УДК 621.397.2

V.G. Abakumov[1], Dr. Sc., T. Ádám[2], Ph. D., B. Formanek[2], A. Kane[2], Ph. D.,
P.V. Popovych[1]

Rate Control in Open-Loop Video Transcoder

Предложен метод управления скоростью видеопотока на основе транскодера без обратной связи. Получена математическая модель транскодирования видеопотока, связывающая количество бит на входе и выходе транскодера с глобальным коэффициентом квантования до и после транскодирования. Проведены экспериментальные исследования предложенного метода управления скоростью видеопотока, результаты которого подтверждают адекватность полученной модели.

Method of video stream rate control on basis of open-loop transcoder is offered. Expression linking the number of bits in transcoder's input and target number of output bits with global quantization scale before and after transcoding is obtained. Experimental researches of offered method are carried out which confirm the functionality of this model.

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. It is intended to provide transmission flexibility to pre-encoded bit streams by dynamically adjusting the bit rate of these bitstreams 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. 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.

1. 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 (Fig.1). It is computationally very expensive, but often used way of video transcoding.

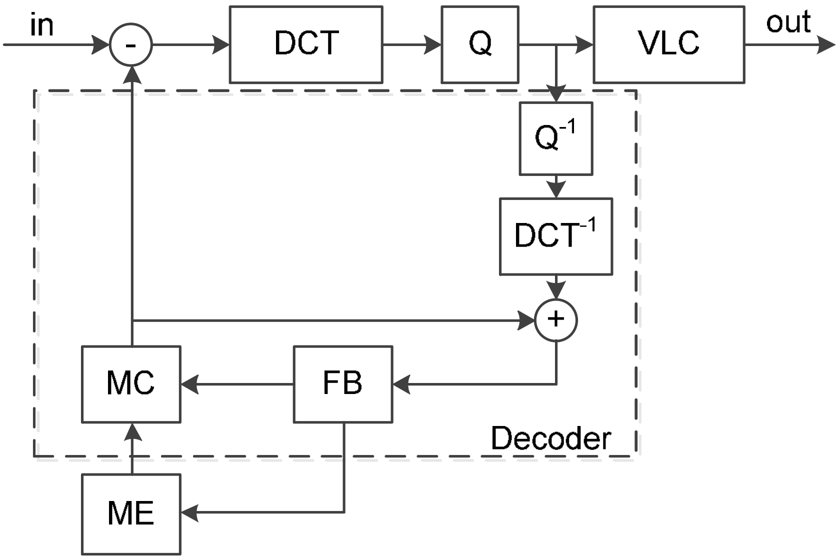

Fig. 1. Video encoder: DCT – Discrete Cosine Transform, Q – Quantizer, DCT^{-1} – Inverse DCT, Q^{-1} – Dequantizer, VLC – Variable Length Coder, FB – Fame Buffer, ME – Motion Estimator, MC- Motion Coder

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 (Fig. 2). 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].

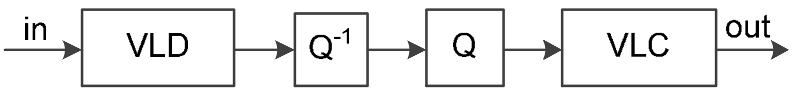

Fig. 2. Open-Loop transcoder: VLD - Variable Length Decoder

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. 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) (Fig. 3), 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».

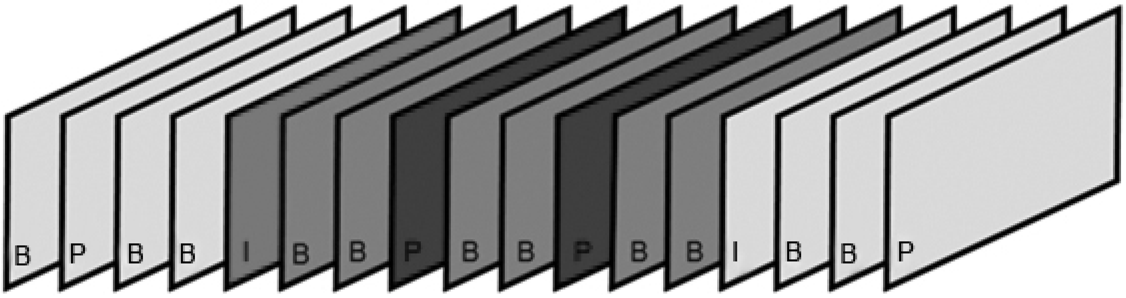

Fig. 3. Picture coding types in MPEG-2

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 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.

