



## Rate Control in Open-Loop MPEG Video Transcoder

Bence FORMANEK<sup>1</sup>, Tihamér ÁDÁM<sup>2</sup>

<sup>1</sup> R&D Department, CableWorld Ltd., Budapest, Hungary,  
e-mail: formanek.bence@cableworld.hu

<sup>2</sup> Department of Automation, University of Miskolc, Miskolc, Hungary,  
e-mail: adam@mazsola.iit.uni-miskolc.hu

Manuscript received March 15, 2009; revised June 06, 2009.

**Abstract:** Video transcoding is intended to provide transmission flexibility to pre-encoded bit streams by dynamically adjusting the bit rate of these bit streams according to new bandwidth constraints that were unknown at the time of encoding. What makes transcoding different from video encoding is that transcoding has access to many coding parameters, which can be obtained from the input compressed stream. They may be used not only to simplify the computation, but also to improve the video quality. In this paper, we propose a low complexity rate-control method for open loop MPEG-2 video transcoders, working entirely in the frequency domain.

**Keywords:** Video, MPEG, MPEG-2, bitrate, transcoder, DVB, codec.

### 1. Introduction

The process of converting between different compression formats and/or further reducing the bit rate of a previously compressed signal is known as transcoding, and it can introduce significant impairments in video quality if performed without due care. There are three basic requirements in transcoding. The information in the original bitstream should be exploited as much as possible. The resulting video quality of the new bitstream should be as high as possible, or as close as possible to the bitstream created by coding the original source video at the reduced rate [1], [2]. In real-time applications, the transcoding delay and memory requirement should be minimized to meet real-time constraints.

## 2. Video transcoding

The most straightforward transcoding architecture is to cascade a decoder and an encoder directly. In this architecture, the incoming source video stream is fully decoded, and then the decoded video is re-encoded into the target video stream with desirable bitrate or format. It is computationally very expensive, but often used way of video transcoding.

A more efficient solution to perform conversion between video bitstreams of the same standard - homogeneous transcoding - is open-loop transcoding. In an open-loop transcoder the process of video coding is reversed until the quantization step, a new quantizer value is calculated for lower bitrate, then the DCT coefficients requantized with this new quantizer value, and the rest of the video coding process is executed again with the new DCT coefficient values. Because of the higher quantizer step, the amount of information contained in each picture will be lower, which means lower bitrate for the entire video stream. Open-loop transcoders are computationally efficient, since they operate directly on the DCT coefficients. However, they suffer from the drift problem [3].



Figure 1: Open-loop transcoder.

The drift problem is explained as follows. A video picture is predicted from its reference pictures and only the prediction errors are coded. For the decoder to work properly, the reference pictures reconstructed and stored in the decoder predictor must be same as those in the encoder predictor. The open-loop transcoders change the prediction errors and, therefore, make the reference pictures in the decoder predictor different from those in the encoder predictor. The differences accumulate and cause the video quality to deteriorate with time until an intrapicture is reached. The error accumulation caused by the encoder/decoder predictor mismatch is called drift and it may cause severe degradation to the video quality [4], [5].

In MPEG-2 video compression, Intra-coded frames (I frames) are encoded without reference frame, MC is not needed in encoding I frames, so the transcoding of I frames is not subject to the drift. Bi-directionally predictive coded frames (B frames) are not used for predicting future frames [6]. Therefore, the transcoding of B frames does not contribute to the propagation and accumulation of the drift. The drift error is only caused by the transcoding operation of predictively coded frames (P frames), because they are used as further reference for prediction and can accumulate through a GOP (Group Of

Pictures). The quality deterioration gradually increases until the next I frame refreshes the video scene, so this error is also called “breathing”.

