Constant Quality Video Encoding*

İsmail Dalgıç Fouad A. Tobagi

Department of Electrical Engineering
Stanford University
Stanford, CA 94305-4055

ABSTRACT: *Lossy video compression algorithms, such as those used in the H.261 and MPEG standards, result in quality degradation seen in the form of tiling, edge busyness, and mosquito noise. The number of bits required to encode a scene so as to achieve a given quality objective depends on the scene content; the more complex the scene is, the more bits are required. Therefore, in order to achieve a given video quality at all times, the encoder parameters must be appropriately adjusted according to the scene content. In this paper, we propose a video encoding scheme which maintains the quality of the encoded video at a constant level. This scheme is based on a quantitative video quality measure, and it uses a feedback control mechanism to control the parameters of the encoder. We evaluate this scheme by applying it to test sequences, and compare it with Constant Bit Rate and Open-Loop Variable Bit Rate schemes in terms of quality and rate. We show that our scheme achieves better quality than the other two schemes for a given total number of bits, particularly when the content has scene changes over time.*

I. Introduction

In order to achieve high compression rates, today's prominent video encoding standards, such as H.261 and MPEG, are based on lossy video compression algorithms. Such loss results in tiling, edge busyness, and mosquito noise [1] in the encoded video, thereby reducing the perceived quality. The number of bits required to encode a scene so as to achieve a given quality objective depends on the scene content; the more complex the scene is, the more bits are required. Therefore, in order to achieve a constant video quality over time, the encoder parameters (in particular, the quantization scale) must be appropriately adjusted according to the scene content.

Currently, there exist two schemes for setting the encoder parameters. The first scheme is called *Constant Bit Rate (CBR) encoding*, where the rate of the encoded video is kept constant at all times by dynamically adjusting the quantizer scale. This is achieved by placing a buffer at the output of the encoder, and using the buffer occupancy level as feedback to dynamically adjust the quantization scale. This scheme is motivated by the fact that some communications technologies, such as ISDN, as well as some storage technologies, such as CD-ROMs, are able to accommodate only constant bit rate streams. It is important to note that in a CBR encoded video stream, the quality varies in time, since the quantizer scale is controlled to achieve a constant bit rate regardless of the scene complexity.

The second scheme is called *Open-Loop Variable Bit Rate (OL-VBR) encoding*, whereby the quantization scale is simply kept at a constant value at all times [2, 3]. The motivation behind this scheme is to presumably provide a more consistent video quality compared to the CBR encoding by taking advantage of the variable bandwidth (and hence, the statistical multiplexing capability) supported by today's LAN technologies, by the newly emerging ATM technology, and by some storage technologies such as magnetic disks. Indeed, with OL-VBR encoding, a more complex scene is encoded using more bits; thus, the quality is less variable in time compared to CBR encoding. Nevertheless, it can be shown that there are still variations in quality.

In order to take full advantage of those technologies that allow statistical multiplexing, video must be encoded at a constant quality at all times, thus producing only as many bits as needed to maintain a given quality objective. Therefore, the objective of our research is to devise a scheme for encoding video at a constant quality. In order to devise a constant quality video encoding scheme, as well as to evaluate the performance of the video encoding schemes, a quantitative video quality measure is required. We use in our study such a measure that has been developed at the *Institute for Telecommunication Sciences (ITS)*[4]. This measure is described in Section 2. In Section 3, the CBR, OL-VBR, and Constant Quality video encoding schemes are described. Section 4 presents an evaluation of these schemes by applying them to test sequences, and Section 5 contains some concluding remarks.

II. ITS Quantitative Video Quality Measure

A quantitative video quality measure has been designed at the ITS that agrees closely with quality judgements made by a large number of viewers. This measure is currently being considered for inclusion in the Digital Video Teleconferencing Performance Standard by the ANSI Accredited Standards Committee T1, Working Group T1A1.5.

To design this measure, the authors first conducted a set of subjective tests in accordance with CCIR Recommendation

*This work was supported in part by NASA under grant NAGW-419, by NSF under grant NCR-9016032, and by Pacific Bell.

500-3 [5], which specifies viewing conditions, rating scales, etc. The viewers were shown a number of original and degraded video pairs, each of them 9 seconds long, and they were asked to rate the difference between the original video and degraded video as either imperceptible (5), perceptible but not annoying (4), slightly annoying (3), annoying (2), or very annoying (1).

