Rate Control Based on Zero-Residue Pre-Selection for Video Transcoding

Chi-Wang Ho Oscar C. Au S.-H. Gary Chan Hoi-Ming Wong Shu-Kei Yip
The Hong Kong University of Science and Technology,
Clear Water Bay, Kowloon, Hong Kong, China
Email: {jodyho, eeau, gchan, hoimingw, sukiyip}@ust.hk

***Abstract*— A common issue in video transcoding for heterogeneous network environment is to efficiently and accurately reduce the bit-rate such that the distortion is minimized under a given rate constraint. To convert the bit-rate of an encoded video to match the channel capacity, in general, re-quantization is done on the DCT coefficients with larger quantization step size. Most existing rate control algorithms for video transcoding in the literature calculate quantization parameters (QPs) of macroblocks (MBs) based on a relationship between certain properties of coded video and bit-rate. They reduce the computational complexity by simplifying the R-D model and reusing the statistics information of input video. In this paper, we propose a Zero-Residue Pre-Selection (ZRPS) mechanism to select only a portion of MBs to apply the rate control in video transcoding. TMN-8 is used to evaluate the impact of ZRPS. Experimental results show that, as compared to the original TMN-8 rate control scheme, TMN-8 with ZRPS achieves up to 1.60 dB gain, in term of PSNR, and requires less than 50% of the computational complexity compared to TMN-8, depending on the characteristics of the video content.**

I. INTRODUCTION

Video delivery over network is a challenging problem because of the heterogeneity in network channel capacity, video formats and devices. Due to such heterogeneity, interoperability between different networks, formats and devices is important. There are two possible interoperability solutions, including scalable video coding and video transcoding. Scalable video coding suffers from poor coding efficiency due to additional overheads and has not been widely used. On the other hand, video transcoding provides high flexibility for the conversion between different bit-rate and formats and is quite popular. Generally speaking, video transcoding can be defined as the conversion of one encoded video to another. Research on video transcoding usually focuses on bit-rate reduction, spatial resolution reduction as well as temporal resolution reduction [1], [2]. In this paper, we will mainly focus on the rate control scheme for bit-rate reduction transcoding.

We briefly present some previous works as follows. Rate control for video transcoding has long been proposed and studied [3]–[9]. Generally, all rate control schemes designed for video coding are applicable to transcoding, such as MPEG-2 TM5 [10], H.263+ TMN-8 [11], etc. Direct application of these algorithms are usually considered as a waste of information since they do not take advantage of the information available in the input video. A number of researchers have proposed various rate control algorithms for transcoding based on the statistics or some properties in the input video with the requirement of some assumptions or offline training. Lei *et al.* proposed a rate control scheme based on the linear relationship between the number of bits produced and the number of quantized non-zero AC coefficients [5]. However, the determination of QP for each MB is done by using a trained table. The performance may depend on the accuracy of the table. Seo *et al.* proposed a piecewise linearly decreasing model for fast rate control in transcoding, which determines the new QP by using the abrupt bit-rate reduction property and the input QP [9] . The performance of this scheme depends on which rate control scheme was applied on the input video. To the best of our knowledge, most of the rate control algorithms for transcoding in the literature mainly studied the simplification of the cost function by reusing the information of input video. There is no paper considering the pre-selection of MBs by using the statistics of input video, and then performing rate control only on the selected sub-set of MBs to reduce the complexity. In this paper, we propose a Zero-Residue Pre-Selection (ZRPS) mechanism to select a sub-set of MBs, which are expected to have non-zero residue after re-quantization. Then rate control will only be applied to these MBs. This can significantly reduce the computational complexity as, normally, only a small portion of MBs need to be processed.

The paper is organized as follows. We first propose the Zero-Residue Pre-Selection (ZRPS) mechanism to select the MBs to apply rate control algorithm. In Section III, we present the modified version of TMN-8 based on ZRPS followed by the analysis of its impact on TMN-8. Finally, we present our experimental results and end with concluding remarks.

II. ZERO-RESIDUE PRE-SELECTION (ZRPS) MECHANISM

In this section, we first define some terms used in this paper and describe the observation we have, followed by our proposed ZRPS mechanism.

