Neighbor-based Fast Rate Control Scheme for Basic Unit Layer for H.264/AVC Standard

Bruno B. Vizzotto, Cláudio M. Diniz, Sergio Bampi
PPGC – Informatics Institute (INF)
Federal University of Rio Grande do Sul (UFRGS)
Porto Alegre – RS, Brazil
{bbvizzotto, cmdiniz, bampi}@inf.ufrgs.br

ABSTRACT

In video encoding H.264/AVC is the state-of-the-art and provides efficient compression for video applications. However, by using the rate control scheme of reference H.264/AVC software, high variations in the visual image quality will be experienced. This paper presents a novel rate control scheme for the H.264/AVC standard in order to ensure the output video quality and maintain a stable output bit-stream. First, an overview of the rate control module is presented. Then a neighbor-fast based model to predict the mean of absolute difference of basic unit layer is proposed. Finally, a comparison between compromise and visual quality with related work described in the literature will be performed.

Categories and Subject Descriptors

I.4.2 [**Image Processing and Computer Vision**]: Compression (Coding) – *Approximate methods*.

General Terms

Algorithms, Performance, Design.

Keywords

Video Encoding, Rate Control, H.264/AVC

1. INTRODUCTION

H.264/AVC [1] is the state-of-the-art video coding standard defined in 2003 by Joint Video Team (JVT). It improves the compression rates over all previous standards by incorporating very efficient coding tools. However, it increased the computational complexity of encoder algorithms compared with previous standards. H.264/AVC is used for video encoding on the Brazilian Digital TV Broadcast System (SBTVD).

In H.264/AVC encoding, the video is partitioned into frames and each frame is partitioned into the basic coding unit that is a macroblock (MB). In turn, each macroblock is transformed by a sequence of steps (prediction, quantization, entropy coding, etc.). These steps results in a variable bit-rate.

Due to the bandwidth constraints imposed by current broadcast communication technology, a maximum bit-rate is defined for the transmission of a video through a communication channel. Thus, the H.264/AVC standard defines a set of levels which imposes restrictions on the rate of output bits of the encoded video. Among these different levels, level 4 is used for encoding high-resolution video (e.g. HD1080, with 1920x1080 pixels) and restricts in 20 Mbps the maximum bit-rate of output coded video.

Therefore, the Rate Control (RC) module present in video encoders aims to restrict volume of data in the output coded video buffer and bandwidth of a given transmission channel [2]. Quantization Parameter (QP) is used in video encoding to control the intensity which the quantization (a process which inserts loss in coding) will be applied to the residue of the prediction. One simple way to control the rate is to maintain QP at a fixed value in which the maximum bandwidth is never reached. However, using this technique, there may be an under-utilization of available bandwidth to transmit video. It is necessary to accurately calculate the QP, dynamically, to avoid this kind of loss. A rate control algorithm aims to dynamically control QP to keep the bit-rate below the maximum required level. Thus, typical video encoder uses a rate control block as a method of regulating the bit-rate variation in order to obtain high quality coding to a given target. Due to its importance, rate control has become a fundamental process in an encoder, and has been widely exploited in video coding standards, such as MPEG-2, MPEG-4, H.263, and H. 264/AVC.

Generally, we can divide the function of RC module in two main procedures: ***(i)*** Bit allocation, which distributes the limit of bandwidth for each basic unit (BU), which can be a group of pictures (GOP), a frame or a macroblock (MB). ***(ii)*** QP values calculation, which is responsible to control the bit production of each basic unit.

The main difference between the rate control block adopted in recent video coding standards, such as H.264/AVC, compared to the previous standards, is the use of the QP as input to the rate distortion optimization (RDO) module [3]. The rate control scheme implemented in H.264/AVC JM reference software [4] performs the RDO for MB of the current frame in such a way that QP must be determined in advance, using the mean of absolute differences (MAD) of the MB of current frame. However, the MAD of current frame or MB is only available after performing the RDO. So, [4] uses a quadratic rate-distortion model to predict the corresponding QP, which is then used for the RDO for each MB in the current basic unit.