As described in [4], the quantitative measure is a linear combination of three quality impairment measures. Those three measures were selected among a number of candidates such that their combination matched best the subjective evaluations. The correlation coefficient between the estimated scores and the subjective scores was 0.94, indicating that there is a good fit between the estimated and the subjective scores. The standard deviation of the error between the estimated scores and the subjective scores was 0.4 impairment units on a scale of 1 to 5; thus, differences below 0.4 should not be considered significant.

The quantitative measure is based upon two quantities, namely, *spatial information* (SI) and *temporal information* (TI). The spatial information for a frame F_n is defined as

$$SI(F_n) = STD_{space}\{Sobel[F_n]\},$$

where STD_{space} is the standard deviation operator over the horizontal and vertical spatial dimensions in a frame, and $Sobel$ is the Sobel filtering operator, which is a high pass filter used for edge detection [6].

The temporal information is based upon the motion difference image, ΔF_n, which is composed of the differences between pixel values at the same location in space but at successive frames (i.e., $\Delta F_n = F_n - F_{n-1}$). The temporal information is given by

$$TI[F_n] = STD_{space}[\Delta F_n].$$

Note that SI and TI are defined on a frame by frame basis. To obtain a single scalar quality estimate for each video sequence, SI and TI values are then time-collapsed as follows. Three measures, m_1, m_2, and m_3, are defined, which are to be linearly combined to get the final quality measure. Measure m_1 is a measure of spatial distortion, and is obtained from the SI features of the original and degraded video. The equation for m_1 is given by

$$m_1 = RMS_{time}(5.81\left|\frac{SI[O_n] - SI[D_n]}{SI[O_n]}\right|),$$

where O_n is the n^{th} frame of the original video sequence, D_n is the n^{th} frame of the degraded video sequence, and RMS denotes the root mean square function, and the subscript *time* denotes that the function is performed over time, for the duration of each test sequence.

Measures m_2 and m_3 are both measures of temporal distortion. Measure m_2 is given by

$$m_2 = f_{time}[0.108MAX\{(TI[O_n] - TI[D_n]), 0\}],$$

where $f_{time}[x_t] = STD_{time}\{CONV(x_t, [-1, 2, -1])\}$, STD_{time} is the standard deviation across time (again, for the duration of each test sequence), and $CONV$ is the convolution operator. The m_2 measure is non-zero only when the degraded video has lost motion energy with respect to the original video.

Measure m_3 is given by

$$m_3 = MAX_{time}\{4.23LOG_{10}(\frac{TI[D_n]}{TI[O_n]})\},$$

where MAX_{time} returns the maximum value of the time history for each test sequence. This measure selects the video frame that has the largest added motion. This may be the point of maximum jerky motion or the point where there are the worst uncorrected errors.

Finally, the quality measure $\hat{s}$ is given in terms of m_1, m_2, and m_3 by

$$\hat{s} = 4.77 - 0.992m_1 - 0.272m_2 - 0.356m_3.$$

III. Constant Bit Rate, Open-Loop VBR, and Constant Quality Video Encoding Schemes

We consider that video streams are encoded according to a coding standard such as H.261 [7] or MPEG [8]. In both MPEG and H.261, a frame is divided into a number of 16x16 "macroblocks." A macroblock consists of three "blocks," two for each color component, and one for the luminance. A Discrete Cosine Transform (DCT) is performed individually in each of the blocks. The resulting DCT coefficients are then quantized by using a quantizer scale q, which can take values between 1 and 31. Larger values of q correspond to coarser quantization, and hence, greater degradation in the quality (of course, smaller data rate as well). The resulting quantized DCT coefficients are then run-length and variable-length coded. Also, in both H.261 and MPEG, a macroblock can be coded differentially with respect to the previous frame; moreover, in MPEG, it can be coded differentially with respect to both a preceding and a following frame.

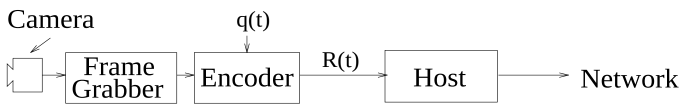


Figure 1: Block diagram of a station which encodes and sends a video stream

In Figure 1, we show the block diagram of a station which encodes and sends a video stream over a communications network. A frame is taken in the video camera, and sent as an analog signal to the frame grabber, where it is digitized. Then it is compressed by the encoder. The data produced by the encoder is passed to the host computer. If the network is of the circuit-switching type, then the encoder has to produce a constant bit rate video stream, and the host simply passes the stream at a constant rate to the network. If the network is of the packet-switching type, and thus is able to accommodate variable bit rate streams, then the encoder may produce either constant or variable bit rate video;

the host computer packetizes the encoded video stream, and sends it over the network. In this paper, we do not consider any particular network, and we assume that all of the video degradation is due to the encoding. In a forthcoming paper, we will evaluate the performance of various networking technologies for supporting video traffic.