2. Video bitrate 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 [9]:

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

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; α and β 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:

$$\frac{b_{ti}}{b_{oi}} = \frac{R_a}{R_o}, \tag{3}$$

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:

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

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; 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:

$$Q_{ti} = mQ_{oi}, \tag{5}$$

where m is a multiplier that controls the bitrate reduction. Using equations (2), (3) and (5), we can write:

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

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

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

where α and β 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:

$$e = \frac{R_t - R_a}{R_a}, \tag{8}$$

where R_t is the effective output bitrate; R_a is the average output bitrate we would like to achieve; e is a weighted difference from the desired output bitrate represented as error signal. 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:

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

Only the definition of α and β remains.

Experiments were carried out to determine the value of α and β. 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. 4 for one of the used video sequence.

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 α and β parameter values are different for different picture types. Fig. 5 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.

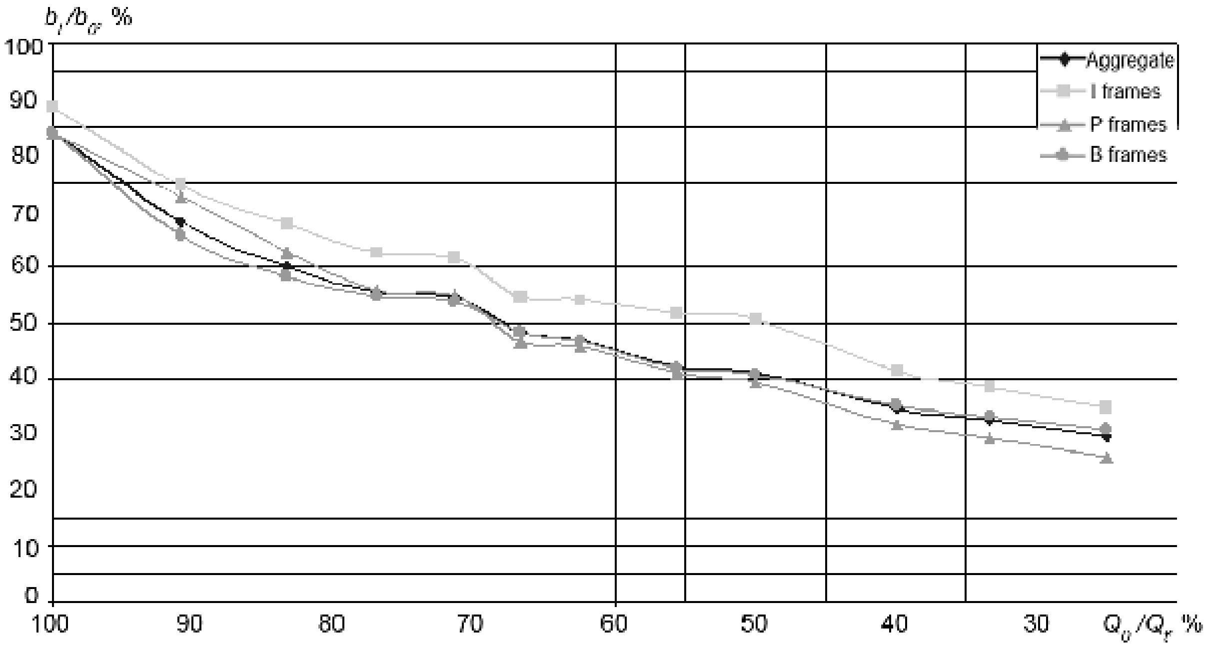

Fig. 4. Picture size and picture global quantization scale

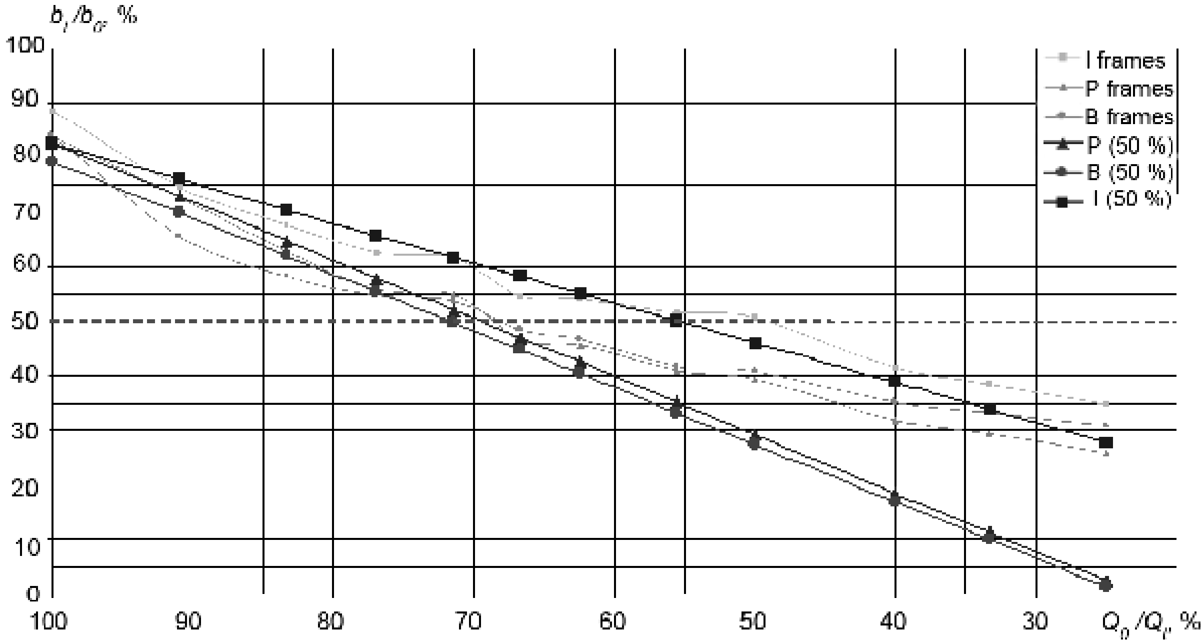

Fig. 5. Picture size and linear approximation of picture global quantization scale

3. Experimental results

We analyzed a transcoder realization based on the concept above with different video streams. Results in Fig. 6 are from a scene of 3000 pictures from a VBR (Variable Bit Rate) live television broadcast. The input average bitrate R_o is

4,5 Mbit/s, the output average bitrate R_a is set to 2 Mbit/s, and it is VBR also. *m* multiplier is shown as changing to keep the set output bitrate.

