Constant Bit Rate For Video Streaming Over Packet Switching Networks

Dr. Nasser N. Khamiss
nasser_alani@yahoo.com
Nahrain University / Information Engineering College

Hamsa A. Abdullah
hamssa_kareem@yahoo.com
Nahrain University / Information Engineering College

ABSTRACT

This paper proposed a technique of bit-rate control for video streaming over constant bit-rate communication channel, where the MPEG-2, standard variable bit-rate (VBR), is adapted to be used as a constant bit-rate (CBR). The target image quality and output rate of the video encoder is controlled by feedback based on the buffer level. A CBR transport over networks, result is a good performance compared with that of VBR. It introduced simplicity of network monitoring and analysis, where the VBR video streaming over CBR channel can be adapted to avoid the problem of congestion of the network. This paper studies the important issue of adapting VBR-compressed video for transport over a CBR channel. The developed systems are implemented using Matlab (Ver 6.5) under Windows XP operating system.

KEY WORD
Code, Video streaming, Compression, Control, and Bit rate.

1. INTRODUCTION

Video compression, which is a necessary process for video communication over networks, removes spatial and temporal redundancies contained in video sequences. The MPEG-2 format can provide extremely high-quality video with a data rate of around 1 MB/sec. However, when transmitted over networks, compressed video streams are very sensitive to bit errors and packet losses, which lead to inevitable spatial and temporal error propagation. Due to the limited storage size or communication bandwidth, quantization is introduced to reduce the bit rate of the compressed video signal such that the size or bandwidth limitation can be met properly Rate control mechanism is responsible to adjust the quantization parameters (QPs) to achieve this purpose see Fig.1 [1].

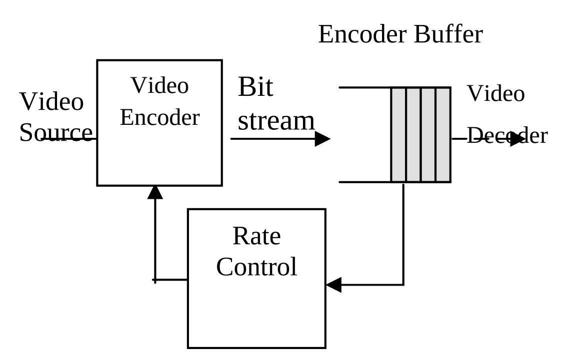


Figure (1) Rate control in video coding system

Generally, the rate control part is an informative part in video coding standards. It leaves the flexibility to designers to develop the suitable scheme for specific applications. Whilst matching the desired bit rate, rate control needs to consider the following challenging issues; distortion, complexity, and constraints.

The Distortion D is a decreasing function of the bit-rate R. There is an inherent tradeoff between the distortion and the bit rate. Based on the rate distortion (R-D) theory, the distortion D is a decreasing function of the bit-rate R. A decreasing in distortion leads to an increasing in rate and vice versa. So the fundamental problem in rate control can be stated that:

$$\min \ D, \quad \text{s.t. } R \le R_{max} \qquad (1)$$

Where R_{max} denotes the maximum bit rate. In other words, rate control is to achieve the maximum picture quality (minimum distortion) without exceeding the maximum permitted bit rate where quality is typically represented by peak signal noise ratio (PSNR).

While the complexity for a different application has different computational requirements. For real-time video coding systems, time-consuming rate control algorithm should be avoided. As a dual problem, the rate control algorithm is expected to be simple enough with acceptable performance or is close to the optimal solution with reasonable cost.

On the other hand, there are various constraints for video compression applications. For end-to-end real-time video communication systems, there is the delay constraint to avoid delay jitter and jerky motion. For constant bit-rate applications, buffer constraint is introduced. The rate control algorithm should guarantee the encoder and decoder buffers do not overflow or underflow. In CBR applications, rate control designers mainly focus on how to improve the matching accuracy between the targets bit rate and actual bit rate and satisfy the low-latency and buffer constraints. Further more, the fluctuation of video quality cannot be avoided due to the varying content in natural scenes. However, by taking the advantage of the encoder buffer, smoothing video quality is possible as the buffer can tolerate limited bit-rate fluctuation whilst the buffer does not either overflow or underflow [2].

