

# AN IMPROVED MACROBLOCK LEVEL RATE CONTROL ALGORITHM FOR MPEG-4 VIDEO OBJECT CODING

Qing Wu and Shing-Chow Chan

Department of Electrical and Electronic Engineering, The University of Hong Kong.

Emails: qingwu@eee.hku.hk, scchan@eee.hku.hk

## ABSTRACT

This paper presents an improved macroblock (MB) level rate control algorithm for MPEG-4 object coding. Motivated by the statistical analysis of the coding of MBs and VOPs for typical sequences, a refined quadratic rate-distortion model is proposed in the new MB rate control algorithm. The basic idea is to exploit the statistical coding property of a group of MBs coded by an identical QP rather than an individual MB to do the modeling, thereby producing more stable quantization parameter to improve the estimation and coding performances. In addition, some improved measures which explore the new model for coding the MBs in a VOP are incorporated. Simulation results show that the proposed algorithm can achieve the much smoother bit rate and better picture quality in terms of PSNR than the MPEG-4 VM18 algorithm for the tested sequences.

## 1. INTRODUCTION

Rate control [1, 2, 3] plays a crucial role in controlling the output bit rate and video quality of video encoders. In frame-based video coding standards such as MPEG-1/2, the rate control algorithm is usually divided into two major levels, the frame level and the macroblock (MB) level. Whereas, for object-based coding standard such as MPEG-4 [4], an additional object level rate control is necessary for allocating available bits among multiple video objects (VOs) [5]. At the frame or object level, the rate control algorithm determines the corresponding quantization parameter (QP) and estimates the bits required for the video frame or object to be encoded. At the MB level, the QP needs to be adjusted further by the rate control algorithm in order to meet accurately the given target bits, which are assigned at the frame or object level. MB rate control is thus very important as it eventually determines whether the target bits will be met and hence the final video quality. An example of such a MB rate control algorithm is the one proposed in MPEG-4 video verification model (VM) v18.0 [6], which adjusts the QP value for each MB within a Video Object Plane (VOP).

In this paper, an improved MB rate control algorithm for MPEG-4 video object coding is proposed. The new MB rate control algorithm employs a refined quadratic rate-distortion model, which is derived from the one proposed in [2, 3]. The model was developed by analyzing the statistical properties of the bit rates for coding MBs and VOPs. It exploits the statistical coding property of a group of MBs encoded by an identical QP, instead of an individual MB, for rate allocation. In particular, the encoded MBs are classified and assembled into an array of different observation data points in terms of the QPs used, and then refined and rearranged before being fed into the linear regression to accurately update the model parameters. Furthermore, a useful measure to

precisely compute the estimated target bits for next MB to be coded is also adopted. Experimental results show that the proposed MB rate control algorithm significantly outperforms the MPEG-4 VM18 algorithm in achieving the target bit rates of VOPs. Besides, considerable better picture quality over VM18 in terms of PSNR for a number of test sequences is obtained. It is mainly attributed to the reduced overhead resulting from the more accurate and stable QP values for the MBs of VOPs produced by our algorithm.

The rest of the paper is organized as follows. Section 2 reviews the MB rate control algorithm for object coding in MPEG-4 VM18. The improved MB rate control for MPEG-4 object coding is presented in Section 3. Experimental results are shown in Section 4 and finally, conclusions are given in Section 5.

## 2. MB RATE CONTROL ALGORITHM IN MPEG-4 VM18

### 2.1 Rate-Distortion Model

The MB rate control in MPEG-4 VM18 [6] employs two different rate-distortion models, which are used respectively for high bit rate and the low bit rate situations. Let  $B_i$  denote the number of bits produced by the  $i$ -th MB of a VOP, then it can be modeled as follows:

$$\text{if (TargetBitRate} > R) \\ B_i = (A_1 / QP_i^2) \cdot MAD_i^2; \tag{1}$$

$$\text{Else} \\ B_i = (A_2 / QP_i^2) \cdot MAD_i + (A_3 / QP_i) \cdot MAD_i; \tag{2}$$

where  $A_1$ ,  $A_2$  and  $A_3$  are the model parameters,  $MAD_i$  is the Mean Absolute Difference of the  $i$ -th MB, and  $R$  is a rate threshold for deciding which one of the two models is to be used. By default,  $R=0.085$ . For (2), given the values of  $A_2$  and  $A_3$ , if the corresponding QP cannot be found at a certain value of  $B_i$ , then the following linear model is used instead, where  $A_{3\_x11}$  is the corresponding model parameter:

$$B_i = (A_{3\_x11} / QP_i) \cdot MAD_i. \tag{3}$$

### 2.2 Target Number Of Bits For Next MB To Be Encoded

Let  $N$  be the number of MBs in a VOP. The target number of bits,  $T_i$ , for the  $i$ -th MB (the next MB to be encoded), is given by:

$$T_i = (W_i \cdot MAD_i) \cdot T_{\text{TEX}} / \sum_{j=1}^N (W_j \cdot MAD_j). \tag{4}$$

where  $W_j$  is a weight indicating the perceptual importance of the  $i$ -th MB. By default,  $W_1=W_2= \dots =W_N=1$ .  $T_{\text{TEX}}$  is the target number of bits for the texture of remaining MBs in the current VOP, which is the target VOP bits  $T$  from the object level rate control minus the bits for coding the texture of the encoded MBs in the current VOP and the header bits for motion, shape, and syntax overhead. The header bits usually come from the previously encoded VOP.

## 2.3 MB Rate Control Algorithm

The algorithm is described using C-like programming language, and the contents following “//” are the related comments.

### Step 1. Initialization:

```

if (current VOP == the first VOP of the VO)
{A1 = A1_prev = 100; A2 = A2_prev = 400; A3 = A3_prev = 0;}
else {A1 = A1_prev; A2 = A2_prev; A3 = A3_prev;}
if (TargetBitRate > R) K = HIGH_MODE;
else K = LOW_MODE;
i = 1; // i is MB index;
skip = 0; // count the number of MBs coded with 0 bit.
S1 = (∑j=1..N Wj · MADj); // By default, W1=W2=...=WN=1.

```

### Step 2. Compute QP for i-th MB:

```

Ti = (Wi · MADi / Si) · TTEX;
if (K == HIGH_MODE) QP = MADi · SQRT(A1/Ti);
else {
QP = Q : (A2 / Q2) · MADi + (A3 / Q) · MADi = Ti;
if ((A3 · MADi)2 + 4 · Ti · A2 · MADi) < 0)
QP = (MADi · A3_x11) / Ti; // using linear model instead.
}
Round QP to value in {1,...,31}, and make sure the difference between
the new QP and that of previous MB is in DQUANT={-2, -1, 0, 1, 2}.

```

### Step 3. Encode MB with QP;

### Step 4. Update Counters:

```

Let Bi' is the actual encoded bits of i-th MB, then
TTEX = TTEX - Bi'; Si+1 = Si - Wi · MADi;

```

### Step 5. Update A1, or A2, A3, A3\_x11:

```

if (Bi' == 0)
{ skip = skip + 1; goto step 6; }
else { // do updating;
if (K == HIGH_MODE) {
A1' = (Bi' / MADi2) · QPi2;
Let A1' = A1' / (i-skip) + A1' · (i-skip-1) / i;
A1 = A1' · i / N + A1_prev · (N-i) / N;
}
else { // K == LOW_MODE;
// using linear regression to update the model parameters.
From the last M (sliding window size) encoded MBs with
non-zero bits (Bi' > 0), three sets are created: {b1, b2, ..., bM},
{q1, q2, ..., qM}, and {m1, m2, ..., mM}, where the bi's, qi's
and mi's correspond to the values of Bi', QPi, and MADi. Let
M not be greater than 20.
// update model parameters A2, A3 and A3_x11 as follows:

$$A_2 = \frac{M \cdot \sum_{i=1}^M (b_i/m_i) - \sum_{i=1}^M (q_i^{-1}) \cdot \sum_{i=1}^M (q_i \cdot b_i/m_i)}{M \cdot \sum_{i=1}^M (q_i^{-2}) - \left( \sum_{i=1}^M (q_i^{-1}) \right)^2};$$


$$A_3 = \frac{\sum_{i=1}^M (q_i^{-2}) \cdot \sum_{i=1}^M (q_i \cdot b_i/m_i) - \sum_{i=1}^M (q_i^{-1}) \cdot \sum_{i=1}^M (b_i/m_i)}{M \cdot \sum_{i=1}^M (q_i^{-2}) - \left( \sum_{i=1}^M (q_i^{-1}) \right)^2};$$


$$A_{3\_x11} = \left( \sum_{i=1}^M (q_i \cdot b_i/m_i) \right) / M;$$

} // end of “else { // K == LOW_MODE;”
} // end of “else { // do updating;”