In the bit-rate reduction transcoding, the residual MBs are de-quantized and later re-quantized by a larger QP. Due to re-quantization, the number of bits spent on the luminance and chrominance tends to be lower in the output video than in the input video. We classify the MBs into four classes namely, Z-Z, Z-NZ, NZ-Z and NZ-NZ. Here ’Z’ means all DCT coefficients in all of the four luminance and two chrominance 8x8 blocks are zero. ’NZ’ means some DCT coefficients in the four luminance and two chrominance 8x8 blocks are non-zero. The first term corresponds to the input MB and the second term corresponds to the corresponding output MB. For example, ’NZ-Z’ means some DCT coefficients of the input MB are non-zero and all DCT coefficients of the output MB are zero. We will ignore the Z-NZ group in the following discussion as it rarely occurs. Experimental results show that the percentage of MBs belonging to the Z-NZ group is negligible (typically only less than 1%). In the Z-NZ group, bits are needed to encode the MBs in the output video but not in the input. This can occur when the distortion from the previous frame is very large resulting in large prediction residue and non-zero DCT coefficients after re-quantization in the current MB. In what follows, a Zero-Residue Pre-Selection mechanism is presented to classify the MBs into two groups: zero-residue group (ZRG) and non-zero-residue group (NZRG). ZRG includes the MBs in Z-Z and NZ-Z group, and NZRG includes the MBs in Z-NZ and NZ-NZ

This work is supported, in part, by Innovation Technology Commission (ITS/122/03) and Research Grant Council (HKUST6156/03E) in Hong Kong.

TABLE I

THE GROUPING OF THE MBS ACCORDING TO THEIR DCT COEFFICIENTS AFTER RE-QUANTIZATION

ZRPS Group	Group	Input DCT coefficients of MB	Output DCT coefficients of MB
ZRG	Z-Z	All Zero	All Zero
	NZ-Z	Some Non-Zero	All Zero
NZRG	Z-NZ	All Zero	Some Non-Zero
	NZ-NZ	Some Non-Zero	Some Non-Zero

group. The grouping of MBs are summerized in Table I. We will claim that the rate control needs only to be applied on the NZRG. This can reduce complexity while achieving superior quality compare to blindly applying rate control to the whole frame.

In the ZRPS mechanism, we first define a zero-residue map for frame t, $ZRM_t\left[i\right]$, where i is the MB index, as follows. If all quantized coefficients of the MB i (including all luminance and chrominance blocks in the MB) are zero after re-quantization, $ZRM_t\left[i\right] = 0$ (ZRG), otherwise, $ZRM_t\left[i\right] = 1$ (NZRG). Since this zero-residue map can be obtained only after re-quantization, we have to predict this before re-quantization in order to use it for rate control. ZRPS mechanism provides a way to predict the $ZRM_t\left[i\right]$ for the current frame t based on the previous input and output frames $t-1$ and the current input frame.

Firstly, we need to define some variables. Let b_i^t be the number of bits spent to code the coefficients of MB i in input frame t, $\tilde{b}_i^t$ be the estimated number of bits needed to code the coefficients of MB i in output frame t and Δ_i^{t-1} is the amount of bit reduction for coding the coefficients of MB i of frame $t-1$ from the input video to output video. T_1 and T_2 are two thresholds, which represents in term of number of bits, used in ZRPS. Then, the ZRPS mechanism is shown as follows:

Step 1: Initialize the ZRM_t for frame t based on the quantized coefficients of frame $t-1$. If all quantized coefficients of MB i in frame $t-1$ are zero, $ZRM_t\left[i\right] = 0$, otherwise, $ZRM_t\left[i\right] = 1$.
Step 2: Estimate the number of bits needed for MB i, $\tilde{b}_i^t$, as $b_i^t - \Delta_i^{t-1}$.
Step 3: Check each MB with $ZRM_t\left[i\right] = 1$. If $\tilde{b}_i^t < T_1$, mark $ZRM_t\left[i\right] = 0$.
Step 4: Check each MB with $ZRM_t\left[i\right] = 0$. If $\tilde{b}_i^t > T_2$, mark $ZRM_t\left[i\right] = 1$.

In step 1, using transcoded output frame $t-1$, $ZRM_{t-1}\left[i\right]$ can be generated and used as a starting point for predicting the $ZRM_t\left[i\right]$. Since there is a relationship between the quantized DCT coefficients and the number of bits generated after encoding, the bit-count information of MB i in frame $t-1$ can be used in step 2 - 4 to predict the resulting bits needed for the MB i in frame t, and hence predict whether the coefficients of MB i is all zero or not.

