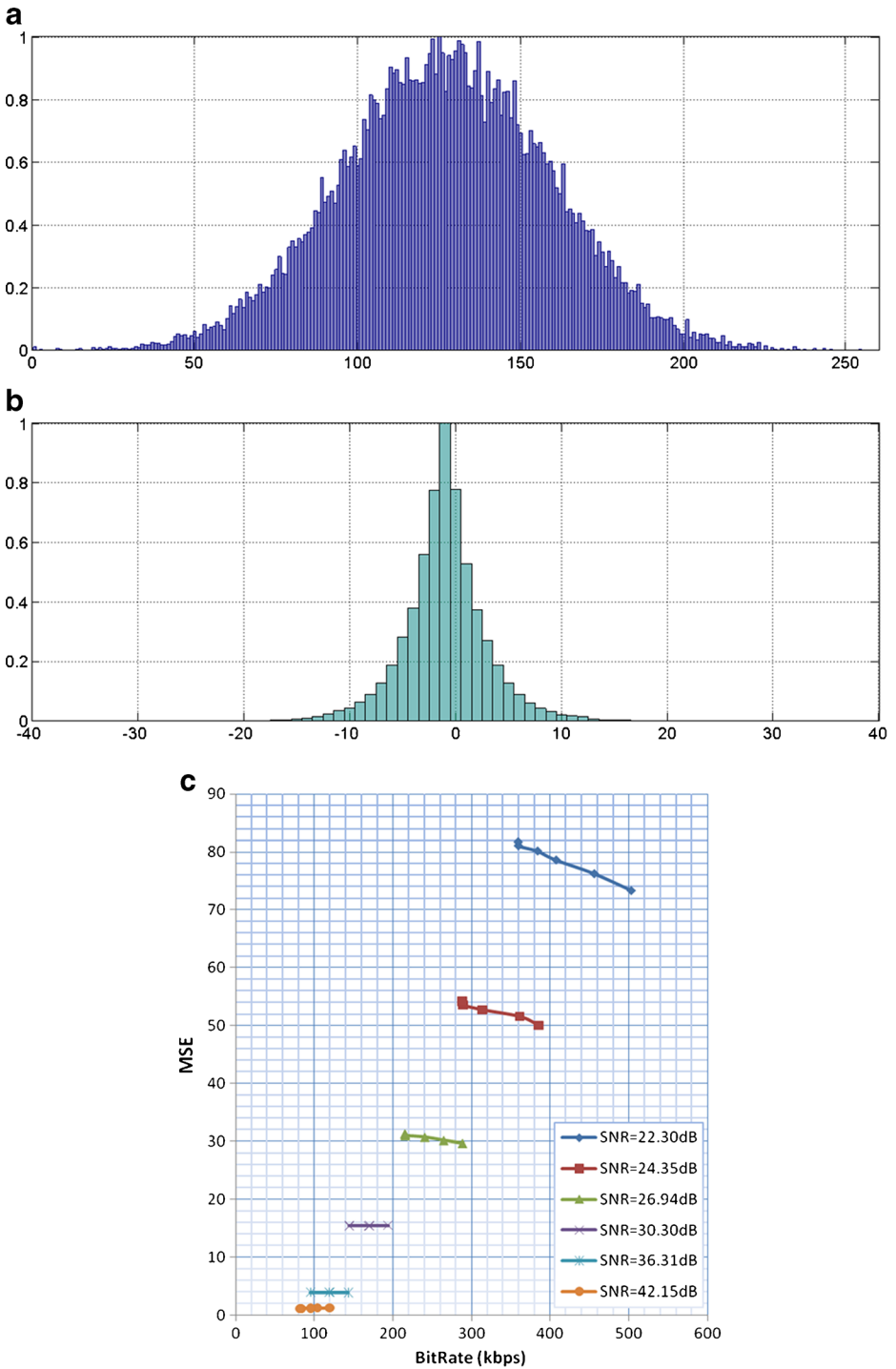
symbols in a frame form a symbol block  $W_q$ . Then, the quantized block  $W_q$  is sent to the rate compatible punctured turbo (RCPT) encoder which consists of two constituent convolutional codes with rate 4/5. After encoding, the system bits of the block are discarded. The parity bits are stored in the buffer and transmitted to decoder upon request successively.

At decoder, the SI  $Y$  is generated by using the motion compensated interpolation (MCI) algorithm [1]. Then, the SI  $Y$  is fed into the turbo decoder module. A Log-MAP algorithm [6] is adopted to decode the quantized symbol block  $W_q$  with the help of the SI  $Y$  until an acceptable bit error rate (BER) is achieved. After turbo decoding, the decoded symbol block  $W_q$  can be used to reconstruct Wyner-Ziv frame  $W$  together with  $Y$ .

### 3.2 Analysis of optimal condition

Quantization can impact on the coding performance of the pixel domain DVC remarkably. In traditional hybrid video coding, quantization is performed on the coefficients of the transformed prediction residue. The distribution of the residual is known to the encoder and it is assumed to be *Laplacian* or *Gaussian* with single variance. So the quantizer accompanied with an entropy encoder can be designed to achieve optimum rate-distortion with the penalty at encoder. While for the pixel domain DVC case, quantization is performed on the original pixel values of the Wyner-Ziv frame. The quantized source signal is encoded individually but decoded depending on the SI at decoder. If the distribution of the Wyner-Ziv frame and the SI in one quantization partition mismatches too much, the BER between the source and the SI will be large and more parity bits will be cost for turbo decoding. This situation may decrease the coding performance. Moreover, the distribution of the pixel values in a frame is a mixture model with multi-variance. And the distribution varies if the content of the frame changes. Hence, there is a higher probability that the error between the source and the SI is bigger if a uniform quantizer is adopted, which is not adaptive to the joint distribution of the source and the SI, especially for low level quantization. If the quantizer is designed according to the joint distribution of the source and the SI with non-uniform quantization partition, the error between the source and the SI in one quantization partition will be small. As a result, the bit rate consumed by Wyner-Ziv decoding can be saved. Accordingly, a better RD trade-off can be achieved. Therefore, how to get the optimal quantizer is one of the most important problems in DVC.

First, we will analyze the relationship between the rate and distortion in the pixel domain DVC. Figure 2 shows the rate and distortion curves when different quantization partitions



**Fig. 2** RD curves of pixel domain DVC by using different quantization partitions (a) Distribution of source  $X$  (b) Distribution of noise: SNR=30.3 dB (c) RD results of DVC by using different quantization partition

are applied to practical pixel domain DVC. In this test, we use the rand functions in Matlab to generate a source  $X$  and a noise signal  $N$ . While the SI  $Y$  is equal to the source  $X$  with the noise  $N$ . Both of the source  $X$  and SI  $Y$  conform to Gaussian distribution as shown in Fig. 2 (a). The noise signal  $N$  complies with the Laplacian distribution as shown in Fig. 2(b). Then, the source  $X$  is encoded by the pixel domain DVC scheme with different quantization partitions bins. Correspondingly  $X$  is reconstructed with the help of the SI  $Y$  at decoder side. The quantizers are comprised of 4-level uniform and non-uniform partition bins. When the SNR equals to 26.94 dB, the applied partition bins are shown in Table 1. For the other five SNR cases, the settings of quantization partition bins are similar. We use SNR to measure the correlation between  $X$  and  $Y$ . In this test, six  $X$ - $Y$  pairs with different SNR are adopted. Accordingly, in Fig. 2(c), each curve corresponds to an  $X$ - $Y$  pair with certain correlation. The SNR is the quotient of the source  $X$  and the noise  $N$  and it is defined as follows,

$$SNR = 10\log_{10} \frac{\sum_{i=1}^M x_i^2}{\sum_{i=1}^M (x_i - y_i)^2} \tag{1}$$

where  $x_i$  is one element of the source and  $y_i$  is corresponding element of the SI.  $M$  is the block size of the generated Gaussian source and it equals to 25344 in the test. In Fig. 2(c), the vertical axis is the mean squared error (MSE) between the original source and the decoded signal. The horizontal axis is the bit rate and the unit is kilobits per second (kbps). The bit rate is calculated as follows,

$$R = \frac{Tb \cdot fr}{fn \cdot 1000} \tag{2}$$

where  $Tb$  is the total bits used in decoding and  $fr$  is the frame rate.  $fn$  is the number of coded frame. Therefore, in Fig. 2(c), each curve show the rate-distortion results of above DVC coding test with different quantizers for a certain  $X$ - $Y$  correlation. From this figure, we can see that the distortion increase is slow compared with the rapid reduce of the rate. Hence, it is possible to design an optimal quantizer to reduce the rate significantly while keeping the distortion increase little, and a better RD trade-off can be achieved if the optimal non-uniform quantizer is applied to such DVC schemes.