2. RATE CONTROL CONSTRAINED QUALITY

Rate control always evolves into constrained problems in practical applications [3]. Since the amount of information in compressed video sequences is inherently variable, a buffer is placed between the video encoder and the transmission channel to smooth out the rate variation. The buffer also dictates the amount of delay in transmission systems. Larger buffer corresponds to long end-to-end delay. In practical video coding applications, it is essential to consider the rate/buffer/delay constraints. Rate control is then responsible for adjusting the coding parameters to achieve maximum picture quality and ensure the buffer never underflows or overflows. The proposed rate control algorithm is designed for bit-rate control includes the requirements of MPEG-2 video coding standard. The TM5 rate control algorithm is designed for bit-rate control in MPEG-2 video coding standard. It consists of the following steps:

(1) Target bit allocation: The target number of bits for the next picture depends on the picture-type and universal weighting factors.

(2) Rate control: The reference value of the QP for each macro block (MB) (Q_j) is set as follows:

$$Q_j = \frac{B_j \times 31}{r} \qquad (2)$$

Where r =2 × R/*f*, **R** denotes the bit rate (bps), *f* denotes the frame rate (fps), for a constant quantization step size of 31, and B_j is the fullness of the buffer.

(3) Adaptive quantization: Finally, the QP for MB *j* is $mquant_j = Q_j \times N\_act_j$ and is clipped to the range, where N_act_j is the normalized spatial activity measure for the MB *j*.

3. MODE SELECTION AND SIZE SELECTION

In video coding, there are different frame types I, P, and B frames with different Macro Block (MB) modes (Intra, Inter, skip, etc.). These various types (modes) relate to different bit rates and different residual errors. Frame-type selection can be optimized by incorporating the Lagrange multiplier method to minimize the distortion under the bit budget constraint [4]. The strategy was to determine both the number and the position of reference frames in a group of pictures whilst choosing the optimal quantizer. The mode selection considers the impact of possible modes on R-D behavior in the group of MBs and the optimal strategy is based on the joint selection of MB modes and quantizer scales to achieve the highest picture quality. Variable block size video coding is to balance the tradeoff between the bit rate and distortion by choosing the optimal block size in the frame. Optimal selection of coding parameters in a video coder involves variable-block-size motion estimation and multimode residual coding.

4. FRAME DROPPING FILTER: RATE SHAPING

The frame-dropping filter is used to reduce the data rate of a video stream by discarding a number of frames and transmitting the remaining frames at a lower rate [5]. Before removing the temporal redundancy from the current Inter-frame, it's compared with the previous frame and the difference is measured between them. If the difference is very small, the current frame will be discarded, and the next frame is used as the current frame. But in the receiver side, the discarded frame will be displayed by repeating the previous frame, (i.e. if we transmit 30 fps, and for example in the encoder side we discard two frames, then in the receiver side the number of frames that will be displayed are 32 frame, because the system was design to transmit 30 fps). But if there are no frames to discard, the number of frames displayed equals to the number of frames transmitted. The procedure of frame dropping is as follows:

Step 1: Determine the similarity between the frames.

Step 2: If the current frame is very similar to the pervious frame, then it's drop.

Step 3: The step 1 and 2 repeated for all frame in video sequence.

The encoder transmits the encoded frame with its number and it would be used in decoder side to find the dropped frame. Fig. 2 shows the frame dropping sequence.

5. RATE CONTROL SCHEME

Rate control is an important issue in video streaming applications for both wired and wireless networks. Rate control schemes can be classified into two major categories:

Constant-bit-rate (CBR) control is used for the constant-channel-bandwidth video transmission, and variable-bit-rate (VBR) control for the variable-channel-bandwidth