Suppose all cofficients of MB i in frame $t-1$ is re-quantized to zero and $ZRM_t\left[i\right]$ is initially set to 0. Then, the reduction of bits of this MB i in frame $t-1$ and $\tilde{b}_i^t$ are computed as described in above steps. The resulting $\tilde{b}_i^t$ can be either positive or negative. If it is positive, the number of bits needed for coding the MB i is expected to be remain positive (non-zero). If it is negative, this means the number of bits reduced after re-quantization of the MB i in frame $t-1$ is larger than the number of bits spent for coding the coefficients of MB i in frame t. If $\tilde{b}_i^t$ is smaller than T_1, then all the coefficients of MB i in frame t will be very likely to be quantized to zero. So it should not belong to NZRG and needs to switch to ZRG. Similar concept is applied to step 4. In the ZRPS, T_1 and T_2 are updated according to the estimation error, which refers to the MBs classify wrongly in ZRPS, after transcoded the frame t.

After the above steps, a predicted zero-residue map, ZRM_t, is obtained. This can be used as an indicator for selecting the MBs to perform rate control algorithm.

III. MODIFIED TMN-8 BASED ON ZRPS

Having illustrated the ZRPS mechanism in the previous section, let us introduce a simple and yet efficient modification of TMN-8 rate control with ZRPS. This can reduce the number of MBs needed to be processed by MB-layer rate control, and hence speeds up the rate control, but, at the same time, provides better PSNR. The frame-layer bit allocation is the same as TMN-8. Then, in the MB-layer rate control, the TMN-8 without ZRPS is applied on the first P frame after I frame since we need to do the initialization of ZRM_t and use the bit-count information to predict the ZRM_{t+1} for the frame $t+1$. According to our experiment, initially, the threshold T_1 and T_2 of ZRPS mechanism is reasonably good to set to -10 and 10 respectively for most of the common video sequences. For the subsequent P frames, the TMN-8 with ZRPS is used, we call this scheme as ZRPS-TMN-8. The ZRPS-TMN-8 is summarized as follows:

Macroblock-layer rate control:
Step 1: *Create the ZRM_t* — Follow the steps of ZRPS mechanism described in previous section.
Step 2: *Compute the sum of weighted standard deviation of all MBs S_1* — Based on the ZRM_t, we compute the variance of the i^{th} MB prediction error if $ZRM_t\left[i\right] = 1$ and compute $S_1 = \sum_{k=1\ \&\ ZRM_t[k]=1}^{N} \alpha_k \sigma_k$. The number of MBs with $ZRM_t\left[i\right] = 1$ is defined as N_{ZRM} and used in the rate control instead of the total number of MBs in a frame. The equation of calculating the weighting α_i is modified and shown as below:

$$\alpha_i = \begin{cases} 2\frac{B}{AN_{ZRM}}(1-\sigma_i)+\sigma_i, & \frac{B}{AN_{ZRM}} < 0.5 \\ 1, & \text{otherwise.} \end{cases}$$

where A is the number of pixels in a macroblock and B is the bit budget for current frame.
Step 3: *Initialize the counter and the model parameter K and C* — This is exactly the same as the TMN-8.
Step 4: *Compute Q_i^* for i^{th} MB* — If $ZRM_t\left[i\right] = 1$, calculate the QP same as TMN-8, otherwise, copy QP used by the previous MB. If we are running out of bits, set Q_i^* = maximum quantization step size. Finally, set $Q_{prev} = Q_i^*$.
Step 5: *Update bit budget for remaining MBs B* — This is exactly the same as the TMN-8.
Step 6: *Compute the model parameters $\hat{K}_i$ and $\hat{C}_i$ for the i^{th} MB* — This is same as the TMN-8 and performed only when $ZRM_t\left[i\right] = 1$.
Step 7: *Update K and C using $\hat{K}_i$ and $\hat{C}_i$ and the counter* — This is exactly the same as the TMN-8.
Step 8: *Update T_1 and T_2 for the ZRPS* — The T_1 and T_2 are updated according to the error experienced in this frame. Then, repeat from step 4 until all the MBs are finished.

We claim that it is sufficient to apply rate control only on non-zero residue group (NZRG). The results in section V show that the performance of the ZRPS-TMN-8 is better than TMN-8 in term of both speed and PSNR.