Then, the reason of applying non-uniform quantizer which can reduce Wyner-Ziv coding rate is explained here. The Wyner-Ziv encoder is equivalent to a quantizer followed by a Slepian-Wolf encoder [1]. The quantization process is to map the pixel values to a certain quantizer index. Let  $Q$  denotes the quantization partition patterns. The  $q_x$  and  $q_y$  denote the quantizer index of source signals  $x$  and SI  $y$ .  $Y$  denotes the SI and  $X$  denotes the source. That

**Table 1** Coding results of DVC with different quantization partition

SNR(dB)	Quantization partition	Bit rate (kbps)	MSE
26.94	[0,78,128,178,256]	288.84	29.67
	[0,72,129,184,256]	264.84	30.18
	[0,66,128,189,256]	240.6	30.72
	[0,61,128,194,256]	215.4	31
	[0,58,128,199,256]	215.64	31.21
	[0,54,128,201,256]	215.04	31.32

has been shown in [18] that when ideal SWC is assumed, the achievable rate  $R$  is given by  $H(Q|Y)$ . The conditional entropy equals to [11],

$$\begin{aligned}
 H(Q|Y) &= - \sum_{q_x} \sum_{q_y} p(q_x, q_y) \log_2 \frac{p(q_y)}{p(q_x, q_y)} \\
 &= - \sum_{i=1}^{QL} \sum_{j=1}^{QL} p(i, j) \log_2 \frac{p(j)}{p(i, j)}
 \end{aligned}
 \tag{3}$$

The joint probability  $p(q_x, q_y)$  is as follows:

$$p(i, j) = p(q_x, q_y) = \int_{x \in [T_{i-1}, T_i]} \int_{y \in [T_{j-1}, T_j]} p(x, y) dx dy
 \tag{4}$$

in which  $(T_0, T_2, \dots, T_{QL})$  means the endpoints of the quantization partition.  $QL$  is the quantization level. The marginal probability  $p(q_y)$  is

$$p(j) = p(q_y) = \sum_{i=1}^{QL} p(q_x, q_y = j)
 \tag{5}$$

$j$  denotes is the  $j$ th quantization bin and the  $j$ th bin is  $[T_{j-1}, T_j]$ . Let's define a joint distribution function to measure the variation of the joint probability distribution under different quantization partition patterns. From equation (4), it is known that  $p(i, i)$  denote the probability of  $x$  and  $y$  is partitioned into the same quantization bin. Therefore, the joint distribution function adopts the sum of the traces of the joint distribution matrix  $p(q_x, q_y)$  to measure the joint probability. The function is defined as

$$f_{joint} = \sum tr(p(q_x, q_y)) = \sum_{i=1}^{QL} p(i, i)
 \tag{6}$$

To investigate the relationship between the Slepian-Wolf coding rate and the quantization partition pattern, we show the numerical result of ideal SWC coding rate in (3), the joint distribution factor in (6) and the actual BER corresponding to different quantization partition types in Table 2. From Table 2 we can see that the joint distribution of  $p(i, i)$  increases and the BER decreases if the quantization bins with high probability density are enlarged. Moreover, the conditional entropy (3) is proportional to BER. Intuitively, we hope to decrease the conditional entropy, equivalently, to increase the joint distribution of  $p(i, i)$ . Subsequently, both of the BER and the rate for Slepian-Wolf coding are decreased.

From the above analysis, it is known that an optimal condition which considers the trade-off between the rate and distortion has to be achieved at encoder side in the optimal non-uniform quantizer generation algorithm. However, since the source and the SI cannot appear

**Table 2** Comparison of the Slepian-Wolf coding results, SNR=26.94 dB

Quantization partition bins	SWC rate	$f_{joint}$	Actual BER
[0,78,128,178,256]	0.5991	0.8473	0.085
[0,72,129,184,256]	0.0887	0.8574	0.077
[0,66,128,189,256]	0.0369	0.8736	0.07
[0,61,128,194,256]	0.0317	0.8818	0.065



simultaneously at encoder or decoder, the accurately conditional probability distribution of the SI given the source is hard to be calculated. So the rate and distortion of Wyner-Ziv coding cannot be precisely computed. In the next subsection, we will propose a method to estimate the rate and distortion for the practical pixel Wyner-Ziv coding.

### 3.3 Rate distortion model design

To achieve a better rate and distortion trade-off while applying different quantization partitions to the practical pixel domain DVC, a Lagrange function is adopted as an optimal condition in decision-making process. The function is described as follows,

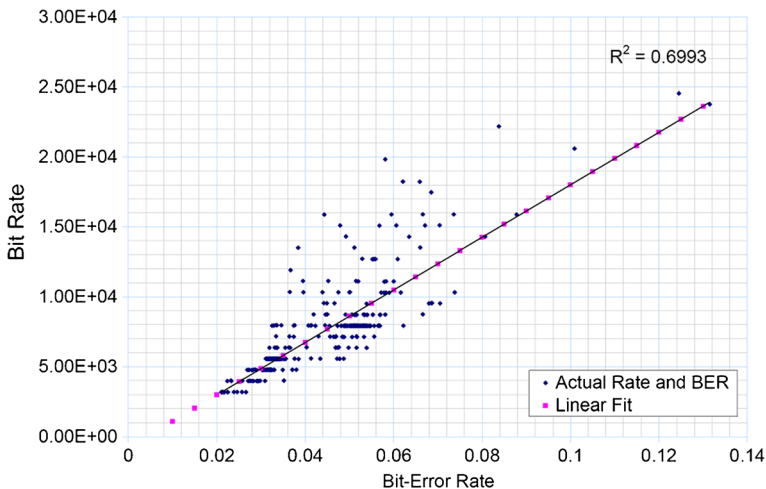
$$J(Q) = \min\{D(Q) + \lambda R(Q)\} \tag{7}$$

In the above function,  $Q$  is a kind of quantization partition pattern.  $R$  represents the rate and  $D$  denotes the distortion;  $\lambda$  is a multiplier of the Lagrange function. In the following, the model of rate and distortion which can be obtained at encoder side is described.

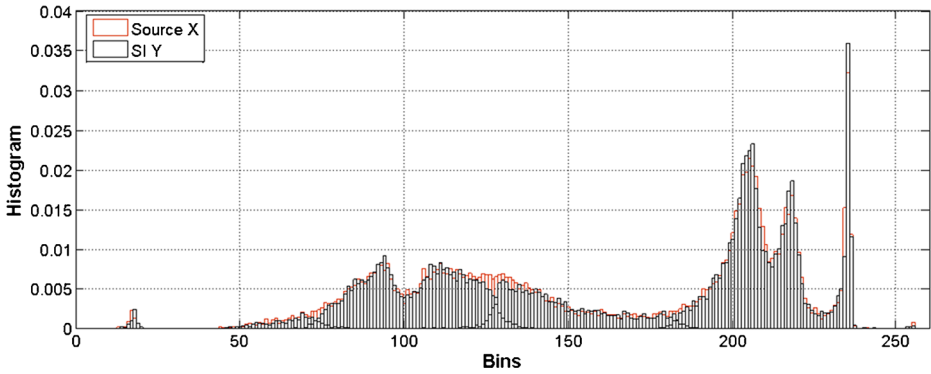
According to the analysis, there is a linear relationship between the actual coding rate and actual bit-error rate. The relationship is shown in Fig. 3. In Fig. 3, three test sequences (Foreman\Akiyo\Mother&Daughter@QCIF) are used and the average quality of key frames is 35.57 dB. Therefore, we propose a method to estimate the conditional probability at encoder. The model of rate and distortion are both based on the estimated conditional probability. Let  $P(Y|Q)$  denote the conditional probability. Since  $Y$  can not be obtained at encoder, we use the distribution of  $X$  to estimate the distribution of  $Y$ . This estimation is reasonable because these two distributions are very similar in general. This is shown in Fig. 4. The histograms of  $Y$  and  $X$  which are the first Wyner-Ziv frames of Forman@QCIF sequence are compared. We use a Gaussian mixture model to estimate the conditional probability density function  $P(Y|Q)$ . It is defined as follows:

$$P(Y|Q) = \sum_k \pi_k N(x_k | \mu_k, \sigma_k) \tag{8}$$

$k = 1, \dots, QL$



**Fig. 3** Linear relationship between the coding rate and the BER



**Fig. 4** Similarity between the histogram of the source  $X$  and that of the SI  $Y$ , the 1st frame of Foreman@QCIF

where  $QL$  is the quantization level.  $P(Y|Q)$  is a linear combination of  $QL$  Gaussian components.  $N(x_k|\mu_k, \sigma_k)$  is the Gaussian distribution model for  $x_k$  in the  $k$ th quantization partition bins.  $\mu_k$  is the mean and  $\sigma_k$  is the variance. The parameter  $\pi_k$  is the *mixing coefficients*. We use the following method to estimate the mixing coefficient.

$$\pi_k = \frac{N_k}{N} \tag{9}$$

where  $N$  is the total pixels number of the source  $x$ .  $N_k$  is the number of the source  $x$  which is contained in the  $k$ th quantization partition bin. The mixing coefficient  $\pi_k$  meets the normalization condition.