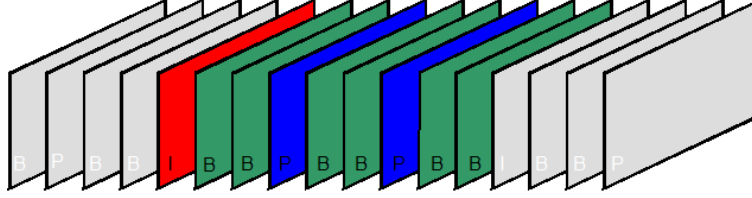


Figure 2: Picture coding types in MPEG-2.

However, the degree of the video quality degradation caused by the drift varies with architectures. In addition, the drift will be terminated by an intrapicture. In the applications where the number of coded pictures between two consecutive intrapictures is small and the quality degradation caused by the drift is acceptable, these architectures, although not drift free, can still be quite useful due to the potentially lower cost in terms of computation and required frame memory.

### 3. Rate Control

The goal of rate control in video coding and transcoding is to achieve a target bitrate with good and consistent visual quality. Rate control is responsible for maintaining consistent video quality while satisfying bandwidth, delay, and memory constraints by determining picture quantization parameters. The particular picture quality and rate of an MPEG-2 encoder is achieved by selecting a specific quantizer scale for each macroblock in picture. This value is calculated for each macroblock. A picture global quantization scale is the average of the macroblock quantizer scale values in that picture.

In the widely used MPEG-2 Test Model 5 (TM5) [7], [8], a picture complexity measure characterizes the difficulty in coding a picture, so that the target number of bits for coding that picture is proportional to its complexity. This complexity measure can be computed from the picture's spatial properties. Measurements have shown that this is actually a fairly accurate model for a large variety of scenes.

$$b_i = f(Q_i) = \frac{c_i}{Q_i}, \quad (1)$$

where  $b_i$  refers to the target bits per picture,  $Q_i$  is the picture global quantization scale,  $c_i$  is the picture complexity measure.

However, decoded images are not available in an open-loop transcoder, so complexity measure can't be calculated in a similar way to encoders. In transcoding, it is yet possible to compute the frame complexities from the input bit-stream, since the quantization step sizes and the number of bits information per picture are available, and then we can use these complexities for the bit allocation in transcoding. From equation (1), we could assume that:

$$\frac{b_{ti}}{b_{oi}} = f\left(\frac{Q_{oi}}{Q_{ti}}\right) = \alpha\left(\frac{Q_{oi}}{Q_{ti}}\right) + \beta, \quad (2)$$

where  $b_{ti}$  is the target bits for transcoded picture,  $b_{oi}$  is the original size of the current picture in bits,  $Q_{oi}$  is the picture global quantization scale of the original picture,  $Q_{ti}$  is the picture global quantization scale for transcoded picture, and  $\alpha$  and  $\beta$  are parameters of the linear equation. We assume that the value above can be approximated by a linear equation.

Since a video sequence consist of successive video frames, the bitrate of the sequence depends on the size of the successive pictures, so we can say that:

$$\frac{b_{ti}}{b_{oi}} = \frac{R_a}{R_o}, \quad (3)$$

where  $R_a$  is the average output bitrate we would like to achieve,  $R_o$  is the input bitrate of the original video sequence.

Since different types of pictures have different sizes and in a group of pictures all picture types can be found, starting with an I picture, usually the GOP size, or GOP bitrate is specified in an encoder.

$$R_{GOP} = \frac{\sum_{n=1}^{N_{GOP}} b_n}{TN_{GOP}}, \quad (4)$$

where  $R_{GOP}$  is the bitrate of the GOP,  $N_{GOP}$  is the number of frames in a GOP,  $b_n$  is the size of the single pictures and  $T$  is the frame time.

A rate-control algorithm usually needs to know the GOP configuration. However, in transcoding, the output GOP configuration is often determined by the input one, since transcoding typically does not change the frame coding types in order to keep complexity low. In real-time transcoding, the input GOP configuration is usually unknown to the transcoder.

To control the output bitrate and the size of the single pictures, our rate control algorithm needs to know the occurrence rate of different pictures types. The input video stream has to be analyzed as long as the frame coding type rate can be calculated. As GOP size is variable it means several GOP times.