IV. Analysis of Modified TMN-8

In the ZRPS-TMN-8, there is one additional step prior to the MB-layer rate control, which selects a sub-set of MBs in the current frame to participate in the rate control. The pre-selection process is according to the proposed ZRPS mechanism. There are two reasons that help the ZRPS-TMN-8 to outperform the original TMN-8.

A. More Accurate Model Parameter K for large residue MBs

Firstly, for a typical TMN-8 rate control algorithm, the model parameter K is updated depending on the actual model parameter $\hat{K}_i$ of the i^{th} MB. This adapts the model according to the statistics of previous coded MBs. However, the value of $\hat{K}_i$ varies a lot across the whole frame. With the ZRPS, a more accurate and stable model parameter can be obtained for MBs with large residue. This is because the MBs with residue coefficients tends to zero, are usually classified into ZRG, and the undesirable effect of these MBs with very little residue on the model parameters are eliminated. We calculate the average absolute difference between K and $\hat{K}_i$ for the whole sequence, and ZRPS-TMN-8 obtains a smaller error than TMN-8. For example, in the children sequence, the errors of ZRPS-TMN-8 and TMN-8 are 0.149 and 0.206 respectively.

B. Smaller Quantization Overhead and QP for large residue MBs

In the TMN-8, the quantization overhead at the low bit-rate situation is controlled by α_i. However, due to the slight difference between the bits spent on different MB, the buffer level may slight fluctuate and the QPs may slightly vary over the whole frame. In the ZRPS-TMN-8, the QP of the MB with $ZRM_t\,[i] = 0$ is copied from the previous MB. There is no quantization overhead across these MBs. Since the percentage of MBs belonging to ZRG in a typical video frame is quite large (usually over 80%) significantly smaller quantization overhead is needed. Experimental results show that the ZRPS-TMN-8 typically can achieve smaller quantization overhead compared to the TMN-8.

In addition to the smaller quantization overhead, we also have more bits for the MBs in NZRG. Since we only consider the MBs with $ZRM_t\,[i] = 1$ in the calculation of S_i, which is the sum of weighted standard deviation of all MBs, the S_i is reduced by about 40% depending on the content of video sequence. It results in a smaller QP for these MBs and hence smaller distortion. Although the number of bits spent by these MBs may increase, but experimential results show that these extra bits, in most cases, can be compensated by the bit saving from the smaller quantization overhead. Overall, the proposed ZRPS-TMN-8 gives higher PSNR.

V. Experimental Results

We implemented the proposed and TMN-8 rate control scheme in a H.263-to-H.263 transcoder based on H.263+ software developed by UBC [12], which is simply a cascaded of a decoder and an encoder. In this transcoder, the motion vectors from the input video are re-used with a small range refinement search. Thirteen QCIF test sequences are used, each with frame rate of 30 Hz and originally encoded in 384kbps. The first frame was intra-coded (I frame) with QP = 20. The remaining frames were all inter-coded (P frames). Then, these video are transcoded to 64kbps and 96kbps.

Table II shows the actual bit-rates achieved and the percentage of MBs processed by the two rate control strategies for converting a set of QCIF video sequences from 384kbps to 64kbps and from 384kbps to 96kbps. Observed that our proposed ZRPS-TMN-8 achieves similar bit-rate compared to TMN-8.

TABLE II
Comparison of bit-rate achieved by TMN-8 and the proposed ZRPS-TMN-8

Name	**384kbps to 64kbps**		**384kbps to 96kbps**	
	TMN-8	**Proposed**	**TMN-8**	**Proposed**
akiyo	64.38	64.15	96.50	96.44
children	64.08	63.64	96.35	96.08
coastguard	64.24	64.24	96.36	96.35
container	64.24	63.56	96.45	96.20
foreman	64.24	64.29	96.36	96.37
hall monitor	64.23	63.79	96.36	96.32
mobile	64.28	64.36	96.35	96.35
m&d	64.26	64.31	96.35	96.42
sean	64.28	64.10	96.45	96.37
silent voice	64.24	64.24	96.35	96.39
stefan	64.52	64.62	96.68	96.66
table	63.90	61.59	96.32	95.28
weather	64.11	64.27	96.39	96.39
Average	**64.23**	**63.94**	**96.41**	**96.28**