The method of moment estimation is used to estimate the probability density functions of each Gaussian component in the Gaussian mixture model (8). The estimation function for mean is

$$\mu_k = \frac{1}{N_k} \sum_{i=1}^{N_k} x_k(i) \tag{10}$$

$$x_k(i) \in [T_{k-1}, T_k) k = 1, \dots, QL$$

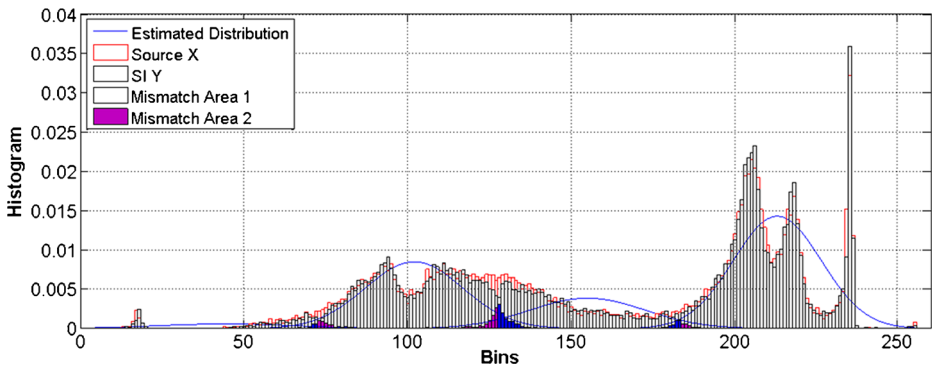
The estimation function for variance is

$$\sigma_k = \frac{1}{N_k} \sum_{i=1}^{N_k} (x_k(i) - \mu_k)(x_k(i) - \mu_k) \tag{11}$$

$$x_k(i) \in [T_{k-1}, T_k) k = 1, \dots, QL$$

Figure 5 shows the pixel domain histograms of  $X$  and  $Y$  which are the 1st frame of Wyner-Ziv frame in test sequence Foreman@QCIF. The curve is the estimated conditional probability density function  $P(Y|Q_k)$ . The area characterized by different colors represents the real mismatch between  $X$  and  $Y$ . It is denoted that  $y_k$  falls outside certain quantization partition bin  $[T_k, T_{k+i})$  while the corresponding  $x_k$  resides in it.  $x_k$  and  $y_k$  are pixels having the same position  $k$  in the original frame and the corresponding SI. The ratio between the color area and the total area of the histogram is the actual BER.

Due to the SI  $Y$  cannot be accessed at encoder. The actual BER cannot be obtained. This means that the color area cannot be achieved. Thus, the area bounded by the curve of  $P \times (Y|Q_k)$ , which falls outside the boundary of quantization bins, and the axis of  $Bins$  is used to



**Fig. 5** Estimated probability distribution of  $P(Y|Q_k)$ , the 1st frame of Foreman@QCIF

fit the BER. Let’s denote the specified area as  $S$ . The linear relationship also exists in the actual bit-error rate and the conditional probability of  $Y$  given the quantization partition pattern  $Q$ . The relationship is shown in Fig. 6. The test sequence is Foreman@QCIF and the average quality of the key frame in the DVC coding is 32 dB. Since the actual coding rate and BER are in a linear relationship, the area  $S$  can be used to estimate the rate.

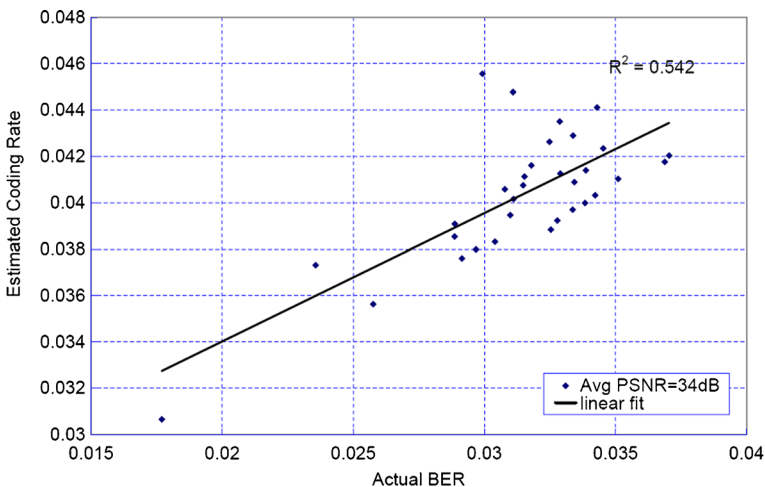
$$R = \sum_k S_k = \sum_k \sum_{i \in \{i|y_k(i) \leq T_{k-1} \text{ or } T_k \leq y_k(i)\}} P(y_k(i)|Q_k) \quad k \in [1, QL] \quad i \in [0, 255] \quad (12)$$

in which  $y_k(i)$  denotes the  $i$ th bin of curve  $P(Y|Q_k)$  in the  $k$ th quantization partition. The bin width is 1. The range of  $y_k(i)$  is  $[0, 255]$  and it is determined by the pixel values.  $P(y_k(i)|Q_k)$  is the probability of  $P(Y|Q_k)$  at  $y_k(i)$  and it can be calculated by equation (8).

Similarly, the conditional probability can also be used to estimate quantization distortion.

$$D = \sum_k \sum_i P(y_k(i)|Q_k) \cdot |Q^{-1}(y_k(i)) - \mu| \quad (13)$$

where  $i \in \{i|P(y_k(i)|Q_k) \geq THR\}$ . It means that the  $i$ th bin in the  $k$ th quantizer partition is used to calculate the distortion if the corresponding probability  $P(y_k(i)|Q_k)$  is larger than a



**Fig. 6** Linear relationship between the estimated coding rate and the actual BER

threshold value.  $Q^{-1}(y_k(i))$  means the inverse quantization value of bin  $y_k(i)$ . The mean value  $\mu$  is defined as follows:

$$\mu = \begin{cases} \frac{1}{M_1} \sum_{x \in [T_{k-1}, \frac{T_{k-1}+T_k}{2})} x & \text{if } y_k(i) \leq \frac{T_{k-1}+T_k}{2} \\ \frac{1}{M_2} \sum_{x \in [\frac{T_{k-1}+T_k}{2}, T_k)} x & \text{if } y_k(i) \geq \frac{T_{k-1}+T_k}{2} \end{cases} \tag{14}$$

where  $M_1 = \text{card}(\{x|x \in [T_{k-1}, \frac{T_{k-1}+T_k}{2})\})$  and  $M_2 = \text{card}(\{x|x \in [\frac{T_{k-1}+T_k}{2}, T_k)\})$ .  $M_1$  and  $M_2$  are the numbers of source elements in the first and the second half-part of each quantization partition respectively. So far, the rate and distortion models have been established.