To set the output bitrate, the picture global quantization scale of the currently transcoded picture has to be calculated from the current picture's known input quantization scale.

$$Q_{ti} = mQ_{oi}, \quad (5)$$

where  $Q_{oi}$  is the current picture's input global quantization scale,  $Q_{ti}$  is the global quantization scale of the currently transcoded picture, and  $m$  is a multiplier that controls the bitrate reduction. Using equations (2), (3) and (5), we can write:

$$\frac{R_a}{R_o} = \alpha \left( \frac{1}{m} \right) + \beta. \quad (6)$$

In equation (6), there is  $m$  that can be expressed as:

$$m = \frac{\alpha}{\left( \frac{R_a}{R_o} \right) - \beta}, \quad (7)$$

where  $R_a$  is the output average bitrate we would like to achieve,  $R_o$  is the input bitrate of the original video sequence, and  $\alpha$  and  $\beta$  are parameters from equation (2).

In our previous calculations we did not take into account the real output bitrate, only the output bitrate that should be achieved. Similarly, we did not take into account that values of the used parameters are dependent on the picture characteristics. So we have to examine the effective output bitrate and correct the calculated parameter values:

$$e = \frac{R_t - R_a}{R_a}, \quad (8)$$

where  $R_t$  is the effective output bitrate,  $R_a$  is the output average bitrate that should be achieved and  $e$  is a weighted difference from the desired output bitrate. With this weighted difference value used as an error signal, we create a feedback to influence the rate control algorithm and the new quantizer value.

From the above and with the modification of equation (7), we get the final equation which contains all the parameters from equations above and gives the  $m$  multiplier value for the actual picture:

$$m = \frac{\alpha(1 + e)}{\left( \frac{R_a}{R_o} \right) - \beta}, \quad (9)$$

where  $R_a$  is the output average bitrate that should be achieved,  $R_o$  is the input bitrate of the original video sequence, and  $e$  is the weighted error signal. Only the definition of  $\alpha$  and  $\beta$  remains.

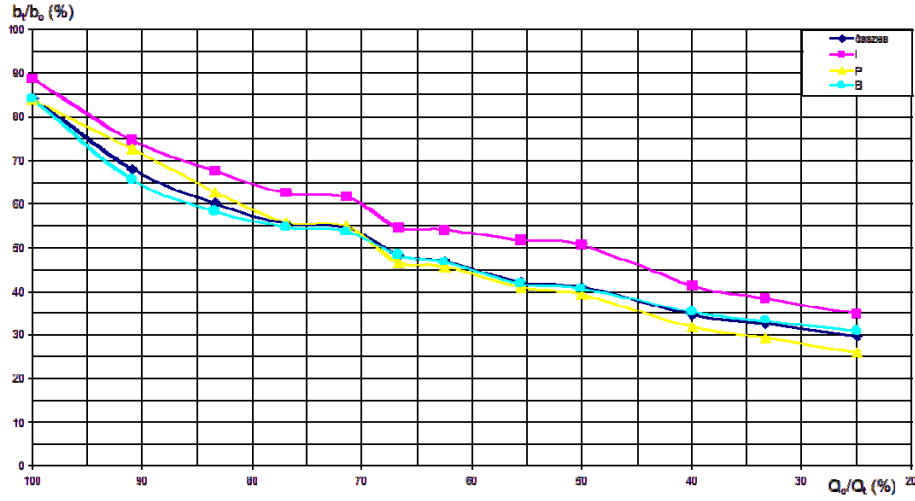


Figure 3: Picture size and picture global quantization scale.