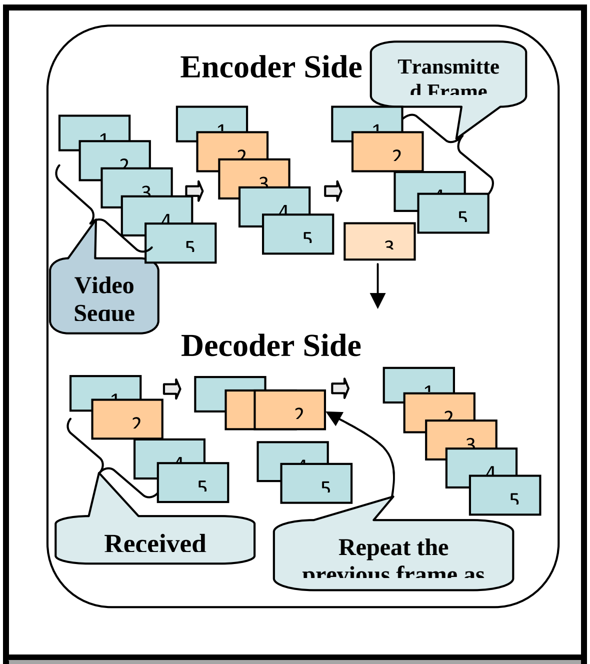


Figure (2) The frame dropping sequence

transmission. These rate-control schemes usually resolve two main problems. The first is how to allocate proper bits to each coding unit according to the buffer status, i.e., rate allocation, and the second is how to adjust the encoder parameters to properly encode each unit with the allocated bits, i.e., quantization parameter adjustment. The rate allocation in the rate control is usually associated with a buffer model specified in the video-coding standard such that used in [6]. Rate control mechanisms are useful for two different things, to send a VBR video stream on a CBR channel or to adapt to congestion on the network. The main idea is to decrease the quality for the more complex scenes and increase it for the less complex ones to achieve a constant bit rate. There are different ways to adjust the output rate from an encoder. Firstly by decreasing the sampling rate, secondly by increasing the quantization step size, finally by decreasing the number of bits used for each pixel or by increasing the movement detection threshold. Common for all reduction techniques is that the quality is reduced [7]. The rate control includes two levels of rate control: frame-level rate control and MB-level rate control as

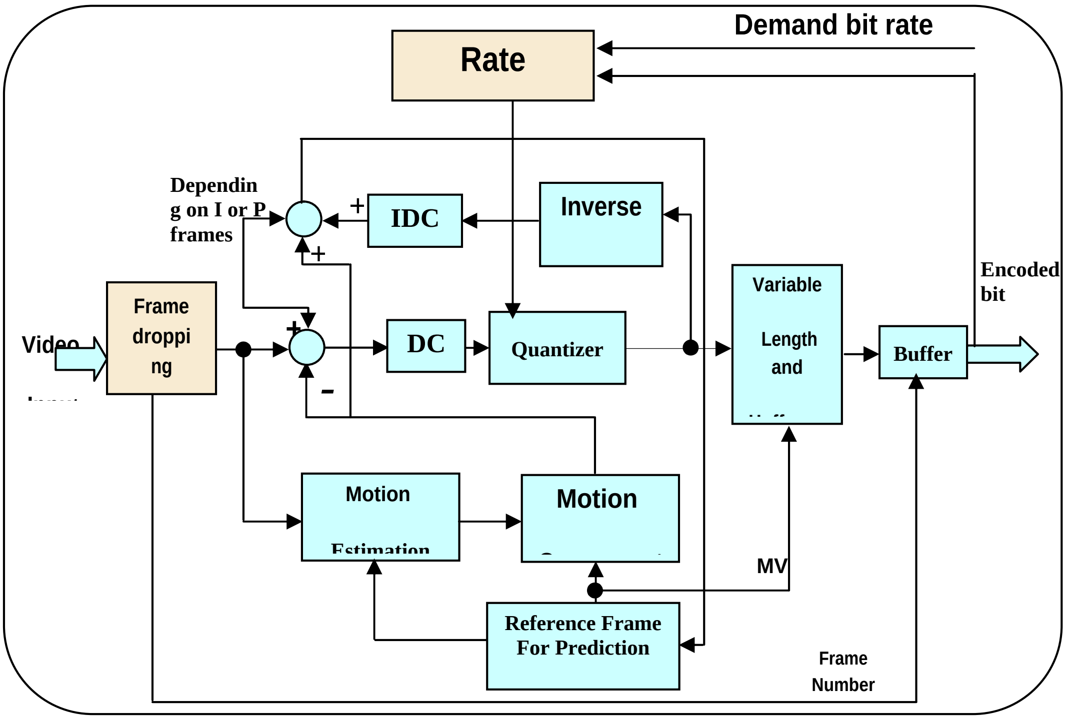


Figure (3) The proposed rate control scheme

used in [8]. In this work, at the frame-level rate control, the estimation target bit of the current frame is used according to Motion Complexity (MC).

