AN OPTIMAL NON-UNIFORM SCALAR QUANTIZER FOR DISTRIBUTED VIDEO CODING

Bo Wu[1], Xun Guo[2], Debin Zhao[1,2], Wen Gao[1,2], Feng Wu[3]

[1]Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China
[2]School of Computer Sciences, Harbin Institute of Technology, Harbin, 150001, China
[3]Microsoft Research Asia, Beijing, 100080, China

ABSTRACT

In this paper, we propose a novel algorithm to design an optimal non-uniform scalar quantizer for distributed video coding, which aims at achieving a coding rate close to joint conditional entropy of the quantized video frames given the side information. Wyner-Ziv theory on source coding is employed as the basic coding principle and the asymmetric scenario is considered. In this algorithm, a probability distribution model, which considers the influence of the joint distribution of input source and side information to the coding performance, is established and used as the optimality condition firstly. Then, a modified Lloyd Max algorithm is used to design the scalar quantizer to give an optimal quantization for input source before coding. Experimental results show that compared to uniform scalar quantization, proposed algorithm can improve coding performance largely, especially at low bit rate.

1. INTRODUCTION

In traditional video coding scheme, e.g. H.264/AVC, asymmetric complexity exits in encoder and decoder. With the motion estimation module, the encoder cost much more time than decoder. However, in many applications such as sensor networks, video compression has to be done in the cameras with constrained power and low processing ability. In these applications, simple and real time encoder is needed. Thus, the complexity which mainly comes from becomes a big burden. We know that most of the encoding complexity mainly comes from the correlation exploitation part, such as motion estimation. To tackle this problem, distributed source coding (DSC) theory has been used in the practical video coding, which can move most complexity from encoder to decoder.

Theory of Slepian-Wolf [1] shows that even if correlated sources are encoded without getting information from each other, coding performance can be as good as dependent encoding if the compressed signals can be jointly decoded. And also, Wyner and Ziv have extended the theory to the lossy source coding with side information [2]. Recently, several practical Wyner-Ziv coding techniques have been proposed for video coding, namely distributed video coding (DVC). In [3], Aaron and Girod proposed a DVC scheme using turbo codes. In [4], Praddhan and Ramchandran proposed a DVC framework based on syndrome of codeword co-set. Video frames are encoded independently and decoded jointly in these schemes.

Quantization is a key point which can affect coding performance largely in above DVC scheme. A lot of work has been done on the scalar quantizer design in DSC recently. In [5], an efficient algorithm for finding global optimal scalar quantizer with continuous code cells was proposed. And in [6], the authors proposed a more efficient quantizer with disconnected partition regions. A general rate distortion optimal scalar quantizer was proposed in [7], which considers different rate measures and output non-uniform disconnected partition regions. However, in practical DVC application, many factors, e.g. algorithm complexity, different contents of video sequences and side information generation, can limit the quantizer efficiency. Thus, uniform scalar quantizer is always adopted by this kind of scheme.

Therefore, in this paper, we propose an algorithm for designing the optimal non-uniform scalar quantizer, which is fit for practical DVC scheme. The scheme proposed in [3] is employed as our basic idea, in which input video frames are classified into two categories in our scheme, namely intra frame and Wyner-Ziv frame. Wyner-Ziv frames are encoded using turbo codes and decoded with side information. The optimal quantizer is used to give a reasonable non-uniform quantization to the Wyner-Ziv frame before coding. An extended Lloyd Max algorithm with rate distortion (RD) model is used to get the optimal quantization partition, which can achieve a good tradeoff between complexity and coding performance.

The remainder of this paper is organized as follows: Section 2 describes DVC scheme employed by this paper including analysis for the quantization. Section 3 gives design process of proposed quantizer in details including analysis of the optimality condition and the quantization steps. Section 4 gives the experimental results. And conclusions are drawn in section 5.

2. DVC SCHEME USING TURBO CODES

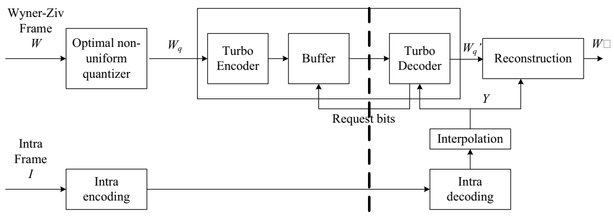

Figure 1. Diagram of DVC scheme based on optimal non-uniform quantizer

Figure 1 shows the coding scheme of the employed DVC system, in which frames of the input video sequence are classified into intra frames and Wyner-Ziv frames. Intra frames, i.e. odd frames, are coded with the traditional DCT based intra coding method. Wyner-Ziv frames, i.e. even frames, are encoded with turbo codes and decoded using logarithm max a posteriori (log-MAP) algorithm together with side information which is generated from the intra frames.