The aim of Lagrange cost function is to minimize the function (7) and the  $\lambda$  is the so called Lagrange multiplier. Consequently, how to determine  $\lambda$  become a key problem. In our proposed method, both of the distortion  $D$  and rate  $R$  are related to quantization partition pattern  $Q$ , Supposing  $R$  and  $D$  to be differentiable everywhere, the minimum of function (7) is given by setting its derivative to zero, i.e.

$$\frac{d\{J(Q)\}}{dQ} = \frac{d\{D(Q)\}}{dQ} + \lambda \frac{d\{R(Q)\}}{dQ} = 0 \tag{15}$$

leading to,

$$\lambda = - \frac{d\{D(Q)\}}{dQ} / \frac{d\{R(Q)\}}{dQ} = - \frac{dD}{dR} \tag{16}$$

Therefore, the above equation derives the Lagrange multiplier and it indicates that  $\lambda$  corresponding to the negative slope of the rate-distortion curve. In our scheme, we use a quantization partition updating algorithms to iteratively derive the optimal quantization partition bins. In each updating cycle, the corresponding partition bin is extended with equal steps. As the quantization partition bins changes, it produces a new pair of rate  $R$  and distortion  $D$ . We use an offline training method to obtain a sequence level  $\lambda$ . In this method, a sequence level rate-distortion curve can be calculated using the RD model (12) and (13) according to the quantization partition updating cycle. Consequently, the negative slope of the sequence level rate-distortion curve is the Lagrange multiplier  $\lambda$ .

### 3.4 Modified Lloyd-Max quantizer design algorithm

In this section, a modified Lloyd-Max algorithm is described in detail. It is designed to iteratively derive the optimal quantizer. In the modified Lloyd-Max algorithm, a novel quantization partition updating algorithm is proposed. These quantization partition bins are updated according to the monotonicity of histogram. The modified Lloyd-Max algorithm in the encoder is described as following:

1. Analyze the histogram of the pixels which are drawn from the Wyner-Ziv frame;
2. Choose the uniform quantization partition bins as the initial quantization partition for a fixed quantization level;
3. Set  $k=1, J^0=\text{MAX\_VAL}$ ;
4. Compute the new optimal reproduction values for these partitions using the centroid condition; then, using the nearest neighbor condition to gather the source elements to different  $QL$  partitions; consequently, new Voronoi quantization partitions are obtained.

5. Calculate the rate and distortion for the particular quantization partition, using equation (10)~(12). Compute the cost function  $J^{(k)} = D + \lambda R$  and compare it with the previous step cost  $J^{(k-1)}$ . If  $J^k < J^{(k-1)}$ , continue to next step; Else, stop and exit the iteration;
6. Perform the quantization partition updating algorithms. It chooses new initial reproductions set which expand the cell that contains those inputs with high probability. Increase iterator,  $k++$  and go back to 4.

The 4th step is to use the Lloyd algorithm to adjust the new quantization partition and calculate the new reproduction value for these new partition bins. First, the Lloyd algorithm uses the nearest neighbor rules to compute the new quantization partition bins.

$$V_i = \{x | d(x, \hat{x}_i) \leq d(x, \hat{x}_k); \text{ all } k \neq i\} \quad k, i \in [1, QL] \quad (17)$$

where  $\hat{x}_i$  is the reproduction value for the  $i$ th quantization partition bin. The initial reproduction value  $\hat{x}_i$  is the midpoint of the quantization partition bin.  $d$  denotes the distance from sample points  $x$  to the neighbor  $\hat{x}_i$  and the Euclidean distance is taken as the distance metric. In our scheme, the sample point is the pixel value and the neighbor is the reproduction value.  $V_i$  is the rearranged  $i$ th quantization partition bins.

$$\hat{x}_i = (T_{i-1} + T_i) / 2 \quad i \in [1, QL] \quad (18)$$

Second, the new reproduction values are computed for the new quantization partition  $V_i$  using the centroid condition rule.

$$\hat{x}_i = \frac{\sum r_i x}{\sum r_i}, \quad (19)$$

in which,

$$r_i = \begin{cases} 1 & \text{if } x \in V_i \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

The mean value in (19) is the new reproduction value for  $V_i$ . The aim of rearranging the quantization partition is that the produced  $V_i$  is beneficial to the Turbo decoding procedure. According to the nearest neighbor condition, it is known that the distance from those pixels in one particular partition to the corresponding reproduction value is shorter than that to other reproduction values. This circumstance improves the joint distribution between the original source and SI [11]. Hence, the bit error rate between the original and the SI is decreased. Consequently, the bit-rate consumed by recovering the error is decreased.

### 3.5 Quantization partition updating algorithm

According to the analysis in section 3.2, if the quantization bins with high probability density are enlarged, the joint distribution of source  $X$  and SI  $Y$  increases. Therefore, the partition bin with the highest probability is extended first. To guarantee the convergence of the optimization function  $J(Q)$ , the boundary of current quantization bin needs to be extended toward the decreasing direction of the monotonic interval. The detailed steps of quantization partition updating algorithm are described as follows:

1. Count the probability distribution for each quantization partition. Firstly, integrate the probability density of each quantization partition bin and the probability for each bin is obtained. Secondly, sort the partition bin according to the probability.

2. Extend the quantization bins based on the sorting order.
3. Determine the monotonic interval within which the boundary of current quantization partition bin is estimated to lie. Afterwards the boundary is expanded toward the direction of monotonic decreasing. When the boundary reaches the minimal value of current monotonic interval, stop the expansion of the current interval and go to Step 2 to deal with the next interval extension. Or the extending boundary coincides with the other boundary of the neighboring quantization partition bin. Merge two quantization partitions and generate a new quantization range. Then, go to Step 2 to deal with the next interval extension.

Above all, we will introduce the first step. To get the probability for quantization partition, the histogram of the pixel intensity should be counted up. Let’s define the histogram of those pixels in one frame as,

$$h(k) = \sum_{x,y} I\{f(x,y)\} = k \quad k = 0, 1, \dots, 255 \tag{21}$$

in which  $I\{f(x,y)\}$  denotes the pixel intensity at the position  $(x,y)$  in one frame  $f$ .  $k$  is the  $k$ th intensity level and its value ranges from 0 to 255. Hence, the bin width of the histogram is 1.  $h(k)$  is the histogram and it indicates the total number of the pixels with intensity level  $k$ . The histogram  $h(k)$  implies the probability density. Then, the histograms are counted up to calculate the probability of each quantization partition bin,

$$P(Q_S) = \sum_{k \in S} h(k) \quad S = [1, QL] \tag{22}$$

$Q_s$  represents the quantization partition  $S$ .  $P(Q_s)$  is the probability of quantization partition  $S$ .

Then, the quantization partition bins are sorted according to  $P(Q_s)$ . Next, the quantization bins are extended based on the sorting order. Subsequently, the determination of monotonic intervals is introduced as follows. To determine the monotonic intervals, the extreme points of the histogram should be derived first. The extreme points are defined as:

$$\begin{aligned} h_{\max}(n) &= \{n|h(n)h(n-1) \& h(n)h(n+1)\} \\ h_{\min}(n) &= \{n|h(n)h(n-1) \& h(n)h(n+1)\}, \end{aligned} \tag{23}$$

where the maximum value of histogram  $h$  is  $h_{\max}(n)$  and the minimum value is  $h_{\min}(n)$ .

Since the original frame  $f(x,y)$  contains high frequency information, there are a lot of noise peaks in the histogram. If the histogram  $h(k)$  is directly used to derive the extreme points, many pseudo extreme values may be obtained, as shown in Fig. 7. The red triangle is the derived extreme points. Therefore, the histogram needs to be smoothed. Considering the computational complexity, the moving window weighed average is used to smooth the histogram. The weighing function is defined as follows:

$$h_s(k) = \frac{1}{N} \sum_{m \in [k-\frac{N}{2}, k+\frac{N}{2}]} h(m)w(m) \tag{24}$$

$h_s$  is the smoothed histogram, and the moving window function  $w(m)$  is defined as:

$$w(m) = \begin{cases} 1 & k - \frac{N}{2} \leq m \leq k + \frac{N}{2} \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

After smoothing, the distribution of the extreme points is shown in Fig. 8. The blue curve is the smoothed histogram. Compared with Fig. 7, the pseudo extreme points are significantly decreased. However, there also exist some pseudo extreme points which are not