To solve such a problem, the Rate Control scheme, Fig.3, is suggested. The propose system is the rate control scheme, where all the encoded bits are put into an output buffer for transmission. The output buffer fullness should be considered in target bit estimation so that the buffer is maintained neither overflow nor underflow. When the output bit rate not match with target bit estimation the rate control provide new quantization step size to encode the data with new quality to be match with target bit rate.

6. PROPOSED BIT-RATE CONTROL REQUIREMENTS:

6.1 Complexity Estimation

A simple metric that reflects the encoding complexity associated with the residuals is Mean Absolute Difference. The MAD is used as the prediction error which is a convenient surrogate for this purpose as follows:

$$MAD = \sum_{ij} |source_{i,j} - prediction_{i,j}| \quad (3)$$

This MAD is an inverse measure of predictor's accuracy and the temporal similarity of adjacent pictures. Ideally, the MAD would be estimated after encoding the current picture, but that would require encoding the picture again after the QP is selected [9]. Instead, we can usually assume that this complexity surrogate varies gradually from picture to picture, and estimate it based upon data extracted from the encoder for previous pictures.

6.2 Virtual Buffer Model

Any compliant decoder is equipped with a buffer to smooth out variations in the rate and arrival time of incoming data. The corresponding encoder must produce a bit stream that satisfies constraints of the decoder, so a virtual buffer model is used to simulate the fullness of the real decoder buffer. The change in fullness of the virtual buffer is the difference between the total bits encoded into the stream, less a constant removal rate assumed to equal the bandwidth (or demanded bit rate). The buffer fullness is bounded by zero from below and by the buffer capacity from above. The user must specify appropriate values for buffer capacity and initial buffer fullness, consistent with the decoder levels supported.

6.3 GoP Bit Allocation

Based upon the demanded bit rate and the current fullness of the virtual buffer, a target bit rate for the entire group of pictures (GoP) is determined, and QP for the GoP's I-picture and first P-picture is also determined. The GoP Target is fed into the next block for detailed bit allocation to pictures or to smaller basic units.

6.4 Basic Unit Bit Allocation

The "Basic Unit” With this approach, rate control may be pursued to different levels of granularity – such as picture, slice, MB row or any contiguous set of MBs. That level is referred to as a "basic unit" at which rate control is resolved, and for which distinct values of QP are calculated.

6.5 Motion Estimation and Compensation

The motion estimation unit is the first stage. The uncompressed video sequence input undergoes temporal redundancy reduction by exploiting similarities between neighboring video frames as used in [10]. Temporal redundancy arises since the differences between two successive frames are usually similar, especially for high frame rates, because the objects in the scene can only make small displacements. With motion estimation, the difference between successive frames can be made smaller since they are more similar. Compression is achieved by predicting the next frame relative to the original frame. The predicted data are the residue between the current and reference pictures, and a set of motion vectors which represent the predicted motion direction. The process of finding the motion vector is optimal or suboptimal depending on the block matching algorithm chosen. Since the correlation between successive frames is inherently very high, the compression in this stage has large impact on the overall performance of the whole system. The motion predicted frames are usually called P-frames (Predicted frames).

7. Simulation Results

The proposed video-compression (VBR) system was developed by using the rate control and frame dropping to keep the output bit rate within target bit rate and the VBR, is adapted to be used as a CBR, see Fig. 4. The codec system (encoder and decoder) has been given input with varying parameters such as quantization matrix, variation in quantization of I-frame and P-frame, and drops the frame according to the error between the current frame and previous frame. The test operation is implemented on video sample with frame size 320 × 240 and the frame rate is 30 frames per second. This video sample contains two types of frames, reference frame (I-frame) and predictive frame (P-frame). The video sample is compressed with two types of systems. The first system is VBR system, which uses constant quality for all video sequence so that the output bit rate is variable with time. The system is tested on these two samples with quality values of (90, 70, 50, 30, and 10). The second system is CBR system, at which the encoder uses different quality for frames to match the encoded bit rate with target bit rate. With the above two systems, the number of frames tested in one GoP is 5, 10, 15, or 30 frames. The performance of the proposed compression system was evaluated by using a popular quantitative measure of image quality known as the peak signal-to-noise ratio (PSNR), which is defined as:

$$PSNR(dB) = 10\, log_{10} \frac{(2^n - 1)^2}{MSN} \tag{4}$$

Where n is the number of bits per image sample and MSE is the Mean Squared Error between the distorted frame and the original frame. The mean square error is found by taking the error squared divided by the total number of pixels in the frame.