At encoder, each pixel of W is quantized using the optimal non-uniform scalar quantizer with 2^M levels to generate a quantized symbol firstly. Then, quantized frame W_q is sent to the turbo encoder which consists of two constituent convolutional encoders of rate 4/5 [8].

At decoder, side information Y is generated using interpolation based on motion compensated prediction. Because Y can be seen as the noisy version of W through a virtual correlation channel, the conditional probability model between W and Y can be defined as

$$f(d) = \tfrac{\alpha}{2} e^{-\alpha|d|},$$

where d is the difference between W and Y. The conditional probability is used to predict which quantization partition the source W most probable lies in. Then, the quantized Wyner-Ziv frame W_q is decoded jointly using side information Y, transmitted parity bits and the probability. After that, decoded $W_q^{'}$ can be generated and be used to reconstruct $W^{'}$ together with Y. Both the interleaver and the decoding process are based on symbol level.

At traditional video encoder, quantization is done in the real prediction residue. While in DVC case, quantization is done in original Wyner-Ziv frame and the quantized Wyner-Ziv frame has to be restored using side information at decoder. Thus, the quantization can impact on coding performance largely. If the distribution of the Wyner-Ziv frame and the side information in each quantization partition mismatches too much, the conditional probability $f(d)$ will be not accurate and the coding performance will be very poor. In most cases, uniform scalar quantizer is adopted because it can limit max prediction error within a fix range. If the quantizer is designed according to the distribution of Wyner-Ziv frame and side information, performance will be increased largely. Therefore, how to get the optimal quantizer is one of the most important problems in distributed video coding.

3. DESIGN OF OPTIMAL NON-UNIFORM SCALAR QUANTIZER

3.1. Analysis of Optimality Condition

In this section, we will analyze the optimal scalar quantization for distributed video coding and prove that the quantization considering both the input source and the side information can achieve local optimality.

X and Y are the bin index of the quantized original sequence and the side information respectively. X and Y can be considered as independent replicas of a pair of dependent random variables (X,Y). Our aim is to exploit the correlation of X and Y to decrease the coding rate close to the Wyner-Ziv bound. When the side information is available at both of the encoder and decoder, the rate-distortion function of the source coding is $R_{X|Y}(D)$. It is the Wyner-Ziv bound under the distortion constraint D. When the side information is only available at the decoder side, the coding rate is $R^{WZ}_{X|Y}(D)$. It can be proved [9] that the rate loss is

$$R^{WZ}_{X|Y}(D) - R_{X|Y}(D) \le \inf_{\{N \in X : E(d(N)) < D\}} I(W; W+N) \tag{1}$$

Where W is X-U, and N is the channel noise. When D approaches to 0, the rate loss goes to 0. Thus the rate of the Wyner-Ziv problem achieves to its bound $R_{X|Y}(D)$. The distortion is

$$E[d(\hat{X}(Z,Y) - X)] \le D \tag{2}$$

Z is a prediction for X. Z and Y forms a reconstruction function $\hat{X} = \hat{X}(Z,Y)$ and $E[d(\hat{X} - X)] \le D$. Z takes over all random variables to achieve the infimum under the distortion constraint. From equation (2), we can see that if the correlation of X and Y increases, the equation will approach to 0.

In the turbo decoding process, the quantizer and the conditional probability distribution model of the quantization index should be designed together to exploit statistical dependencies. We use $P(q_k \mid y)$ to describe the conditional probability, where q_k is the bin index of the quantized original source at time k. We have the following equation

$$P(q_k \mid y) = \frac{\alpha}{2}\exp(-\alpha \mid Q^{-1}(q_k) - y \mid), \quad (3)$$

where $Q^{-1}(q_k)$ is an estimate reconstruction value for q_k and y is the side information. Through the trellis based SISO algorithm, the *a posteriori* probabilities $P(y \mid q_k)$ of the output symbols are deduced based on the bin index's conditional probability $P(q_k \mid y)$ and the probability of the parity bits. Experientially, $P(q_k \mid y)$ has great impact on $P(y \mid q_k)$. That is, more accurate $P(q_k \mid y)$ leads to fewer bits for recovering the quantizer index sequence. Consequently, this problem is a cluster problem and for one particular cell q_k, it should consist of all elements closer to its reproduction $Q^{-1}(q_k)$ than to any other cell's output reproductions.