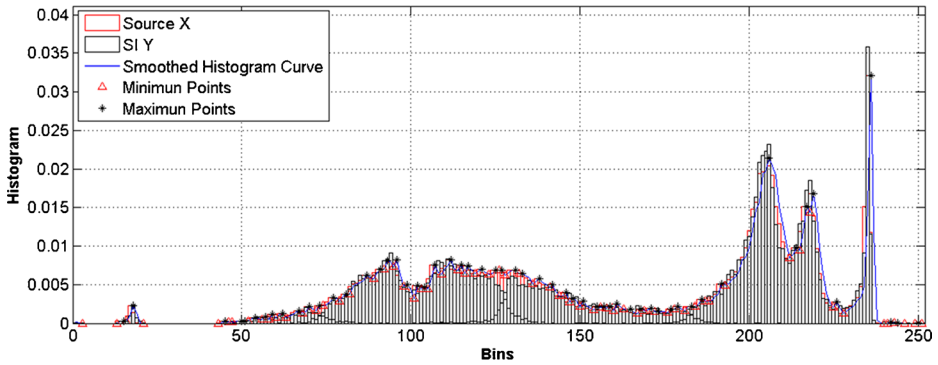


Fig. 7 Extreme points of histogram  $h$

located at the peak or valley of the histogram wave. Therefore, the pseudo extreme points need to be further removed. From Fig. 8, it can be seen that the gradient of extreme points gathered in the peak and valley is smooth, whereas the gradient of other groups of extreme points descends steeply. So a threshold of the slope over a cluster of extreme points used to remove those points which are not located at the peak or valley of the histogram. First, a clustering method resembling  $k$ -means algorithm is used to gather the extreme points into different sets. Then, the maximal slopes over different sets are calculated. The values of the slopes are compared with the threshold to judge whether the points in each set are real extreme points.

$K$ -means clustering is a method which partition  $n$  observations into  $k$  clusters are based on a specific criterion. The time complexity of  $k$ -means is  $O(tkn)$ , in which  $n$  is the number of the samples.  $k$  is the number of the cluster sets and  $t$  is the iterative times.  $k$  and  $t$  are often much smaller than  $n$ . Therefore, it is an efficient algorithm. The weakness of  $k$ -means is that the number of cluster sets  $k$  should be predefined. However, determining the correct number of clusters for an unknown sample set is difficult. In our method, the number of clusters  $k$  does not need to be determined first, but to determine a radius  $R$ . In our scheme, an empirical radius  $R$  is taken. The radius  $R$  is obtained by an offline training method and it is consistent at the sequence level. For any point  $p$  in the sample space, the distance from  $p$  to the already known centers

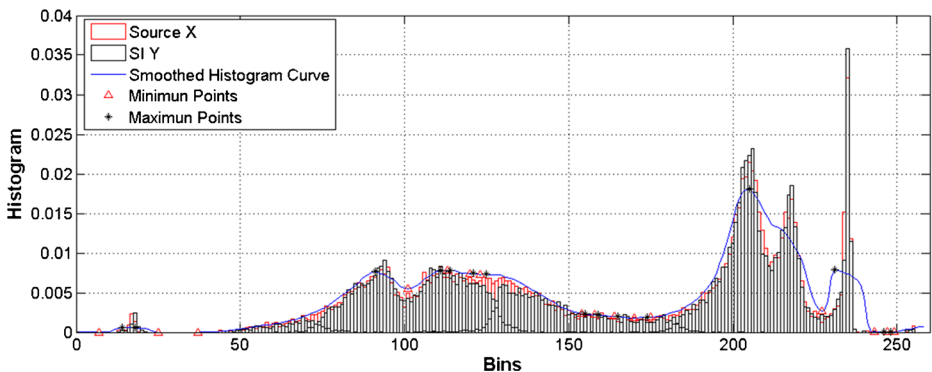


Fig. 8 Extreme points of smoothed histogram

is measured. If the distance to center  $k$  is smaller than  $R$ , then  $p$  is added into cluster set  $k$ . Otherwise if the distances from  $p$  to any center are all larger than  $R$ ,  $p$  is performed as a central point for a new cluster set  $k+1$ . The number of cluster set is increased by 1. The two dimensional information that includes the pixel value and the histogram of the extreme points  $p = (k, h_s(k))$  are adopted as a point in the sample space. The distance  $r$  from  $p$  to any already known centers in our cluster method is calculated as follows,

$$r = \left[ (k - k_c)^2 + (h_s(k) - h_s(k_c))^2 \right]^{\frac{1}{2}} \tag{26}$$

where  $(k_c, h_s(k_c))$  is the center of the clustering set. The generated clustering sets are shown as Fig. 9.

The slope of the cluster set is determined by the maximal histogram  $h_s^{\max}$  and the minimal histogram  $h_s^{\min}$  in one cluster set. The function is defined as:

$$l = \left| h_s^{\max}(k_1) - h_s^{\min}(k_2) / k_1 - k_2 \right| \tag{27}$$

The concrete steps of the cluster algorithm are described as followed:

1. Set the cluster set as  $CSet = \emptyset$  and the set number  $k=0$
2. While  $(i < PointSet.size)$  {
  - a) Select the sample point  $p_i$  from the sample space  $PointSet$ ;
  - b) Calculate the radius  $r$  from  $p_i$  to all already known centers  $p = (k, h_s(k))$  of existing cluster sets.
    - If there exists a cluster set  $k$  that satisfies  $r_k < R$ , the sample point  $p_i$  is added into the  $k$ th cluster set  $CSet(k)$ . Update the center of the  $k$ th cluster set  $CSet(k)$ ;
    - Else if all radius  $r$  is larger than  $R$ , then  $p_i$  is performed as a center for a new cluster set.  $k++$ .
  - c)  $i++$ ; }

After performing the proposed clustering method, the pseudo extreme points which are not located at the peak or valley of the histogram are removed. The maximum and minimum values of the remaining extreme points are determined. The interval between two adjacent maximum and minimum values is a monotonic one.

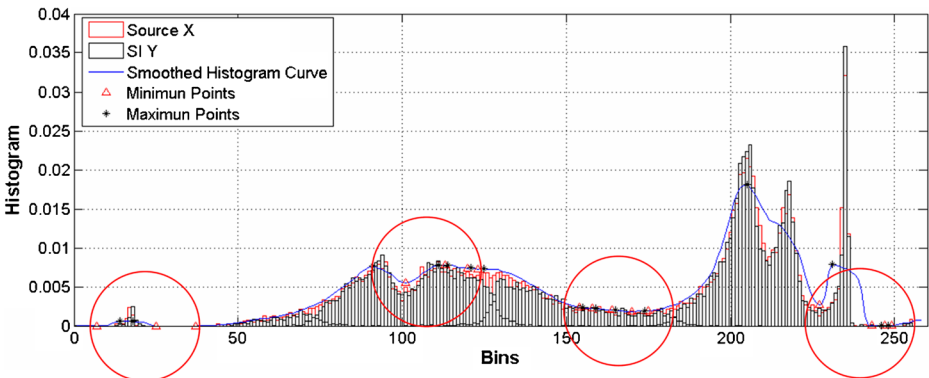


Fig. 9 Clusters of extreme points using K-MEANS clustering algorithm



At last, the interval extension is introduced. First, the monotonic interval in which the boundary  $T_i$  of the current quantization bin  $[T_i, T_{i+1})$  lies is determined. Then,  $T_i$  extends toward the direction of monotonic decreasing, which is defined as follows:

$$T_i' = \begin{cases} T_i + \text{Step} & \text{if } T_i \in \text{MD} \\ T_i - \text{Step} & \text{if } T_i \in \text{MI} \end{cases} \quad (28)$$

where MD represents a monotonically decreasing interval and MI represents a monotonically increasing interval. When  $T_i$  reaches the minimal value of current monotonic interval, stop the extension of  $T_i$  and begin to extend  $T_{i+1}$ . Until both the boundaries of  $[T_i, T_{i+1})$  reach the minimal values, the current quantization bin  $[T_i, T_{i+1})$  stops extending and the extension of next bin begins. Moreover when the boundary  $T_i$  coincides with that of adjacent quantization bin  $T_{i-1}$ , two adjacent bins are merged together. After that the quantization level is decreased by 1 and a new quantization partition pattern is produced. To keep the quantization level unchanged, a new quantization bin should be generated. First, the new quantization partition is sorted according to the probabilities of each quantization bins. Second, the quantization bin  $[T_m, T_{m+1})$  with the minimal probability is chosen for the generation of new quantization bin. The quantization  $[T_m, T_{m+1})$  bin is divided into two partition bins  $[T_m, (T_m + T_{m+1})/2)$  and  $[(T_m + T_{m+1})/2, T_{m+1})$ .