```

### Step 6. Termination Decision:

```

if (i == N) // all MBs are encoded.
{ A1_prev = A1; A2_prev = A2; A3_prev = A3; stop algorithm; }
else { i = i + 1; goto Step 2; }

```

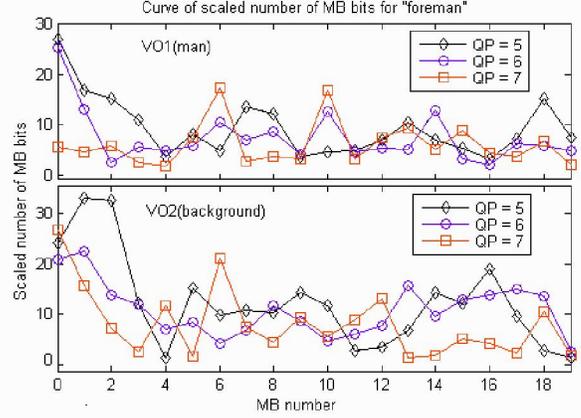


Fig. 1. Scaled numbers of MB bits at different constant QPs for “foreman”.

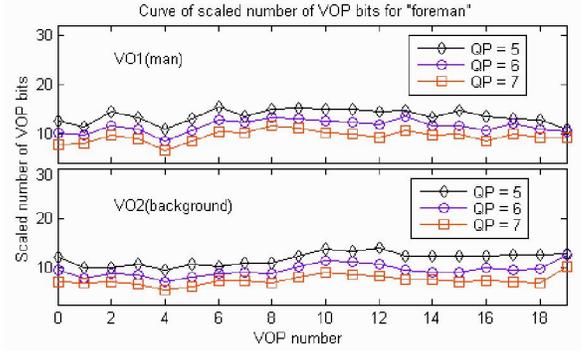


Fig. 2. Scaled numbers of VOP bits at different constant QPs for “foreman”.

## 3. IMPROVED MB RATE CONTROL

### 3.1 Statistical Analysis Of The Coding of MBs and VOPs

The main problem in a rate control algorithm is to model accurately the bit required for encoding a MB due to the complicated quantization and variable length coding (VLC) processes. In the MB rate control model of MPEG-4 VM18, the number of bits required for an MB is assumed to be proportional to the variance of the MB in the form of MAD (cf. low rate model (2) and linear model (3)) or MAD<sup>2</sup> (high rate model (1)), and it decreases with increasing values of QP. Fig. 1 shows typical curves of number of bits (scaled) for 20 consecutive MBs from a P-VOP. Each curve corresponds to a constant QP, and the corresponding group of MBs belong to the VOPs of VO1 (man) or VO2 (background) of the “foreman” sequence. The scaled number of bits for a MB is equal to the number of encoded bits divided by the MB MAD. It can be seen that the scaled number of bits varies violently with different MBs. It suggests that the VM models cannot properly model the required bits at the MB level. Fig. 2 depicts typical curves of scaled number of bits for 20 consecutive P-VOPs belonging to VO1/VO2 of “foreman”, and each one is plotted at a constant QP. It can be seen that the scaled number of VOP bits, which is equal to the encoded VOP bits divided by the sum of MADs of all MBs in the VOP, appear to be much smoother and flatter than those in Fig. 1. More importantly, these curves are separated clearly with different QPs. Similar observations are obtained for other sequences such as “news”, “coastguard”, “container” and so on. It must be noted that all the MBs used in Fig. 1 and Fig. 2 are encoded with non-zero bits. Actually, many MBs of VOPs with small MADs are coded with zero bits, especially for those MBs at the VOP boundaries. For

instance, the percentages of zero-bit MBs encoded at  $QP = 5$  for VO1 and VO2 of “foreman” are 31.8% and 19.9%, respectively.

### 3.2 Improved MB Rate Control Algorithm

The objective of the proposed MB rate control algorithm is to make the QPs of the MBs vary smoothly around the mean QP of the entire VOP, and to meet the VOP target bits as much as possible.