In Table III, we show the performance comparison between the two rate control schemes in terms of PSNR gain and speed. Comparing the total number of P frames encoded by the two rate control schemes, the proposed ZRPS-TMN-8 performs similarly and consistently as TMN-8. The average PSNR achieved by ZRPS-TMN-8 outperforms the one achieved by TMN-8, especially in sean and weather. Up to 1.60 dB PSNR gain is observed in comparison with TMN-8. Figure (1) and (2) show the PSNR over the test sequence 'weather' and 'children' respectively. The curves of ZRPS-TMN-8 are significantly higher than the one of TMN-8. In term of speed, since only a small portion of MBs is involved in MB-layer rate control algorithm, the speed up factor is defined in terms of the number of MBs processed by the rate control.

$$\text{Speed up factor} = \frac{\text{the total number of MBs in the sequences}}{\text{the number of MBs processed}}$$

We can see that the speed up factor ranges from 1.41 to 4.55 times of the original TMN-8 among all of the test sequences. This significantly speeds up the rate control in video transcoding process.

VI. Conclusion

We have presented a Zero-Residue Pre-Selection (ZRPS) mechanism to select the MBs which are expected to have non-zero quantized coefficients after re-quantization. With the ZRPS, a sub-set of MBs are selected to execute the rate control algorithm. A modified TMN-8 with ZRPS is implemented for H.263-to-H.263 video transcoder. The experimental results can be used to verify the effectiveness of the ZRPS mechanism. Indeed, ZRPS mechanism can be applied to most of the existing rate control algorithm to reduce the amount of MBs needed to be processed and hence speed up the algorithm. In comparison with TMN-8, the sequences coded with ZRPS-TMN-8 can achieve similar or higher visual quality and PSNR and require a significantly lower computational complexity.

Acknowledgment

This work is supported, in part, by Innovation Technology Commission (ITS/122/03) and Research Grant Council (HKUST6156/03E) in Hong Kong.

TABLE III
COMPARISON OF PERFORMANCE ACHIEVED BY TMN-8 AND THE PROPOSED ZRPS-TMN-8 WHEN BIT-RATE CONVERSION FROM 384KBPS TO 64KBPS

Name	PSNR (dB)			Encoded Frame		Speed
	TMN-8	Proposed	Gain	TMN-8	Proposed	
akiyo	39.22	39.63	+0.41	292	295	+3.70
children	26.40	26.60	+0.20	276	276	+3.33
coastguard	28.38	28.54	+0.16	295	295	+2.17
container	34.01	34.55	+0.54	292	293	+3.13
foreman	29.74	29.87	+0.13	283	281	+2.08
hall monitor	35.06	36.14	+1.08	294	294	+4.17
mobile	23.07	23.08	+0.01	268	266	+1.69
m&d	37.00	37.20	+0.20	296	296	+2.38
sean	35.26	36.30	+1.04	291	295	+3.70
silent voice	32.71	33.24	+0.53	295	295	+2.94
stefan	24.12	24.15	+0.03	205	205	+2.04
table	30.74	31.34	+0.60	273	273	+3.57
weather	29.33	30.93	+1.60	284	283	+4.55
Average	**31.16**	**31.66**	**+0.50**	**280.31**	**280.54**	**+3.03**

TABLE IV
COMPARISON OF PERFORMANCE ACHIEVED BY TMN-8 AND THE PROPOSED ZRPS-TMN-8 WHEN BIT-RATE CONVERSION FROM 384KBPS TO 96KBPS

Name	PSNR (dB)			Encoded Frame		Speed
	TMN-8	Proposed	Gain	TMN-8	Proposed	
akiyo	41.31	41.52	+0.21	297	296	+3.23
children	27.98	28.61	+0.63	293	294	+2.94
coastguard	30.26	30.30	+0.04	297	297	+1.64
container	35.91	36.28	+0.37	295	296	+2.27
foreman	31.67	31.73	+0.06	294	294	+1.59
hall monitor	37.59	37.86	+0.27	296	296	+2.63
mobile	24.15	24.15	+0.00	292	292	+1.41
m&d	38.80	38.89	+0.09	298	298	+1.96
sean	38.14	38.73	+0.59	295	297	+3.03
silent voice	35.01	35.42	+0.41	297	297	+2.38
stefan	25.07	25.14	+0.07	262	261	+1.72
table	32.52	33.09	+0.57	294	295	+2.70
weather	32.01	33.43	+1.42	290	288	+3.85
Average	**33.11**	**33.47**	**+0.36**	**292.31**	**292.38**	**+2.41**