In order to minimize loss of quality in digital video, this paper proposes a novel rate control scheme named NFS-RC (Neighbor-based Fast Scheme Rate Control). This proposal aims to reduce the QP difference between neighboring BUs and to avoid the use of a quadratic model to ensure a fast MAD prediction. Therefore, the MAD will be predicted by a weighted model, as alternative to the linear model [6]. Our model adopts constant weight parameters, reducing the dynamic update and reducing the

---

WebMedia'11: Proceedings of the 17th Brazilian Symposium on Multimedia and the Web. Short Papers.
October 3 -6, 2011, Florianópolis, SC, Brazil.
ISSN 2175-9650.
SBC - Brazilian Computer Society

complexity. The algorithm is implemented by modifying the JM 17.2 software [5].

2. RATE CONTROL ALGORITHM

In the reference software [5], a linear model is used to predict the MAD of current BU. This prediction is given from the BU in the related position in the reference frame [2] and with some variables update by rate distortion quadratic model.

The rate control algorithm for the H.264/AVC standard is divided into three main levels: ***(i)*** GOP layer, ***(ii)*** frame layer, and ***(iii)*** basic unit layer. These layers make the use of rate control scheme described below to a target bit-rate. First, the MAD calculation is based on the equation defined in (1). In turn, the quantization parameter is adjusted to maintain visual quality of the video according to the bit-rate.

2.1 GOP Layer Rate Control

In the GOP layer rate control there are two main tasks. The first task is the allocation of target bit-rate of current GOP in accordance with the following parameters: channel bandwidth, current frame rate and number of frames in GOP. Based on these values, the second task aims to define the QP value to next iteration of the GOP layer.

In the reference software, the initial quantization parameter QP_0 can be defined by the user and this will be the first QP of the first frame of the GOP. I frame and first P frame of GOP will be based on QP_0 value. Then, the larger is the available channel bandwidth, the lower is the QP_0 and vice versa. Thus, the bandwidth is regulated to QP_0 decrease by one when GOP length is increased by 15.

2.2 Frame Layer Rate Control

In the frame layer rate control there are also two main tasks. First task calculates QP for each frame. The second task inputs QP to the RDO module.

Considering that B type frames are not used as reference for other types of frames (which is the case of I and P frames) and in order to preserve bits to encode I and P frame types, larger value of QP are assigned for I and P frames and lower QP values are assigned for B type frames. However, to ensure the minimum disparity in the visual quality of encoded video, the QP difference between two adjacent B frames cannot be greater than two.

In turn, the QP value calculation for P frames is in accordance with the following parameters: buffer capacity, frame rate, available bandwidth channel and current occupation of the buffer. Thus, the obtained QP can be used earlier for the RDO process of each MB of the current frame.

2.3 Basic Unit Layer Rate Control

In the BU layer RC, three steps are performed: ***(i)*** prediction of MAD value, ***(ii)*** objective bits allocation ***(iii)*** QP value of current BU calculation.

In the first step, the MAD value prediction presented by [4] is calculated using a linear model proposed to solve the "chicken and egg" dilemma (QP and RDO interdependence). This model is based on linear tracking theory [6]. Thus, the MAD prediction of BU of the current frame is calculated according to the MAD of the previous position related.

Assuming that MAD_{cb} denotes the MAD of the current frame of BU and MAD_{pb} denotes the MAD of BU of the reference frame, the linear prediction model is given by (1).

$$MAD_{cb} = a_1 * MAD_{pb} + a_2 \tag{1}$$

Where, a_1 and a_2 are coefficients of the prediction model with initial values in 1 and 0 respectively. They are updated every encoding of each BU by linear regression.

The prediction model requires the update of the parameters a_1 and a_2 in real time. The accuracy of prediction model is ensured by few points of input, thereby at least ten points on a coordinate plan are allocated to determine these coefficients. After allocating, the points with great error are removed, thus generating a plan with points to a_1 and a_2 determination. This procedure has a quadratic computational cost.

In the second step, the objective bits allocation includes allocation of texture bits of the current basic unit in accordance with MAD prediction done in the first step, which includes the calculation of the bits of header of BU, where texture bits of BU are the difference between objective and bits of header according (2).

$$R_i = \frac{f_{rb}}{N_{ub}} - m_h \tag{2}$$

In equation (2), R_i denotes the bits of texture BU, f_{rb} denotes residual of all the BU bits not encoded in the current frame, N_{ub} denotes the number of BUs not encoded and finally m_h denotes the bits of header information of the BU.