#### 3.2.1. Rate-distortion model

From the study of the MB/VOP coding statistics in the previous section, it can be observed that for a set of MBs (e.g., all MBs in a VOP), the scaled sum of encoded bits for these MBs can also be accurately model by the quadratic rate-distortion model. Therefore, a quadratic model similar to (2) is adopted in the proposed MB rate control algorithm, except that a group of MBs encoded by an identical QP are considered in the proposed algorithm instead of an individual MB.

Let  $Rs_n$  be the scaled number of bits produced by encoding a group of MBs with an identical  $QP_n$  which takes value in the QP level set  $Q = \{1, 2, 3, \dots, 31\}$  (namely,  $QP_n = n, n \in Q$ ), then the rate-distortion model of the new MB rate control algorithm can be given by

$$Rs_n = (SB_n / SMAD_n) = (A_2 / QP_n^2) + (A_3 / QP_n), \quad (5)$$

where  $SB_n$  is the sum of bits produced for that group of MBs,  $SMAD_n$  is the sum of MADs (weighted with perceptual importance  $W_j$ ) for that group of MBs,  $A_2$  and  $A_3$  are the model parameters.

The linear model, which is used when the expected QP for a MB cannot be found through (5), is correspondingly adjusted by

$$Rs = (SB / SMAD) = A_{3\_x11} / \overline{QP}, \quad (6)$$

where  $Rs$  is the scaled sum of bits for those encoded MBs in the current VOP,  $\overline{QP}$  is the mean value of QPs for those encoded MBs in the current VOP, and  $A_{3\_x11}$  is a model parameter.

#### 3.2.2. Target number of bits for next MB to be encoded

The scaled target number of bits for the  $i$ -th MB (the next MB to be encoded),  $Ts_i$ , is

$$Ts_i = T_i / (W_i \cdot MAD_i) = T_{\text{TEX}} / \sum_{j=1}^N (W_j \cdot MAD_j). \quad (7)$$

It is actually another form of (4). We can also see from (7) that the scaled target number of bits for next MB to be encoded is equal to the scaled remaining target number of bits for those remaining MBs to be encoded in the VOP. This is exactly in agreement with the statistical analysis in section 3.1 when the model in (5) is employed.

#### 3.2.3. MB rate control algorithm

The proposed MB rate control algorithm is implemented in steps similar to the MPEG-4 VM18 in Section 2.3, except that several improved measures are incorporated in the refined model, as highlighted below.

- a) The improved quadratic model is applied for different target bit rates, no matter whether they are high or low.
- b) For updating the model parameters  $A_2$ ,  $A_3$  and  $A_{3\_x11}$ , a pool is used to store a certain amount (equal to the number of MBs in the current VOP) of encoded MBs.
- c) Among the encoded MBs employed to update the parameters  $A_2$  and  $A_3$  using linear regression, at most 31 different QPs may be used, each corresponding to one quantization level of  $\{1, 2, \dots, 31\}$  and an observation data point. To update the model parameters more accurately in a relatively small range of QP

centering around the mean QP of the entire VOP, we first compute  $\overline{QP}$ , the mean value of QPs for those encoded MBs in the current VOP. Then, from all of the observation data points we select five (if more than five points are available) ones whose QPs center around  $\overline{QP}$  in value, as the final observation data points to be fed into the linear regression. This means that the sliding window size  $M$  in step 5 is not greater than five.

- d) In a VOP, some of the MBs may be encoded with zero bits. To accurately compute the target bits for next MB to be encoded using (7), we estimate the sum of MADs for those MBs to be encoded with zero bits, and subtract them from the sum of MADs for the remaining MBs. For this, a threshold of MAD is calculated from the previously encoded VOP such that most of the MBs with the MAD lower than this threshold are encoded with zero bits. Then this threshold is used for the current VOP to estimate and compute the effective sum of MADs.