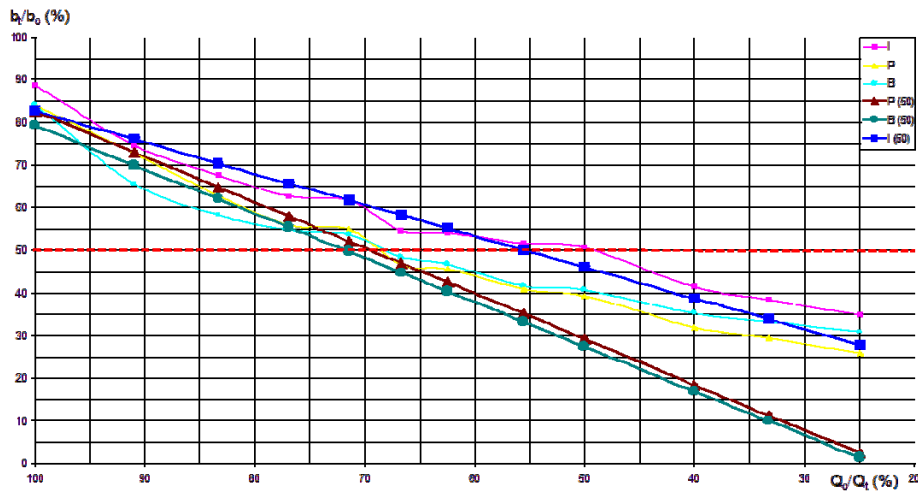


Figure 4: Picture size and linear approximation of picture global quantization scale.

Experiments were carried out to determine the value of  $\alpha$  and  $\beta$ . We analyzed the average ratio between the number of bits in corresponding input

and output images, with different  $m$  values in different video sequences. The experimental results are shown in *Fig. 3* for one of the used video sequence (Mill).

It can be seen from the experimental results, that the ratio can be approximated by a linear function (2) between the number of bits in corresponding input and output images, and that  $\alpha$  and  $\beta$  parameter values are different for different picture types. *Fig. 4* shows the lines of a linear approximation with a further assumption that – in general –, the degree of bitrate reduction will not exceed 50%. To achieve a more accurate approximation on that section only points belonging to ratios above 50% are taken into consideration.

#### 4. Experimental results

We analyzed a transcoder realization based on the concept above with different video streams.

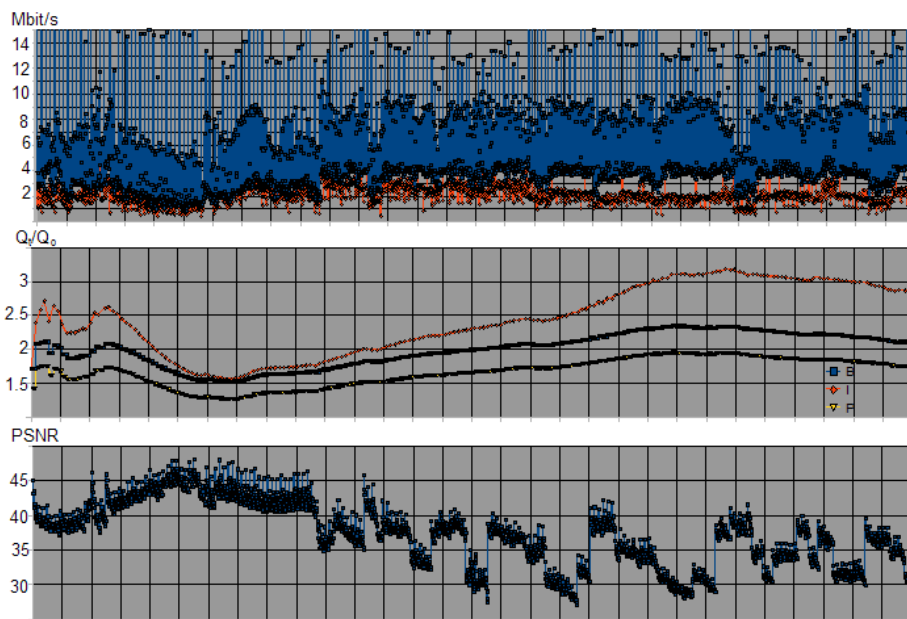


Figure 5: 3,000 frame, bitrate, value of  $m$  and picture quality (PSNR (dB)).