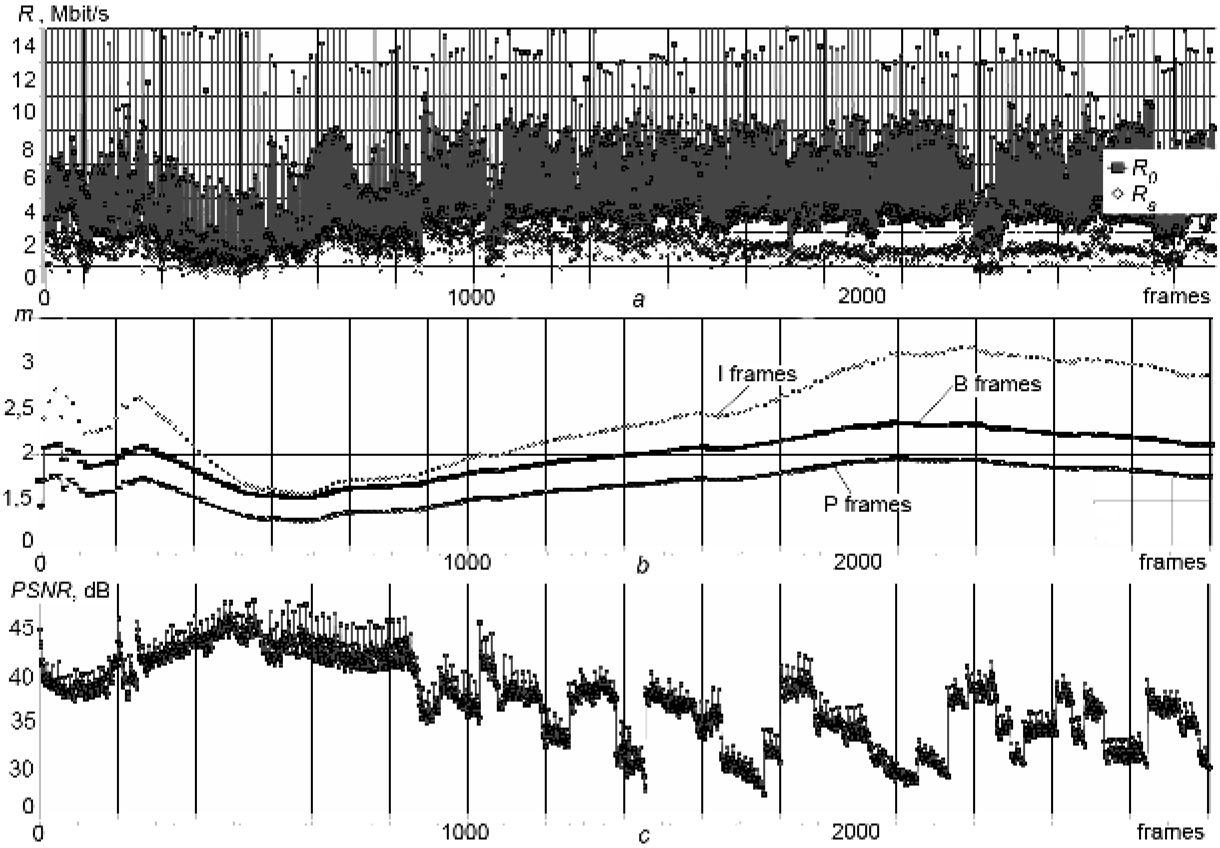

Fig. 6. Bitrate (*a*), value of *m* (*b*) and picture quality (*c*) for 3000 frames