## 4. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed MB rate control algorithm and compare it with MPEG-4 VM18 algorithm, both the frame level and object level rate control algorithms proposed in VM18 are employed in our encoder. The VOPs of the first frame in a test sequence are encoded as I-VOPs, while the VOPs in the following frames are encoded as P-VOPs. All the perceptual importance weights ( $W_j$ ) for the MBs are set to 1. The test CIF sequences include “foreman” and “news”, whose frame rates are both 24 fps. Figs. 3 to 6 depict the encoded bits for “foreman” and “news” at both high and low target bit rates using the proposed MB rate control algorithm and the VM18 algorithm. It can be seen that the proposed MB rate control algorithm can achieve a much smoother and closer bits rates to the target for each VO and entire video than that using VM18. From Fig. 4 and Fig. 6, it can be seen that the performance improvements at low bit rates is less significant compared with those at high bit rate. It is because the available bits, i.e., the bits available for texture coding, are considerably less than the header bits including shape, MV and syntax overheads at low bit rates. A detailed summary of the simulation results are provided in Table 1. It can be seen that the picture quality in terms of PSNR using the proposed MB rate control algorithm is considerably better than the VM18 algorithm. Furthermore, the QP values of the MBs in a VOP produced by our algorithm are more stable than those in the VM18 algorithm. As a result, less overhead bits are required to specify the change of QPs during rate control, leading to better coding efficiency.

## 5. CONCLUSION

An improved MB level rate control algorithm for MPEG-4 video object coding is presented. It uses a refined quadratic rate-distortion model, which is motivated by the statistical analysis of the coding of MBs and VOPs for typical sequences. The statistical coding property of a group of MBs encoded by an identical QP instead of an individual MB are studied and exploited in the new algorithm to produce more stable quantization parameters, so as to improve the estimation and coding performances. Some other improved measures are also incorporated into the new algorithm. Simulation results show that the proposed algorithm can achieve a much smoother bit rate and better picture quality than those using the MPEG-4 VM18 algorithm for the tested sequences.

## REFERENCES

[1] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 9, pp. 172-185, Feb. 1999.

[2] T. Chiang and Y.-Q. Zhang, "A new rate control scheme using quadratic rate-distortion modeling," *IEEE Trans. on Circuits and System for Video Technology*, vol. 7, pp. 246-250, Feb. 1997.

[3] H. J. Lee, T. Chiang, and Y.-Q. Zhang, "Scalable rate control for MPEG-4 video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol.

10, pp. 878-894, Sept. 2000.

[4] ITU-T Recommendation ISO/IEC 14496-2:2001, "Information Technology- Coding of audio-visual objects -- Part 2: Visual".

[5] A. Vetro, H. Sun, and Y. Wang, "MPEG-4 rate control for multiple video objects," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 9, pp. 186-199, Feb. 1999.

[6] MPEG-4 video verification model v18.0, ISO/IECJTC1/SC19/WG11 Coding of Moving Pictures and Audio N3908, Pisa, Jan. 2001.

Table 1. Comparison of simulation results for proposed MB rate control algorithm and VM18 algorithm.

Test Sequence	Video Bit Rate (bps)	VO	Average Bit Deviation (bits)			PSNR(dB)			Average QP Deviation	
			VM18	Proposed	Improvement	VM18	Proposed	Gain	VM18	Proposed
Foreman	$5.8 \times 10^5$	Entire video	1340.5	740.6	44.75%	35.31	35.98	+0.67	1.67	0.69
		VO1	1058.6	516.2	51.24%	35.18	36.19	+1.01	3.59	0.87
		VO2	587.1	513.0	12.62%	35.41	35.89	+0.48	1.69	0.69
Foreman	$3.2 \times 10^5$	Entire video	936.4	762.6	18.56%	33.34	33.73	+0.39	4.14	2.12
		VO1	555.0	422.2	23.93%	33.53	33.86	+0.33	6.29	2.29
		VO2	584.8	530.8	9.23%	33.25	33.68	+0.43	4.43	2.23
News	$3.6 \times 10^5$	Entire video	1769.3	636.5	64.03%	37.89	39.05	+1.16	2.95	0.96
		VO1	1345.0	309.2	77.01%	36.57	37.57	+1.00	4.10	1.21
		VO2	1120.7	484.8	56.74%	39.02	40.25	+1.23	4.20	1.00
News	$1.6 \times 10^5$	Entire video	955.6	671.7	29.71%	34.92	35.45	+0.53	3.66	2.98
		VO1	542.7	269.5	50.34%	33.37	34.08	+0.71	3.87	3.19
		VO2	691.5	510.8	26.13%	36.18	36.52	+0.34	3.83	3.05