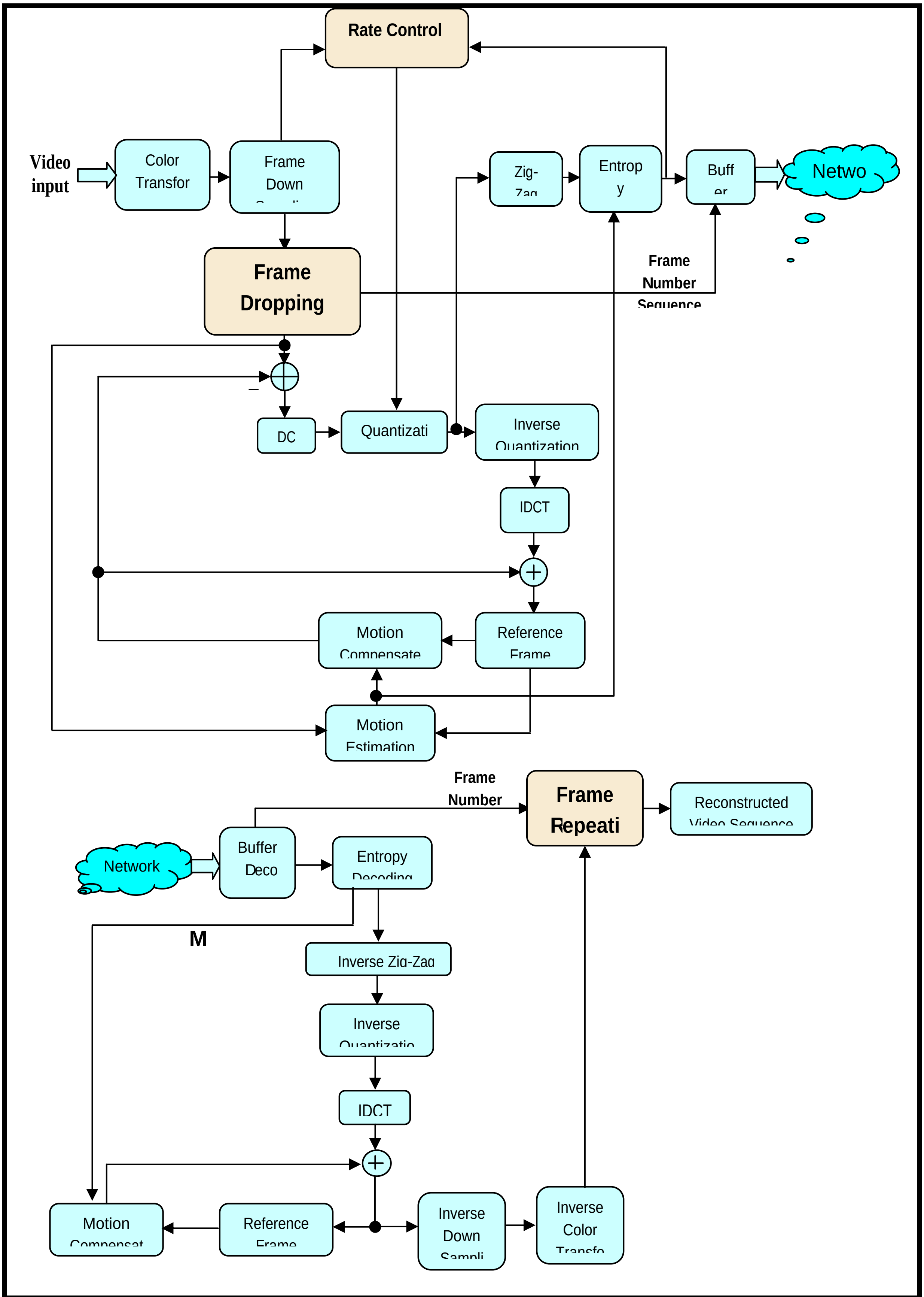


Figure (4) Block diagram of the proposed system

$$MSE = \frac{1}{W \times H} \sum_{r=0}^{H-1} \sum_{c=0}^{W-1} \left(I(r,c) - \hat{I}(r,c)\right)^2 \qquad (5)$$

Where *I(r,c)* is the pixel value of the original frame at the (r,c) location, *I(r,c)* is the value pixel of the reconstructed frame at the same location (r,c), H is the number of line, and W is the number of pixel per line. PSNR is relatively easy to calculate and provides a “rough” approximation to the visual quality of the video frame. In general, a high PSNR indicates a “high-quality” frame.