In the third step the QP calculation is performed using the quadratic rate distortion model. This calculation is given by (3).

$$R_i(j) = X_1 \times \frac{MAD_i(j)}{Q_{step,i}(j)} + X_2 \times \frac{MAD_i(j)}{Q^2_{step,i}(j)} - m_h(j) \tag{3}$$

In (3), X_1 and X_2 are parameters of quadratic rate distortion model.

3. PROPOSED NEIGHBOR-BASED FAST RATE CONTROL SCHEME

The algorithm that can be considered optimal for rate control is impractical, because it considers the variation of the output bit-rate of an MB of encoded video to determine the rate control scheme to encode this MB [7]. Therefore, we propose a novel scheme, called *Neighbor-based Fast Rate Control Scheme* (NFS-RC), based on fast MAD prediction using neighbor BU data.

3.1 Weighted Neighbor-based MAD Prediction

In order to predict the MAD value for each BU, we analyzed MAD correlation on two inherent characteristics of video sequences: temporal and spatial redundancy. Temporal redundancy, a.k.a. *inter-frame redundancy*, is the correlation between neighbor frames. Spatial redundancy, a.k.a. *intra-frame redundancy*, is the correlation between neighbor pixels.

Observing the MAD correlation presented in [8] we selected four neighbors for MAD prediction, as shown in blue in figure 1. For each neighbor, a constant weight is assigned. Based on off-line analysis of various video sequences, the most correlated MAD is of the one of BU on the same position of previous frame (linked by dotted line in figure 1). For this MAD is assigned the higher weight, corresponding at least 82 percent of prediction. Lower weights are assigned for other BUs according to equations (4), (5) and (6).

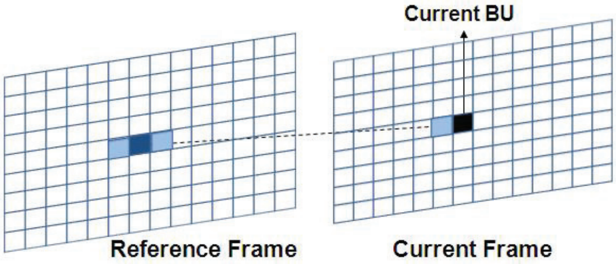

Figure 1 - Proposed NFS-RC Algorithm weighted Model

Then, the MAD of current BU is obtained by a weighted sum given by (4).

$$MAD_{c(i)} = (1-6\alpha)*MAD_{p(i)} + 4\alpha * MAD_{c(i-1)} + \alpha * MAD_{p(i-1)} + \alpha * MAD_{p(i+1)} \quad (4)$$

In the case the current BU is the first of the current frame, left neighbor BU is not available. Then the calculation is simplified by the equation (5).

$$MAD_{c(i)} = (1-2\alpha)*MAD_{p(i)} + 2\alpha * MAD_{p(i+1)} \quad (5)$$

In the case the current BU is the last of the current frame, the right neighbor of co-located MB is not available, so we simplified by the equation (6).

$$MAD_{c(i)} = (1-5\alpha)*MAD_{p(i)} + 3\alpha * MAD_{c(i-1)} + 2\alpha * MAD_{p(i-1)} \quad (6)$$

In (4), (5) and (6), $MAD_{c(i)}$ denotes the MAD of current BU, $MAD_{c(i-1)}$ is the MAD of previous BU, $MAD_{p(i)}$ is the MAD of co-located BU in the previous (reference) frame, and $MAD_{p(i-1)}$ and $MAD_{p(i+1)}$ are respectively the MAD placed after and before the co-located BU in the previous (reference) frame. Finally, α is a constant with value 0.03, obtained empirically through experiments. Note, our weighted neighbor-based MAD prediction is based only in one parameter (MAD) and a constant (α) which eliminates at all the quadratic model for MAD update.