Results in *Fig. 5* are from a scene of 3,000 pictures from a VBR (Variable Bit Rate) live television broadcast. The input average bitrate is 4.5 Mbit/s, the

output average bitrate is set to 2Mbit/s, and it is VBR also.  $m$  multiplier is shown as changing to keep the set output bitrate.

Fig. 6 shows an enlarged part of Fig. 5, it is a scene of 300 pictures. The typical open-loop transcoder error called “breathing” can be seen on the picture quality graph.

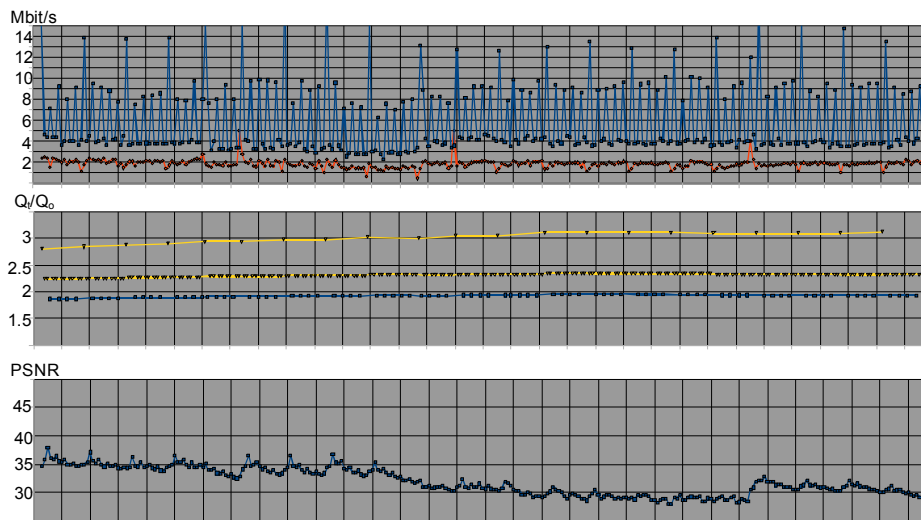


Figure 6: 300 frame, bitrate, value of  $m$  and picture quality (PSNR (dB)).

## References

- [1] Ahmad, I., Wei, X., Sun, Y., Zhang, Y.-Q., “Video transcoding: an overview of various techniques and research issues”, *IEEE Trans. on Multimedia*, Vol. 7, Issue 5, pp. 793-804, Oct. 2005.
- [2] Tudor, P.N., and Werner, O. H., BBC R&D “Real-time transcoding of MPEG-2 video bit streams” *IBC'97 Amsterdam*, pp. 286-301, 1997.
- [3] Lastein, L., and Reul, B., “Transrating of MPEG-2 streams - various applications and different techniques”, *BarcoNet - a Scientific Atlanta Company*, Denmark.
- [4] Xin, J., Lin, C.-W., Sun, M.-T., “Digital Video Transcoding”, *Proceedings of the IEEE*, Vol. 93, Issue 1, pp. 84 – 97, Jan. 2005.
- [5] Assuncao, P. A. A.; Ghanbari, M., “A frequency-domain video transcoder for dynamic bit-rate reduction of MPEG-2 bit streams”, *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 8, Issue 8, pp. 953 – 967, Dec 1998.
- [6] Int. Standards Org./ Int. Electrotech. Comm., (ISO/IEC), “Information technology - Generic coding of moving pictures and associated audio information: Video (MPEG-2 video)”, *ISO/IEC 13818-2, 2nd ed.*, 2000.
- [7] “ISO/IEC-JTC1/SC29/WG11, Test model 5, MPEG93/457”, Apr. 1993.
- [8] Westerink, P. H., Rajagopalan, R., Gonzales, C. A., “Two-pass MPEG-2 variable-bitrate encoding”, *IBM J. Res. Develop.*, Vol. 43, No. 4, July 1999.