We investigate the Lloyd Max algorithm and found that the nearest neighbor condition of the Lloyd quantizer makes this condition possible. That is,

$$R_i \subset \{x : d(x, \hat{x}_k) \le d(x, \hat{x}_i); all \quad k \ne i\}, \quad (4)$$

where $\hat{x}_k = Q^{-1}(q_k)$ and $\hat{x}_k$ is the reproduction of the bin index q_k

Intuitively, we hope that input symbols with high probability can be gathered in one large partition cell. Since the side information has high correlation with the original source, they are similar in distribution. Thus, we can refer the distribution of original source the at the encoder side to find the optimal quantizer partition.

The rate measure is important because it impacts the choosing of the quantization partition directly. We adopt the conditional entropy $R = H(\hat{X} \mid Y)$ as the rate measure. The entropy is defined as follows,

$$H(\hat{X} \mid Y) = -\sum_{\hat{x}}\sum_{y} P(\hat{x}, y)\log P(\hat{x} \mid y) \quad (5)$$

We predefine a non-accurate Gaussian distribution as the *a posteriori* information $P(y \mid \hat{x})$:

$$P(y \mid \hat{x}) = \frac{1}{\sigma_1\sqrt{2\pi}}\exp(-\frac{(y-\hat{x})^2}{2\sigma_1^2}) \quad (6)$$

This is reasonable because those source elements, in a certain cell, whose position is closer to the cell's reproduction value, will be considered with higher probability to lie in this cell. Thus, the rate at time *k* finally equals to

$$\begin{aligned} R &= -\sum_{q_k}\sum_{Y} P(\hat{x}_k, y)\log P(\hat{x}_k \mid y) y \\ &= -\sum_{q_k} \int_{y \in C_k} f(\hat{x}_i) P(y \mid \hat{x}_i)\log P(\hat{x}_i \mid y) dy \\ &= -\sum_{q_k}\sum_{y \in C_k} f(\hat{x}_i) P(y \mid \hat{x}_i) g(y)\Delta y, \quad (7) \end{aligned}$$

where f is the distribution of the source. Because the probability $P(\hat{x}_k \mid y)$ is less than 1, the item $-\log P(\hat{x}_k \mid y)$ is larger than 0. The integral part can be seen as the area of a partition region. Therefore, we can say that the rate is directly proportional to the sum area of all partition regions.

3.2. Quantization Algorithm

The optimal quantization algorithm for DVC is as follows:

1) Analyze the distribution of Wyner-Ziv frame by histogram.
2) Choose an initial reproductions set $\{i \le n : \hat{x}(q_i)\}$ for certain quantization level n. Set the iteration counter k to 1.
3) Using the nearest neighbor condition to gather the source elements to different partitions. Compute the new optimal reproduction values for these partitions with the centroid condition.
4) Calculate the entropy and the distortion for this quantization partition, using equation (7). Compute the cost function $J^{(k)} = D + \lambda R$ and compare it with the previous step cost $J^{(k-1)}$ to decide continue or stop.
5) Choose a new initial reproductions set which expand the cell which contained those inputs with high probability. Increase counter k and go back to 2).

The definition of the conditional probability model $P(y \mid \hat{x})$ is almost general. Thus, this algorithm can be used in both pixel domain scheme and transform domain scheme.

4. EXPERIMENTAL RESULTS

In order to verify the efficiency of proposed algorithm, some experimental results are presented. Firstly, we test our method using Gaussian sources. We adopt memoryless Guassian source Y and X. Y = X + Z, where $Z \sim N(0, \sigma_Z^2)$ is memoryless white Gaussian noise and independent with Y. The side information with CSNR 9.54, 14.47 and 19.55 are tested. For each CSNR, 4 different 4-level non-uniform quantizers are used. The relationship between distortion and rate needed for decoding is shown in figure 2. We can see that various non-uniform quantizer can reduce coding bits significantly without increasing distortion too much. Then, we compare the uniform quantization and proposed method using the source with CSNR 14.47 dB and result is shown in table 1. We can see that with proper quantization, the decoding bits needed in proposed method decreased very much, while the PSNR is nearly the same.