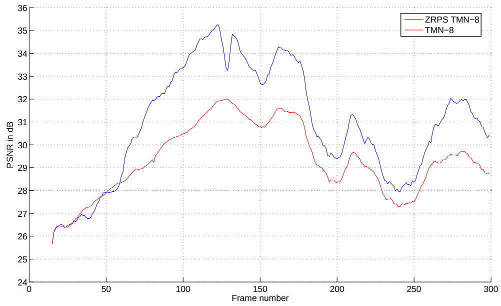

Fig. 1. The PSNR of test sequence ’weather’ converted from 384kbps to 64kbps with +1.60 dB PSNR gain

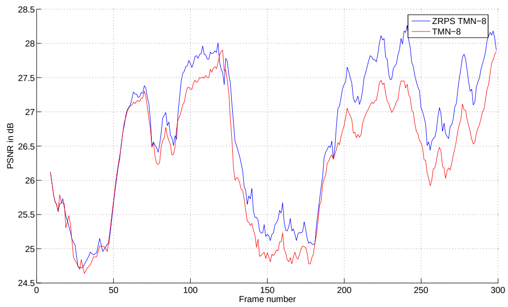

Fig. 2. The PSNR of test sequence ’children’ converted from 384kbps to 64kbps with +0.20 dB PSNR gain

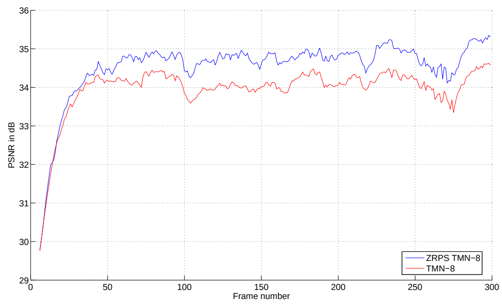

Fig. 3. The PSNR of test sequence ’container’ converted from 384kbps to 64kbps with +0.54 dB PSNR gain

REFERENCES

[1] J. Xin, C.-W. Lin, and M.-T. Sun, “Digital video transcoding,” in *Proceedings of the IEEE*, vol. 93, no. 1, Jan. 2005, pp. 84–97.

[2] A. Vetro, C. Christopoulos, and H. Sun, “Video transcoding architectures and techniques: an overview,” *IEEE Signal Processing Magazine*, vol. 20, no. 2, pp. 18–29, Mar. 2003.

[3] H. Kasai, T. Hanamura, W. Kamayama, and H. Tominaga, “Rate control scheme for low-delay MPEG-2 video transcoder,” in *Proceedings of the IEEE International Conference on Image Processing*, vol. 1, Sept. 2000, pp. 964–967.

[4] W.-N. Lie, M.-L. Tsai, and T. C. L. Lin, “Rate-distortion optimized DCT-domain video transcoder for bit-rate reduction of MPEG videos,” in *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, May 2004, pp. 969–972.

[5] Z. Lei and N. D. Georganas, “Accurate bit allocation and rate control for DCT domain video transcoding,” in *IEEE Canadian Conference on Electrical and Computer Engineering*, vol. 2, May 2002, pp. 968–973.

[6] ——, “Rate adaptation transcoding for precoded video streams,” in *Proceedings of the tenth ACM International Conference on Multimedia*, no. 4, Dec. 2002, pp. 127–136.

[7] B. Xie and W. Zeng, “Sequence-based rate control for constant quality video,” in *Proceedings of the IEEE International Conference on Image Processing*, vol. 1, Sept. 2002, pp. 77–80.

[8] Y. Sun, X. Wei, and I. Ahmad, “Low delay rate-control in video transcoding,” in *Proceedings of the IEEE International Symposium on Circuits and Systems*, vol. 2, May 2003, pp. 660–663.

[9] K.-D. Seo, S.-H. Lee, J.-K. Kim, and J.-S. Koh, “Rate control algorithm for fast bit-rate conversion transcoding,” *IEEE Transactions on Consumer Electronics*, vol. 46, no. 4, pp. 1128–1136, Nov. 2000.

[10] *Test Model 5*, ISO-IEC/JTC1/SC29/WG11, Apr. 1993, Document AVC-491b, Document 2.

[11] J. Ribas-Corbera and S. Lei, “Rate control in DCT video coding for low-delay communications,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 1, pp. 172–185, Feb. 1999.

[12] Image Processing Lab, University of British Columbia, “TMN (H.263+) encoder/decoder, version 3.2,” September 1997. [Online]. Available: www.ee.ubc.ca/image