Quality degradations due to, for example, high compression and/or transmission errors, lead to a drop in PSNR. It is possible to obtain an approximate comparison of the quality of two video sequences by comparing the PSNR of the frames in each sequence, relative to the original video sequence. Calculating the average PSNR of all the frames in a video sequence gives a measure in dB of the “quality” of the sequence. While Compression Ratio (CR) is the degree of video file size reduction due to compression process. This ratio represents the size of the original uncompressed video to the size of compressed video.

$$CR = \frac{\text{Uncompressed file size}}{\text{Compressed file size}} \qquad (6)$$

The t esting table (Table 1) of video compression system (VBR system) consists of five main parts, the first is total bit rate of uncompressed video and numbers of frame transmit per second. The second part is number of frame per GoP, the third part is quality, the fourth part is the size of compressed video in seconds (or bit rate), and the final part is the compression ratio.

- This table shows the relationship between the quality and bit rate, when the quality increases, the bit rate increases as shown in Fig. 5.

- This table shows the effect of quality on the compression ratio (CR). By increasing the quality, the compression ratio decreases. With quality=90, the CR=23.0814, and with quality =10, the CR =54.6200 as shown in Fig. 6.

- This table shows also the relation between the number of frames per GOP and compression ratio, when the number of frames per GoP is maximum (30) the compression ratio is high (39.3809), and when the number of frames per GoP is minimum (5) compression ratio is lower (22.5931) as shown in Fig. 6.

- The table shows the effect of number of GoPs on bit rate, when the number of frames per GoP is maximum (30), the bit rate is small (1.404129) Mbps and when the number of frames per GoP is minimum (5), the bit rate is larger (2.447472) Mbps as shown in Fig. 5.

- The compression ratio of the video compression system can be achieved about 30:1.

Table (1) Testing Table for VBR system

Original data rate (Mb/s)	No. of Frame per GOP	Quality	Bit rate (Mb/s)	CR	PSNR (dB)
55.296	30	90	2.39573	23.0814	29.90
		70	1.59835	34.5955	28.93
		50	1.40412	39.3809	28.10
		30	1.19718	46.1884	26.84
		10	1.01237	54.6200	24.84
	15	90	2.74561	20.1397	31.72
		70	1.82745	30.2584	30.34
		50	1.58350	34.9199	29.27
		30	1.32177	41.8345	27.67
		10	1.07871	51.2609	25.29
	10	90	3.08366	17.9319	33.18
		70	2.04682	27.0155	31.36
		50	1.75649	31.4808	30.07
		30	1.44426	38.2865	28.19
		10	1.14363	48.3513	25.53
	5	90	4.42440	12.4980	36.22
		70	2.98255	18.5398	33.09
		50	2.44747	22.5931	31.35
		30	1.84956	29.8967	28.94
		10	1.33580	41.3954	25.87

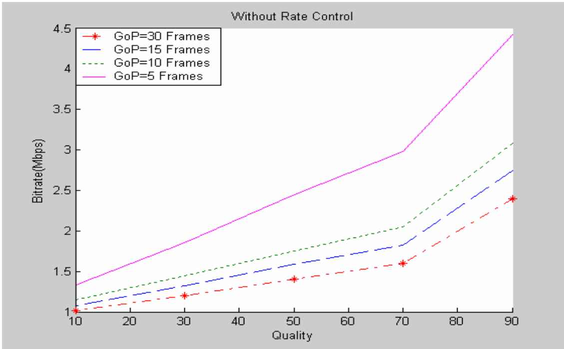


Figure (5) Bit rate versus the different quality

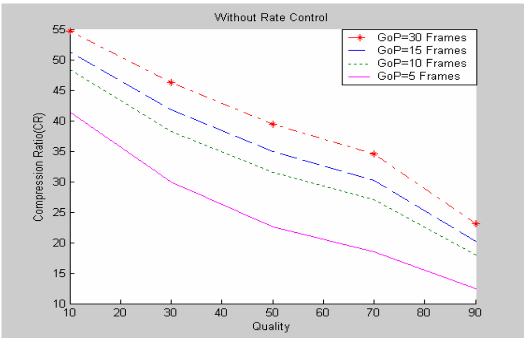


Figure (6) Compression Ratio versus the different quality

The second system is compression video system with rate control (CBR system), table (2) shows the testing result of this system, this table consists also the same parameters of the previous system but the Bit rate of this system is controlled to be equal or smaller than Threshold that is estimated according to channel bandwidth and decoder buffer size:

1. Table (2) shows also the bit rate equal to or smaller than (2 Mbps) for video sequence.
2. The testing table shows the range of the compression ratio in the first video sample between (30 to 37) according to number of frames per GoP.

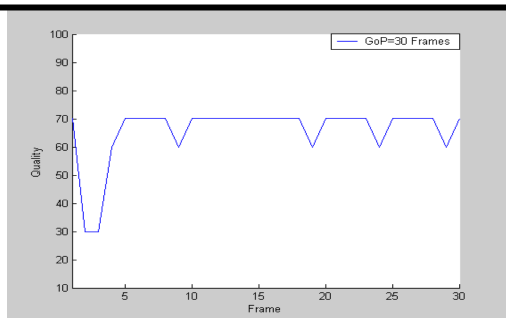
Table (2) Testing Table for CBR system

Original data rate (Mb\s)	Number of Frame per GOP	Bit rate (Mb/s)	CR	PSNR (dB)
55.296	30	1.475310	37.4809	28.47
	15	1.683894	32.8382	29.70
	10	1.735818	31.8559	30.60
	5	1.817670	30.4214	30.77

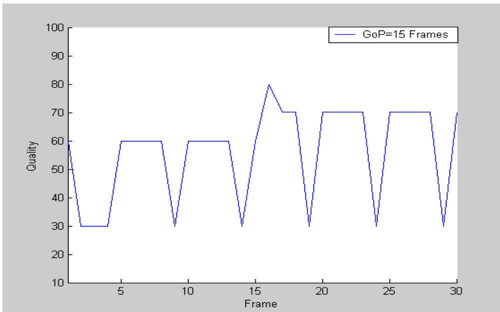
The last system is video compression with rate control and frame dropping, table (3) shows the results of this system. This system is designed to drop the frames according to the difference between the current frame and the previous frame. And in the receiver side the dropped frames are reconstructed by displaying the previous frame (reference frame). So in this system the bit rate and compression ratio are actually greater than the bit rate and compression ratio of previous system, because the number of frames that are displayed at the receiver side is more than the frames that transmitted from sender. But that isn't happen all times. The Fig. 7 shows the quality variation within time for frame sequence in one second for CBR system.

Table (3) Testing Table for CBR system and frame dropping

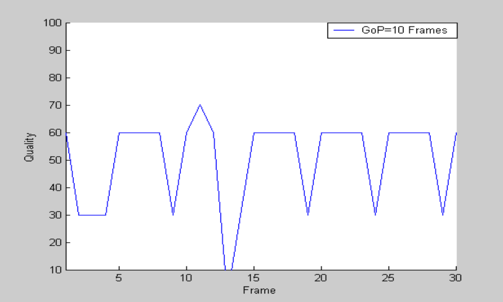
Original data rate (Mb\s)	Number of Frame per GOP	Bit rate (Mb/s)	CR	PSNR (dB)
55.296	30	1.462369	37.8126	29.7439
	15	1.676476	32.9835	31.2461
	10	1.762466	31.3742	31.7505
	5	1.819940	30.3834	32.9889



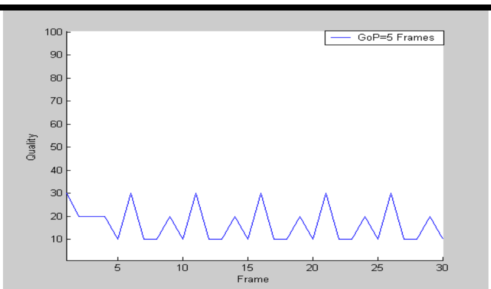
(a)



(b)



(c)



(d)

Figures (7) The quality variation for CBR system

(a) 30 frames per GoP (b) 15 frames per GoP(c) 10 frames per GoP (d) 5 frames per GoP