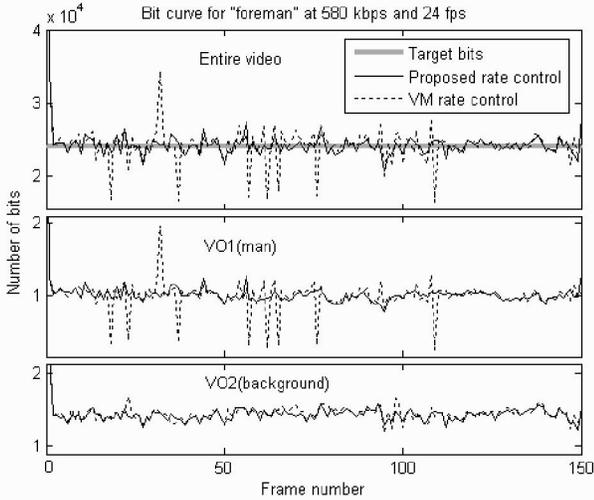


Fig. 3. Comparison of encoded bits at high target bit rate for "foreman".

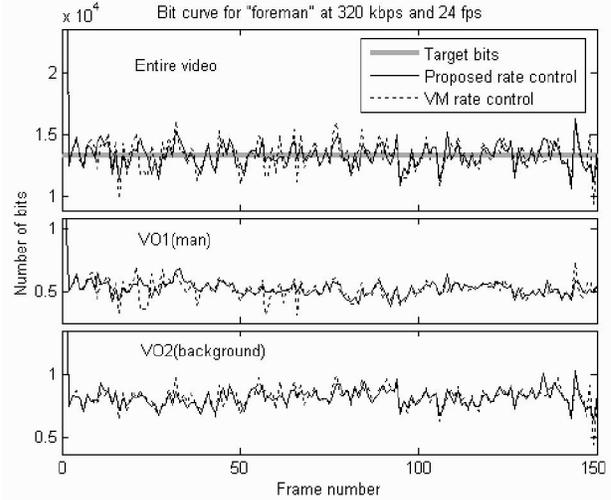


Fig. 4. Comparison of encoded bits at low target bit rate for "foreman".

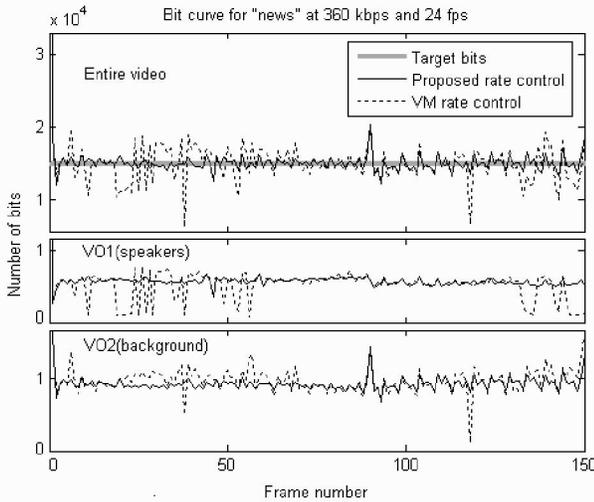


Fig. 5. Comparison of encoded bits at high target bit rate for "news".

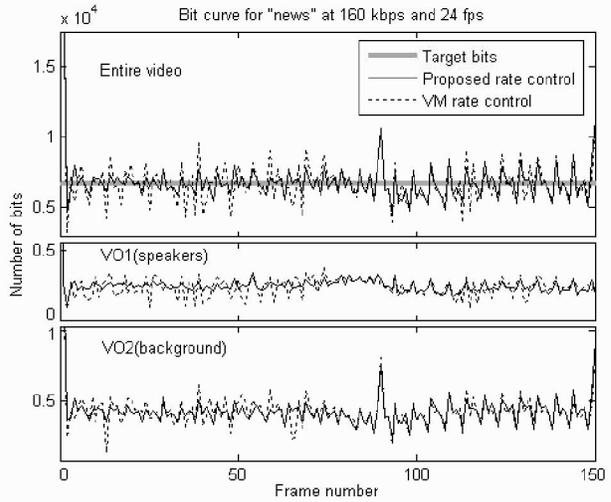


Fig. 6. Comparison of encoded bits at low target bit rate for "news".