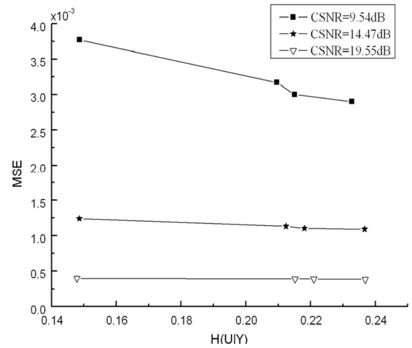

Figure 2. Trade off between MSE and H(U|Y)

Table 1. Wyner-Ziv coding results for uniform and proposed optimal quantization

	Quantization Partition	Reproductions	Bit Rate (kbits)	PSNR
uniform	[0,63,127,191,255]	32,96,160,224	403.92	33.24
non-uniform	[0,10,22,77,255]	2,18,26,128	71.76	33.16

We also compare the performance of proposed algorithm two video sequences (foreman, mother and daughter) in QCIF format are used to evaluate the performance of proposed quantizer. For foreman sequence, 100 frames are tested. And for mother & daughter sequence, 200 frames are tested. In the DVC scheme employed by our simulation, I frames are coded with H.263+ scheme and W frames are coded with turbo codes using generation matrix in [8]. We use the IWIW structure with GOP size 30. Four bit rate points are tested, i.e. QP 24, QP 28, QP 32 and QP 36, and 4-level quantization is used at each point. Results are shown in Figure 4 and up to 2 dB gain can be achieved for foreman sequence and 1 dB gain for mother & daughter sequence.

5. CONCLUSION

In this paper, we have presented an algorithm for designing an optimal non-uniform scalar quantizer for practical DVC scheme. The optimal quantization condition is analyzed and a modified Lloyd Max algorithm with a rate distortion model, which considers the correlations between quantized Wyner-Ziv frame and side information, is used in the quantization process. Without increasing encoder complexity too much, the proposed algorithm achieves good performance.

6. ACKNOWLEDGMENTS

This work has been supported by Key Technologies for High Efficient Video Coding under the contact no. 60333020 and Key Technologies for Dealing with Streaming Media under the contact no. 4041003.

REFERENCES

[1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, pp.471-480, July 1973.

[2] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. 22, pp.1-10, Jan.1976.

[3] A. Aaron, S. Rane, R. Zhang, and B. Girod, "Wyner-Ziv coding for video: Applications to compression and error resilience," in *IEEE Data Compression Conference,* DCC-2003, Mar. 2003.

[4] R. Puri and K. Ramchandran, "PRISM: a new robust video coding architechture based distributed compression principles," in *Allerton Conference on Communication*, Control, and Computing, Oct. 2002.

[5] D. Muresan and M. Effros, "Quantization as histogram segmentation: Globally optimal scalar quantizer design in network systems," in *Proc.IEEE Data Compression Conference (DCC)*, Snowbird, UT, Mar. 2002, pp. 302–311.

[6] M. Effros and D. Muresan, "Codecell contiguity in optimal fixed-rate and entropy-constrained network scalar quantizers," in *Proc. IEEE Data Compression Conference (DCC)*, Snowbird, UT, Apr. 2002, pp. 312–321.

[7] D. Rebollo-Monedero, R. Zhang, and B. Girod, "Design of optimal quantizers for distributed source coding," in *Proc. IEEE Data Compression Conf.*, Snowbird, UT, Mar. 2003, pp. 13–22.

[8] A. Aaron and B. Girod, "Compression with side information using turbo codes," in *Proc. IEEE Data Compression Conf.*, 2002, pp.252-261.

[9] Ram Zamir , "The Rate Loss in the Wyner-Ziv Problem," IEEE Trans. Inform. Theory , vol. 42, no. 6, pp. 2073-2084, November 1996.

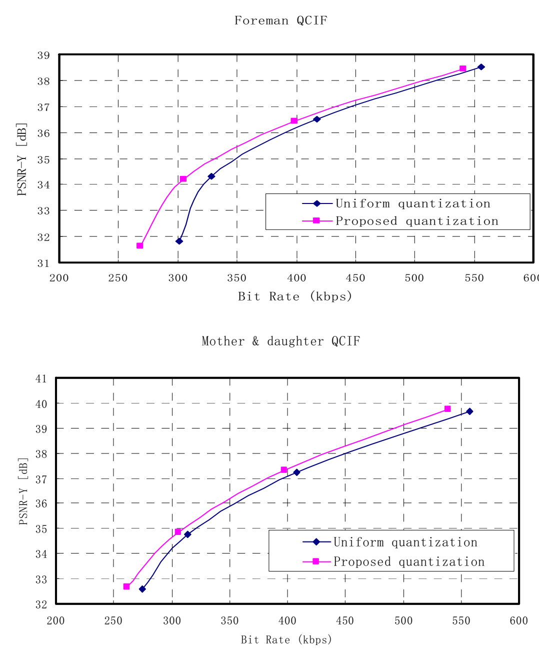

Figure 4. Experimental results for foreman and mother & daughter