8. Conclusions

The results of VBR system show that the number of frames in GoP has significant effect on performance of decoded video sequence, bit rate, and compression ratio. When the number of frames in GoP is decreased, the coded bit rate increases and compression ratio decreases, but this has the advantage that it significantly affect the error in decoded video sequence. In fact the error will decrease when there is a clear decreasing in the number of frame in GoP. And another advantage of this system is that the quality of all frames in one second is the same, which decreases the error in the reconstructed video. The results of CBR system indicates that the effect of rate control on the quality. The output rate is adjusted by increasing the quantization step size, and variation the quality. Other encoding parameter has less effect on the coded bit rate; the size of each GoP affects the bite rate to a lower extent. The results also show that when frame dropping occurs in the transmitter side, it may cause an increase of the number of displayed frames at the receiver side, which might be more than the frames rate. This type of rate shaping is very effective because it removes the repeated frame in encoder side and it's reconstructed it in the decoder side. Finally, the proposed system shows that the performance of CBR system is better than VBR in term of controlling the bit rate and avoiding the packet loss over network, this can achieved by keeping the bit rate within the demand rate.

9. References

[1] Z. Chen, K. N. Ngan, “Recent advances in rate control for video coding”, Signal Processing: Image Communication,, Vol. 22, No. 1, pp. 19–38, 2007.

[2] A. Ortega, “Variable bit-rate video coding, in: M.-T. Sun, A.R. Reibman (Eds.), Compressed Video over Networks”, Marcel Dekker, New York, NY, pp. 343–382, 2000.

[3] A. Ortega, K. Ramchandran, “Rate-distortion methods for image and video compression”, IEEE Signal Process. Magazine, pp. 23–50, 1998.

[4] J. Lee, B.W. Dickinson, Rate-distortion optimized frame type selection for MPEG encoding, IEEE Trans. Circuits Systems Video Technol. Vol. 7, pp. 501–510, 1997.

[5] D. Wu, Y. T. Hou, W. Zhu, Y. Zhang, and J. M. Peha, “Streaming Video over the Internet: Approaches and Directions”, IEEE Transactions On Circuits And Systems for Video Technology, Vol. 11, No. 3, pp. 282-300, 2001.

[6] S. Ma, W. Gao, and Y. Lu, “Rate-Distortion Analysis for H.264/AVC Video Coding and its Application to Rate Control”, IEEE Transactions On Circuits And Systems for Video Technology , Vol. 15, No. 12, pp.1533-1544, Dece. 2005.

[7] A. Engman, “High Performance Video Streaming on Low End Systems”, M.Sc. Thesis in Electrical Engineering, Telia Research AB, Nattjanster Division, Alexander Engman KTH, Nov. 1999.

[8] H. Yu, Z. Lin, and F. Pan, “Applications and Improvement of H.264 in Medical Video Compression”, IEEE Transactions on Circuits and Systems-I: Regular Papers, Vol. 52, No. 12, Dec. 2005.

[9] G. Sullivan, T. Wiegand and K.P. Lim, "Joint Model Reference Encoding Methods and Decoding Concealment Methods " JVT-I049, San Diego, Sep. 2003.

[10] L. M. Ho, “Variable Block Size Motion Estimation Hardware for Video Encoders”, M.Phil. Thesis. The Chinese University of Hong Kong, Hong Kong, Nov. 2006.

الخلاصة

هذه الورقة اقترحت تقنية السيطرة على ارسال الاشارة الفيديوية كمعدل ارسال ثابت بدلا من المتغير عبر قنوات الاتصالات. حيث استخدمت للتعامل مع صيغة MPEG-2 كمعدل ارسال ثابت. وللحفاظ على جودة الاشارة الفيديوية مع معدل ارسال يتناسب وسعة القنوات المستخدمة، تم استخدام مسيطرات مبنية على اساس التغذية المعرفية العكسية. ان استخدام التقنية المقترحة اعطى نتائج جيدة من ناحية الاداء مقارنة بتلك التقنيات التي تعتمد على معدل ارسال متغير. فمسألة التزاحم في قنوات الارسال نتيجة تغير معدلات الارسال يمكن تلافيها وبالتالي تجاوز الاخفاق في الاتصال. تم تنفيذ المنظومة بأستخدام برنامج Matlab (Ver. 6.5).