3.2 Proposed NFS-RC Algorithm

The NFS-RC model considering the BU size ranging from MB to frame size is presented by pseudo-code below in figure 2, where *w_size* denotes the size of BU on the weighted model and *frame_size* denotes the current video resolution. The constants *c* or *p* indicates whether the variable is in current or previous frame respectively, as *zero* and *w_size* means that the variable is the first and the last of frame. *RDcost* denotes the total rate distortion cost, *Predct* denotes the prediction cost and finally *BUcost* denotes the cost of BU refit to next MAD prediction.

```
//MAD prediction (current BU)
if (BU = Frame)
    w_size ← (frame_size/MB);
else
    w_size ← (frame_size/BU_size*MB);
    if (BU[c] = BU[0])
        MAD[c] ← (0.94*MAD[p])+(0.06*MAD[p+1]);
        RDcost ← Predct(MAD[c])+BUcost[c];
    else if (BU[current] = BU[w_size])
        MAD[c] ← (0.85*MAD[p])+(0.09* MAD[c-1])+
        (0.06*MAD[p-1]);
        RDcost ← Predct(MAD[c])+BUcost[c-1]+ BUcost[p-1];
    else
        MAD[c]←(0.82*MAD[p])+(0.12*MAD[c-1])+
        (0.03*MAD[p-1])+(0.03*MAD[p+1]);
        RDcost ← Predct(MAD[c])+BUcost[c-1]+
        BUcost[p-1]+ BUcost[p+1];
    endif
// Updates using Shift
    for (BUs of current frame)
        Shift(RDcost);
        Shift(MAD values);
    end for
    Return (MAD[c]);
Endif
```

Figure 2 - Pseudo-code of NFS-RC Scheme

The MAD is predicted using four multiplications, three additions, and the max number of BUs on a frame shift operations in worst case. With these steps, the algorithm proposed in this work reaches a more accurate MAD value than the quadratic model.

4. RESULTS AND COMPARISON WITH RELATED WORK

In this section we present the obtained results with the execution of our NFS-RC algorithm under JM 17.2 reference software.

Peak Signal-to-Noise Ratio (PSNR) is a well-know way to evaluate image distortion, commonly used as a measure of quality of video reconstruction to evaluate compression losses. In this case, the original data is the signal and the noise is the error introduced by compression. Thus, a higher PSNR indicates that the reconstruction is of higher quality.

Table 1 – CIF comparison NFS-RC algorithm

Sequences	Encoder	PSNR (dB)	
		QCIF	CIF
Akiyo	[4]	45.45	48.16
	NFS-RC	46.31 (+0.86)	48.91 (+0.75
Flower	[4]	28	32.57
	NFS-RC	28.83 (+0.83)	32.50 (-0.07)
Silent	[4]	39.48	42.85
	NFS-RC	40.21 (+0.73)	43.12 (+0.27)
Singer	[4]	40.09	43.82
	NFS-RC	40.67 (+0.58)	44.8 (+0.98)

In table 1 we present the PSNR values obtained by the technique in [4] (which is implemented in JM [5]), compared with our NFS-RC. The tests were performed using QCIF (176x144) and CIF resolution (352x288) videos, to enable fair comparison. The GOP structure used in this test is IPP and the total number of frames was 150 with 15 fps (frames per second) encoding rate. Finally, the target bit-rate was 512 kbps for QCIF and 1024 kbps for CIF resolution.

Results in table 1 show our NFS-RC achieved better visual quality compared to the scheme in [4]. Considering the target bit-rate, NFS-RC obtained 512.04 kbps in average which is 1.17 kbps more accurate compared with [4] in QCIF resolution. When comparing with CIF resolution results, NFS-RC obtained 1024.64 kbps in average, which means 0.99 kbps more accurate than [4].

In [9] is presented a method to compute the MAD ratio for frame complexity prediction. In addition to MAD, in [10], the PSNR is considered as a factor of the rate control mechanism. In [11] is shown a complete frame layer rate control scheme for H.264/AVC that computes the Lagrange multiplier for mode decision by using a quantization parameter.

Table 2 presents a comparison with [11] for QCIF resolution video considering a total of 300 frames with 30fps encoding. The target bit-rate is 24kbps for Carphone and Foreman sequences and 48kbps for Akiyo sequence.

Table 2 – QCIF comparison NFS-RC against [11]

Sequences	Encoder	PSNR (dB)
Akiyo	[11]	42.06
	NFS-RC	43.16 (+1.10)
Foreman	[11]	30.57
	NFS-RC	31.14 (+0.57)
Carphone	[11]	31.94
	NFS-RC	32.02 (+0.08)

In [12], a novel mode-based rate control algorithm applying a two-stage encoding was proposed. In addition, a MB-based bit allocation method to improve the video quality was proposed.