## 4 Experimental results

### 4.1 Evaluation of RD performance using PSNR

In this section, the proposed optimal non-uniform quantizer is implemented on the pixel domain DVC to verify the coding efficiency. The key frames are H.263 intra coded. The test sequences are Foreman@QCIF, daughter@QCIF and Akiyo@QCIF. The parameter of the test sequences are given in Table 3. The last column of Table 3 gives the average quality of key frames which is used in the Wyner-Ziv coding. A rate-compatible punctured turbo encoder (RCPT) is adopted in Slepian-Wolf codec and the acceptable bit error rate at the decoder is set to  $10^{-3}$ . The parameter of Laplacian distribution model is obtained by offline fitting the difference between the original frame and its SI frame. The test results are shown as Fig. 10. In Fig. 10, “Proposed Quantization” denotes the RD curve produced by the DVC scheme which adopts the proposed optimal non-uniform quantizer in this paper. “Uniform Quantization” denotes the RD performance of the DVC scheme which adopts the uniform quantizer. “H.263-I” indicates the RD results of H.263 intra frame coding. “H.263-IBI” indicates the RD results of I-B-I coding of H.263. The GOP size is 2 and the coding structure is I-B-I-B. The fixed quantization 4-level is adopted by the uniform quantization in pixel domain DVC. The range of pixel values is partitioned into four quantization intervals. The level of non-uniform quantization is also four, but the lengths of intervals are unequal. In each sequence, 150 frames are encoded and the coding structure is I-W-...-W-I. The GOP size of Wyner-Ziv coding is 2.

From Fig. 10, we can see that our proposed optimal non-uniform quantizer outperforms the uniform quantizer in practical pixel domain WZVC scheme. The average gain of the Foreman sequence is 0.55 dB. The average gain of the Mother and Daughter sequence is 0.42 dB. The average gain of the Akiyo sequence is 0.46 dB. The average gain of the Silent sequence is 0.47 dB at low bit-rate end. The average gain of the Mobile sequence is 0.12 dB at low bit-rate end. The gain of Mobile sequence is not as significant as other test sequences.

**Table 3** Parameters of the test sequences

Test sequence	Resolution	Luminance/ Chroma	Frame number	Average quality (PSNR)
Foreman	176×144	4:0:0	150	35.84 dB
				34.02 dB
				31.60 dB
				30.71 dB
				29.98 dB
Mother and Daughter	176×144	4:0:0	150	29.35 dB
				44.8 dB
				40.39 dB
				37.71 dB
				35.97 dB
Akiyo	176×144	4:0:0	150	45.58 dB
				41.35 dB
				38.67 dB
				36.73 dB
Silent	176×144	4:0:0	150	34.3 dB
				32.94 dB
				31.85 dB
				30.97 dB
Mobile	176×144	4:0:0	150	30.27 dB
				31.89 dB
				30.06 dB
				28.64 dB
				27.5 dB
				26.56 dB
				25.76 dB

It is because the pixel domain distribution of the Mobile is not so consistent with the mixture Gaussian model.

#### 4.2 Evaluation of image quality using SSIM

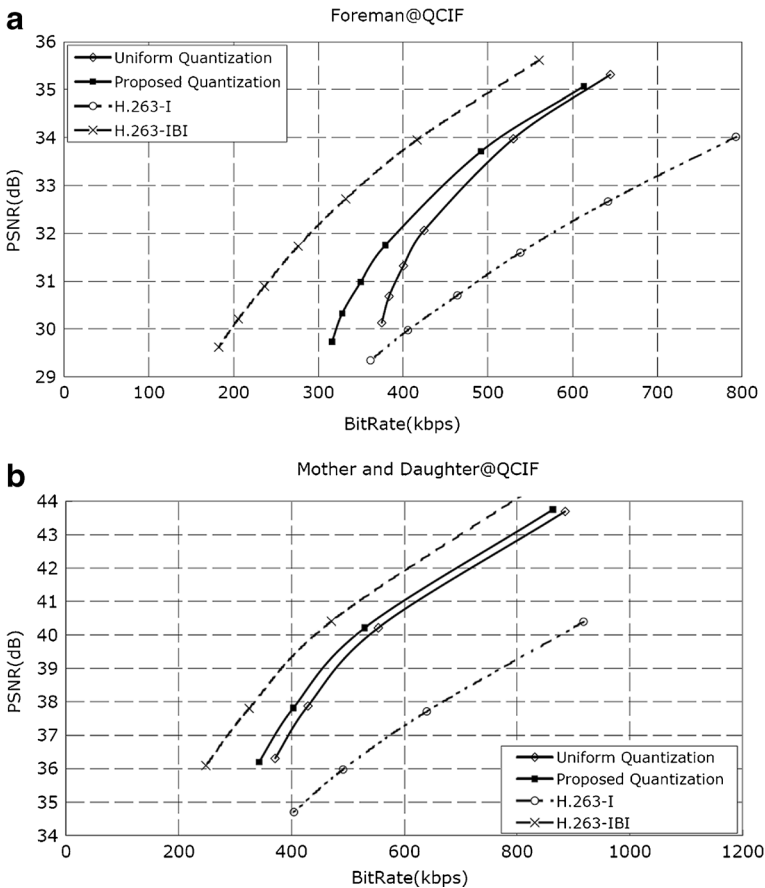
In this section, we use the Structural Similarity Index metric (SSIM) to evaluate the image quality of the decoded video sequence. The SSIM quality assessment is based on the degradation of structural information. It is under the assumption that human visual perception is highly adapted for extracting structural information from a scene. The SSIM has been verified that it is sensitive to the structural information [24]. And it has been applied in many fields to assess the video quality from the perspectives of human visual system, such as the perceptual-based rate-distortion optimization (RDO) [17].

The equation of SSIM is defined as (13) in [24]. In our experiments, the block size of SSIM measurement is typically  $8 \times 8$ , and the final SSIM value is the averaged SSIM of all blocks. The maximum SSIM index is 1, which occurs when the two images are identical. The more similar two images are, the closer SSIM index is to 1. Table 4 shows the average

SSIM index, and average PSNR of various sequences. Compared with the uniform quantizer, the SSIM gain of the proposed optimal non-uniform quantizer is from  $-0.016$  to  $0.001$ , and PSNR gain is between  $-0.44$  and  $0.08$  dB. For most sequences, both of the SSIM and PSNR are loss. However, comparing with the rate savings, the PSNR decrease is not significant. This result is consistent with the discussion in section 3.2. Since the distortion increases slow compared with the rapid bit-rate reductions, if the optimal non-uniform quantizers are applied, a little PSNR can be sacrificed to save more coding bits to achieve a better R-D trade off for Wyner-Ziv coding. For the SSIM index, more than half of the cases (10 of 16) the loss is smaller than 0.01. Therefore, the sacrificed objective quality brings little harm to the subjective quality.

### 4.3 Computational overhead

The computational overhead of the proposed scheme mainly comes from the modified Lloyd-Max quantizer design algorithm, which iteratively derive the optimal quantizer



**Fig. 10** RD results of DVC using optimal non-uniform scalar quantizer (a) Foreman@QCIF (b) Mother and Daughter@QCIF (c) Akiyo@QCIF (d) Silent@QCIF (e) Mobile@QCIF