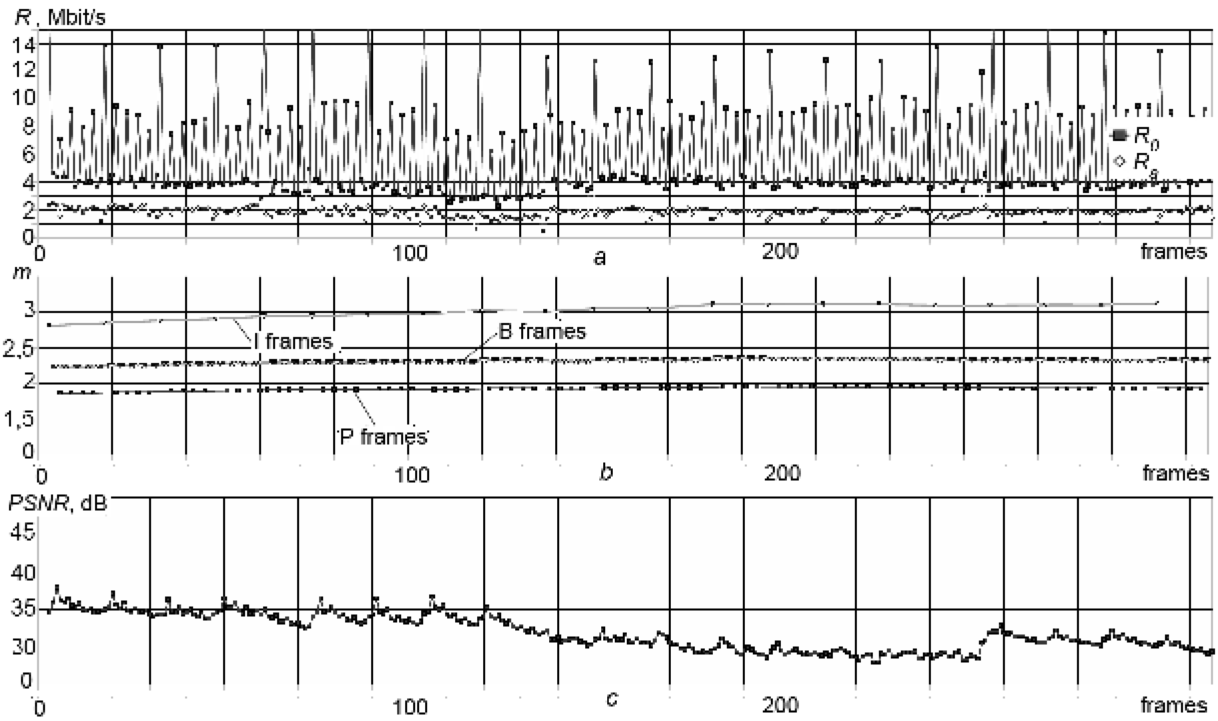

Fig. 7. Bitrate (*a*), value of *m* (*b*) and picture quality (*c*) for 300 frames

Fig. 7 shows an enlarged part of Fig. 6, it is a scene of 300 pictures. The typical open-loop transcoder error called «breathing» (presented by peak signal-to-noise ratio (PSNR)) can be seen on the picture quality graph.

Conclusions

Experimental results show good functionality of open-loop video transcoding method that can be linearly described. This is simply to implement on the basis of modern IC and microprocessor components. In future it is possible to control the video bit rate using other objective picture quality metrics such as SSIM, VQM, etc.

References

1. *Ahmad I.,* Xiaohui Wei, Yu Sun, Ya-Qin Zhang. Video transcoding: an overview of various techniques and research issues // Multimedia IEEE Transactions. – 2005. – Vol. 7, № 5. – P. 793-804.

2. *Tudor P. N.*, Werner O. H. Real-time transcoding of MPEG-2 video bit streams // BBC R&D: IBC'97. – Amsterdam, 1997. – P. 286-301.

3. *Lastein L.*, Reul B. Transrating of MPEG-2 streams - various applications and different techniques // Broadcast Asia 2002, June 2002. – http://www.broadcastpapers.com/whitepapers/IBCBarconetTransratingEfficiency.pdf

4. *Xin, J.,* Lin C.-W., Sun M.-T. Digital Video Transcoding // Proceedings of the IEEE. – 2005. – Vol. 93, № 1. – P. 84-97.

5. *Assuncao P.A.A.*, Ghanbari M. A frequency-domain video transcoder for dynamic bit-rate reduction of MPEG-2 bit streams // Circuits and Systems for Video Technology: IEEE Transactions. – 1998. – Vol. 8, № 8. – P. 953-967.

6. Information technology - Generic coding of moving pictures and associated audio information: Video (MPEG-2 video). Int. Standards Org./Int. Electrotech. Comm. (ISO/IEC), ISO/IEC 13818-2, 2nd ed., 2000. – http://webstore.iec.ch/preview/info_isoiec13818-2%7Bed2.0%7Den.pdf

7. ISO/IEC-JTC1/SC29/WG11. Test model 5, MPEG93/457, Apr. 1993. – http://www.mpeg.org/MPEG/MSSG/tm5/Overview.html

8. *Westerink P. H.*, Rajagopalan R., Gonzales C. A. Two-pass MPEG-2 variable-bitrate encoding // IBM J. Res. Develop. – 1999. – Vol. 43, № 4. – P. 471-488.

9. *Formanek B.*, Ádám T. Rate control in MPEG-2 video bitrate transcoder // Proc. of the 10th International Carpathian Control Conference. – 2009. – P. 297-300.

Ключевые слова: видеотранскодер, без обратной связи, управление скоростью, битрейт, качество изображения, квантование.

[1] *Национальный технический университет Украины «КПИ»*

[2] *Университет г. Мишкольц, Венгрия*