Table 3 – QCIF comparison NFS-RC against [13]

Sequences	Encoder	PSNR (dB)
Akiyo	[13]	45.96
	NFS-RC	45.98 (+0.02)
Carphone	[13]	36.37
	NFS-RC	36.30 (+0.07)
Foreman	[13]	35.57
	NFS-RC	35.50 (-0.07)
Silent	[13]	39.64
	NFS-RC	39.78 (+0.14)

Finally, work in [13] presents an adaptive rate control mechanism for H.264/AVC to produce a more stable output video stream, given a burst video sequence as the input, avoiding the possible buffer overflow and quality decrement on the way of video streaming over networks. In table 3 we present a comparison with [13] for QCIF resolution video with 15fps encoding. The bit-rate target for these sequences is 64kbps.

5. CONCLUSIONS

This paper proposed a new rate control scheme (called NFS-RC) based on neighbor MAD correlation to improve quality and complexity for H.264/AVC video coding. NFS-RC was implemented in JM H.264 reference software. Experimental results indicates that the improved algorithm controls rate more accurately than existing techniques and also achieves higher video quality for the same rate. An effective algorithm for BU layer rate control to solve the problem of rate variation was implemented. The results demonstrate that a fast MAD prediction based on neighboring is a good way to reduce the computational costs presented by existing RC techniques. Finally, NFS-RC obtained higher video quality than related works in most cases. Furthermore, our method not only provides improvements in video quality but more accurately allocate bits to respect target channel bandwidth.

6. REFERENCES

[1] ITU-T H.264 Recommendation (03/2010): advanced video coding for generic audiovisual services, 2010.

[2] Shafique, M.; Sun, H. *Non-linear rate control for H.264/AVC video encoder with multiple picture types using image-statistics and motion-based macroblock prioritization,* ICIP 2009.

[3] Sullivan, G.J.; Wiegand, T. *Rate-distortion optimization for video compression*. IEEE Signal Processing Magazine, vol.15, no. 6, pp.74-90, Nov 1998.

[4] Li, Z., Pan, F., Lim, K. P., Feng, G., Lin, X., and Rahardja, S., *Adaptive Basic Unit Layer Rate Control for JVT*, document JVT-G012-r1, 7th Meeting, Pattaya II, Thailand, Mar. 2003.

[5] Suhring, K. H.*264/AVC Reference Software*. In: Fraunhofer Heinrich-Hertz-Institute. <http://iphome.hhi.de/suehring/tml/download/>.

[6] Chen, C.-T. *Linear system theory and design*. Rinehart and Winston, New York, 1984.

[7] Malewar, A.; Singh, R. K.; Gadre, V.; *Novel Target Bit-rate estimation for improved frame level H.264 rate control. Signal Processing and Communications*, 2010.

[8] Song, L.-X., Zhao, R., Li, X.-E., *An Optimized Rate Control Algorithm for H.264*, CECNet 2011

[9] Jiang, M., Yi, X., and N. Ling, *Improved frame-layer rate control for H.264 using MAD ratio*, in Proc. IEEE ISCAS, 2004, pp. 813–816

[10] Jiang, M. and Ling, N., *On enhancing H.264/AVC video rate control by PSNR-based frame complexity estimation,* IEEE TCE, vol. 51, no. 1, pp. 281–286, Feb. 2005.

[11] Jiang, M. and Ling, N., *On Lagrange multiplier and quantizer adjustment for H.264 frame-layer video rate control*, IEEE TCSVT, vol. 16, no. 5, pp. 663–669, May 2006.

[12] Yang, X., Tan, Y., and Ling, N., *Rate control for H.264 with two-step quantization parameter determination but single-pass encoding*, EURASIP J. Appl. Signal Process., pp. 35–47, 2006.

[13] Chen, J.-Y., Chiu, C.-W., Li, G.-L., Chen, M.-J., *Burst-Aware Dynamic Rate Control for H.264/AVC Video Streaming*. IEEE Trans. on Broadcasting, vol.57, no.1, pp.89-93, Mar. 2011.