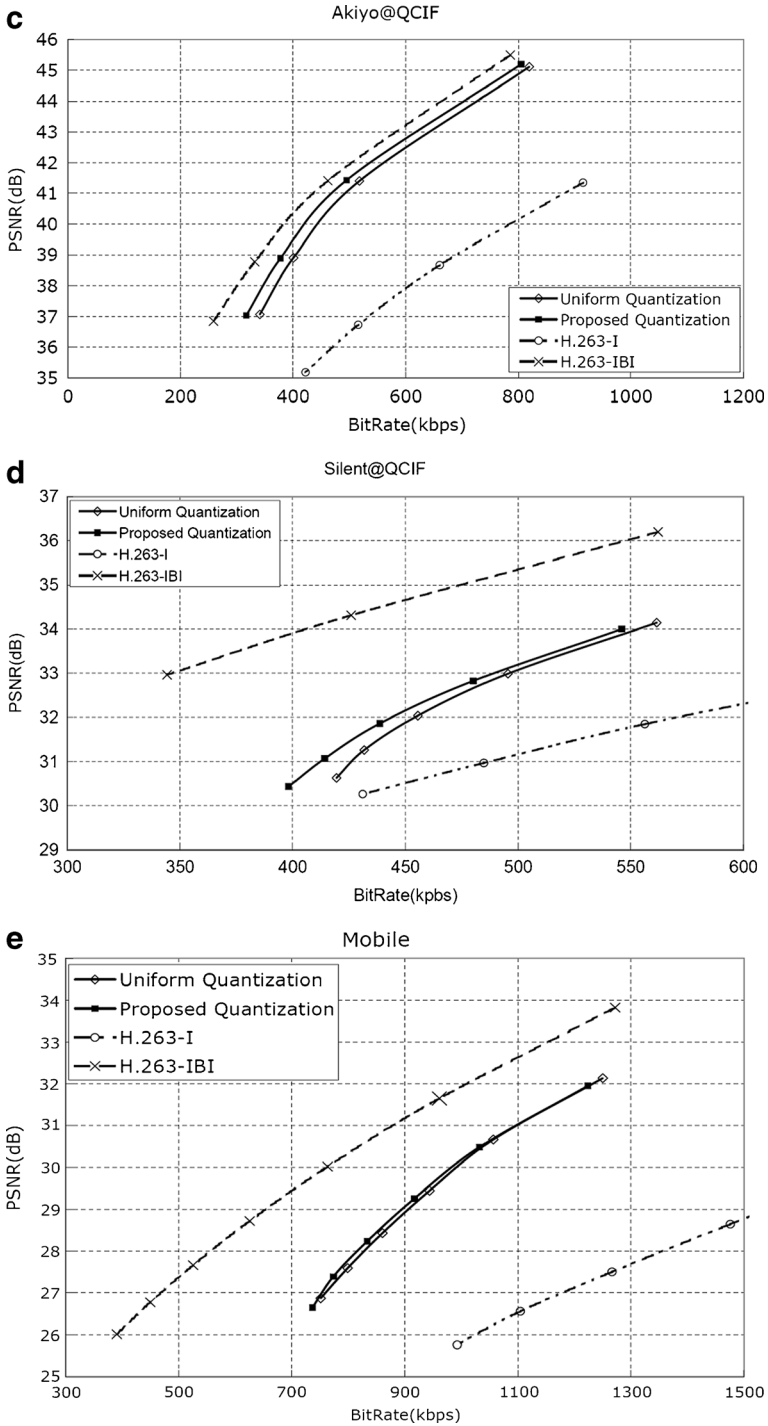


Fig. 10 (continued)

**Table 4** Comparisons of SSIM and PSNR obtained for Wyner-Ziv coding with uniform quantizer and the proposed optimal non-uniform quantizer at different quality levels of key frames

Sequence	Average PSNR of key frames (dB)	Uniform quantizer		Proposed optimal non-uniform quantizer	
		SSIM	PSNR (dB)	SSIM	PSNR (dB)
Foreman	35.84	0.941179	35.31317	0.935474 (-0.005705)	35.07112 (-0.24)
	34.02	0.930352	33.97601	0.922936 (-0.007416)	33.71171 (-0.26)
	31.60	0.907405	32.05752	0.896616 (-0.010789)	31.74862 (-0.31)
	30.71	0.896536	31.31767	0.883815 (-0.012721)	30.98456 (-0.33)
	29.98	0.885639	30.68598	0.871051 (-0.014588)	30.32642 (-0.36)
	29.35	0.875731	30.13102	0.858948 (-0.016783)	29.73194 (-0.39)
Mother and Daughter	44.8	0.986778	43.69433	0.987188 (0.00041)	43.74867 (0.05)
	40.39	0.973329	40.21276	0.973141 (-0.000188)	40.20593 (-0.007)
	37.71	0.959655	37.87372	0.958332 (-0.001323)	37.81196 (-0.06)
	35.97	0.95011	36.30006	0.946973 (-0.003137)	36.1952 (-0.1)
Akiyo	45.58	0.990301	45.11039	0.990958 (0.000657)	45.19135 (0.08)
	41.35	0.980876	41.40303	0.981132 (0.000256)	41.42901 (0.03)
	38.67	0.969836	38.9031	0.969821 (-0.000015)	38.89532 (-0.008)
	36.73	0.956825	37.06971	0.955745 (-0.00108)	37.02954 (-0.04)
Silent	34.3	0.935772	33.9837	0.931597 (-0.004175)	33.70077 (-0.28)
	32.94	0.920514	33.04094	0.91393 (-0.006584)	32.71387 (-0.33)
	31.85	0.904564	32.22912	0.896268 (-0.008296)	31.88112 (-0.35)
	30.97	0.887804	31.55599	0.876676 (-0.011128)	31.16755 (-0.39)
	30.27	0.873632	30.99687	0.859653 (-0.013979)	30.60383 (-0.39)
Mobile	31.89	0.966267	32.38275	0.96368 (-0.0026)	32.01314 (-0.37)
	30.06	0.955692	31.26217	0.952945 (-0.0027)	30.90456 (-0.36)
	28.64	0.944253	30.24766	0.940841 (-0.0034)	29.8754 (-0.37)
	27.5	0.931721	29.3644	0.927363 (-0.0044)	28.97355 (-0.39)
	26.56	0.920065	28.6263	0.914652 (-0.0054)	28.21834 (-0.41)
	25.76	0.908376	27.97749	0.901004 (-0.0074)	27.5329 (-0.44)

partitions for Wyner-Ziv coding according to the estimated rate distortion model. Table 5 shows the encoding time comparison between the uniform and the proposed optimal non-uniform quantizer based WZVC. However, comparing with the state-of-art hybrid H.264/AVC encoder, the encoding time of the optimal non-uniform quantizer based WZVC keeping at millisecond level is reasonable. The motion estimation time of the encoder of the JM reference software exceeds 100 s [12, 22].

## 5 Conclusion

This paper presents a novel optimal non-uniform quantization design method for the pixel domain DVC scheme. A better RD trade-off can be achieved in pixel domain DVC by applying non-uniform quantizer. The non-uniform quantizer considers that the joint distribution between

**Table 5** Encoding time comparison

Sequence	Average PSNR of key/intra frames (dB)	WZVC with uniform quantizer (Millisec)	WZVC with optimal non-uniform quantizer (Millisec)
Foreman	35.84	5	371
	34.02	6	374
	31.60	5	376
	30.71	5	376
	29.98	5	376
	29.35	5	379
Mother and Daughter	44.8	5	622
	40.39	5	623
	37.71	5	622
	35.97	5	626
Akiyo	45.58	5	394
	41.35	5	391
	38.67	6	387
	36.73	6	392
Silent	34.3	5	359
	32.94	5	363
	31.85	5	362
	30.97	5	366
	30.27	5	372
Mobile	31.89	5	331
	30.06	6	332
	28.64	6	333
	27.5	6	333
	26.56	4	332
	25.76	5	333

the source  $X$  and the SI  $Y$  can increase the joint probability distribution  $P(X, Y)$ . In the Wyner-Ziv coding scheme, as the joint probability distribution increases, the bit rate can be reduced accordingly. Compared with the rapid reduce of the rate, the increase of the distortion is slow. So a better rate and distortion trade-off can be achieved when the non-uniform quantization is applied. Therefore, a rate-distortion optimization function is adopted in the modified Lloyd-Max quantizer design algorithm which iteratively derives from the optimal quantization partition. An estimation of the rate and distortion model at encoder side is presented in this paper. In the modified Lloyd-Max quantizer design algorithm, a novel quantization partition updating method is proposed. The experimental results show that the non-uniform quantization based on RD optimization effectively improves the coding performance of the pixel domain DVC system. The efficiency of the non-uniform scalar quantizer should be further explored in the case of transform domain WZVC, such as the DCT domain WZVC and the wavelet domain WZVC.

**Acknowledgments** This work is supported by the National Natural Science Foundation of China (No. 61103064) and the Science and Technology Program of Beijing Municipal Commission of Education (No.KM201010005011).

## References

1. Aaron A, Setton E, Girod B (2003) Towards practical Wyner-Ziv coding of video. Proc. IEEE International Conference on Image Processing, ICIP 2003, Barcelona, Spain, Sept.
2. Aaron A, Rane S, Setton E, Girod B (2004) Transform-domain Wyner-Ziv codec for video. Proc. Visual Communications and Image Processing (VCIP) 2004, San Jose, CA, Jan.
3. Artigas X, Ascenso J, Dalai M, Klomp S, Kubasov D, Ouaret M (2007) The DISCOVER codec: architecture, techniques and evaluation. Proc. Picture Coding Symposium (PCS), Lisboa, Portugal, Nov.
4. Becker-Lakus A, Leung K-M, Ma Z (2010) Bitwise prediction error correction for distributed video coding. Picture Coding Symposium (PCS) 2010, pp382–385, 8–10 Dec.
5. Berger T (1972) Optimum quantizers and permutation codes. IEEE Trans Inf Theory 18:759–765
6. Berrou C, Glavieux A (1996) Near optimum error correcting coding and decoding: turbo-codes. IEEE Trans. Comm., pp. 1261–1271, Oct.
7. Brites C, Ascenso J, Pereira F (2006) Modeling correlation noise statistics at decoder for pixel based Wyner-Ziv video coding. Proc. Picture Coding Symposium (PCS), Beijing, China, April.
8. Bruce R (1964) Optimum quantization. Sc.D. thesis, M.I.T., May 14
9. Fang S, Li X, Zhang L (2007) A Lloyd-Max-based non-uniform quantization scheme for distributed video coding. Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, SNPD 2007, Vol. 1, pp. 848–853, July.
10. Huang B, Ma J (2007) On asymptotic solutions of the Lloyd-Max scalar quantization. Information Communications & Signal Processing, 6th International Conference on, pp. 1–6.
11. Kraskov A, Stogbauer H, Grassberger P (2004) Estimating joint distribution. Phys Rev E 69:066138
12. Kwong S, Xu L, Zhang Y, Zhao D (2012) Low-complexity encoder framework for window-level rate control optimization. IEEE Trans Ind Electron PP(99): Publication Year: Page(s): 1
13. Lahsini C, Zaibi S, Pyndiah R, Bouallegue A (2011) Distributed video coding in pixel domain using spatial correlation at the decoder. Data Compression Conference (DCC) 2011, pp. 382–385, Mar.
14. Lloyd S (1982) Least squares quantization in PCM. IEEE Trans Inf Theory IT-28:129–137
15. Max J (1960) Quantizing for minimum distortion. IEEE Trans Inf Theory IT-6:7–12
16. Muresan D, Effros M (2008) Quantization as histogram segmentation: optimal scalar quantizer design in network systems. IEEE Trans Inf Theory 54:344–366
17. Ou T-S, Huang Y-H, Chen HH (2011) SSIM-based perceptual rate control for video coding. IEEE Trans. Circuits Syst Video Technol
18. Rebollo-Monedero D, Zhang R, Girod B (2003) Design of optimal quantizers for distributed source coding. in Proc. IEEE Data Compression Conf., Snowbird, UT, pp. 13–22, Mar.
19. Roca A, Morbee M, Prades-Nebot J, Delp E (2008) Rate control algorithm for pixel-domain Wyner-Ziv video coding. Proc. Visual Communications and Image Processing (VCIP) 2008, San Jose, USA, Jan.
20. Sheinin V, Jagmohan A, He D (2006) Uniform threshold scalar quantizer performance in Wyner-Ziv coding with memoryless, additive laplacian correlation channel. Proc. IEEE Conf. Acoust. Speech Sig. Proc., pp. 217–221, May.
21. Slepian D, Wolf J (1973) Noiseless coding of correlated information sources. IEEE Trans Inf Theory 19:471–480
22. Swaroop KVS, Rao KR (2010) Performance analysis and comparison of JM 15.1 and Intel IPP H.264 encoder and decoder. System Theory (SSST), 2010 42nd Southeastern Symposium on, Publication Year: Page(s): 371–375
23. Tu Z, Li TJ, Blum RS (2006) On scalar quantizer design with decoder side information. Information Sciences and Systems, 2006 40th Annual Conference on, pp. 224–229, Mar.
24. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13:600–612
25. Weerakkody WAR, Fernando WAC, Badem MB, Kondo AM (2009) Nonlinear quantisation for pixel domain DVC. Electron Lett 45:261–262
26. Wu X (1991) Optimal quantization by matrix searching. J Algorithms 12(4):663–673
27. Wu X, Zhang K (1993) Quantizer monotonicities and globally optimal scalar quantizer design. IEEE Trans Inf Theory 39:1049–1053
28. Wu B, Guo X, Zhao D, Gao W, Wu F (2006) An optimal non-uniform scalar quantizer for distributed video coding. Proc. IEEE International Conference on Multimedia and Expo, ICME 2006, Toronto, Ontario, Canada, pp. 165–168, Jul.
29. Wyner A, Ziv J (1976) The rate-distortion function for source coding with side information at the decoder. IEEE Trans Inf Theory IT-22(1):1–10
30. Zhang Y, Xiong H, He Z, Yu S, Chen CW (2011) Reconstruction for distributed video coding: a context-adaptive markov random field approach. IEEE Trans Circuits Syst Video Technol 21(8):1100–1114



**Bo Wu** received the B.S. degree in computer science from Jilin University, Changchun, China, in 2001, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2010. Then she joined the School of Biomedical Engineering, Capital Medical University, Beijing, China. Her research interests include image and distributed video coding, image and video coding, three-dimensional video representation and coding.



**Nan Zhang** received the B.S. and M.S. degrees from Henan University, Kaifeng, China, in 1997 and 2001, respectively. She received the Ph.D. degree from the College of Computer Science and Technology at Beijing University of Technology, Beijing, China, in 2006.

From 2007 to 2009, she was a Post-doctorate with the Institute of Digital Media, Peking University, Beijing, China. Then she joined the School of Biomedical Engineering, Capital Medical University, Beijing, China. Her research interests include image and video coding, three-dimensional video representation and coding, and three-dimensional video application in telemedicine.





**Siwei Ma** received the B.S. degree in computer science from Shandong Normal University, Jinan, China, in 1999, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005. From 2005 to 2007, he was a Post-doctorate with the University of Southern California, Los Angeles. Then he joined the Institute of Digital Media, School of Electronics Engineering and Computer Science, Peking University, Beijing, China, where he is currently an Associate Professor. He has published over 30 technical articles in refereed journals and proceedings. His current research interests include image and video coding, video processing, video streaming, and transmission.



**Debin Zhao** received the B.S., M.S., and Ph.D. degrees in computer science from the Harbin Institute of Technology (HIT), Harbin, China, in 1985, 1988, and 1998, respectively.

Dr. Zhao was a Lecturer from 1989 to 1993 and an Associate Professor from 1993 to 2000 in the Department of Computer Science, HIT. He is currently a Professor with the Department of Computer Science, HIT, and also an Adjunct Professor with the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China. He has published over 200 technical articles in refereed journals and conference proceedings. His current research interests include image and video coding, video processing, video streaming and transmission, and pattern recognition. Dr. Zhao was the recipient of three National Science and Technology Progress Awards of China (Second Prize).



**Wen Gao** (M'92–SM'05–F'09) received the B.S. and M.S. degrees in computer science from Harbin University of Science and Technology and Harbin Institute of Technology (HIT), China, in 1982 and 1985, respectively, and received the Ph.D. degree in computer science from HIT, China and in electronics engineering from the University of Tokyo, Tokyo, Japan, in 1988 and 1991, respectively. Before joining Peking University, Beijing, China, he was a Full Professor of computer science with the Harbin Institute of Technology, Harbin, China from 1991 to 1995, and with the Chinese Academy of Sciences (CAS), Beijing, China, from 1996 to 2005. With CAS, he served as a Professor (1996–2005), the Managing Director of the Institute of Computing Technology (1998–1999), the Executive Vice President of the Graduate School of CAS (2000–2004), and a Vice President (2000–2003) of the University of Science and Technology of China, Hefei, China. He is currently a Professor of computer science with Peking University. He has published extensively, including four books and over 500 technical articles in refereed journals and conference proceedings. His current research interests include image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics. Dr. Gao is the Editor-in-Chief of the *Journal of Computer* (a journal of the China Computer Federation), an Associate Editor of the *IEEE Transactions on Circuits and Systems for Video Technology*, an Associate Editor of the *IEEE Transactions on Multimedia*, an Associate Editor of the *IEEE Transactions on Autonomous Mental Development*, an Area Editor of the *EURASIP Journal of Image Communications*, and an Editor of the *Journal of Visual Communication and Image Representation*. He chaired a number of prestigious international conferences on multimedia and video signal processing, and also served on the advisory and technical committees of numerous professional organizations.