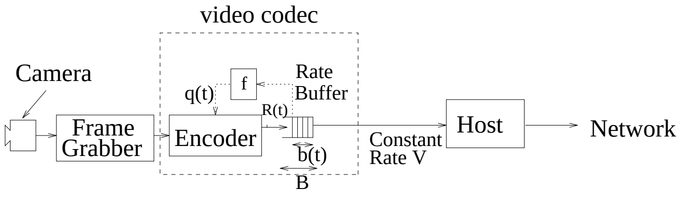


Figure 2: Block diagram of the sending station for CBR encoding

In the following, we describe the existing video encoding schemes, namely, CBR and Open-Loop VBR encoding, as well as our proposed Constant-Quality VBR encoding scheme.

(i) Constant Bit Rate (CBR) Encoding: In Figure 2, we show the block diagram of a station which encodes and sends a constant bit rate video stream. As shown in the figure, to generate a constant bit rate stream, a buffer is placed at the output of the encoder, which is drained at a constant rate; in order to ascertain that the buffer does not overflow or underflow, the buffer occupancy level is used as feedback to control the quantizer scale, q. Typically, the feedback function is chosen such that q is a linear function of the buffer occupancy (within the allowed limits) [9]. That is, if we denote the instantaneous rate at the output of the encoder at time t by $R(t)$, the channel rate by V, the buffer size by B, the buffer occupancy at time t by $b(t)$, and a constant by α, then the relationship between $q(t)$ and $b(t)$ can be written as follows:

$$\begin{aligned} q(t) &= \min\{\max\{\alpha b(t)/B, 1\}, 31\}, \\ b(t) &= \int_0^t (R(t) - V) dt. \end{aligned}$$

Clearly, in CBR encoding, the choice of the buffer size is a trade-off between delay and quality: If the buffer size is large, the encoder can handle variations in the scene complexities without compromising the quality, at the cost of a large delay. Conversely, if the buffer size is small, there will be less delay, but the encoder will frequently have to reduce the compression quality when there are complex scenes that require more bits to be compressed for a given quality. Furthermore, note that when a scene is too complex, even when q=31, the encoder may produce bits at a higher rate than the target constant bit rate; when this situation persists long enough, the buffer may overflow, thereby causing some macroblocks to be completely lost, which leads to a large quality degradation. Clearly, the likelihood of such an overflow occuring is greater when the buffer size is smaller.

(ii) Open-Loop VBR Encoding: In Figure 3, we depict the block diagram of the sending station for Open-Loop VBR. As shown in the figure, the quantizer scale q is simply kept at a constant value, q_0. Thus, the feedback loop is "open," and hence the name Open-Loop VBR.

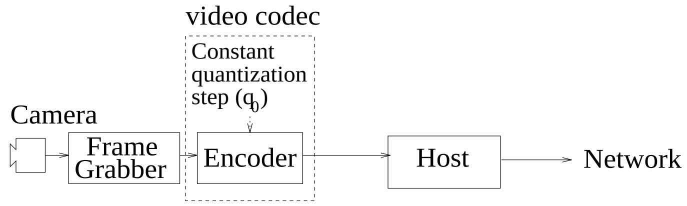


Figure 3: Block diagram of the sending station for Open-Loop VBR encoding

(iii) Constant-Quality VBR (CQ-VBR) Encoding: To encode video streams at a constant quality $\hat{s}_{target}$, we have devised the following scheme. We measure the quality $\hat{s}$ at every frame interval, and use the difference ($\hat{s}_{target}$-$\hat{s}$) as feedback to adjust q. The block diagram of the sending station for this scheme is depicted in Figure 4. Recall that $\hat{s}$ as defined above had a single value for a range of frames; in particular, for the test sequences used in ITS, this range was 9 seconds, or 270 frames (i.e., the entire length of the test sequence). In order to measure the quality for every frame k, we define $\hat{s}(k, w)$ as the quality measured from frame $\max\{k - w + 1, 1\}$ to frame k. Thus, part of the design of the feedback function is to choose an appropriate estimation interval w to measure quality. Too small an estimation interval could cause overreaction, and too large an estimation interval could cause the quality of the recent frames to be washed out by the past ones, leading to instability. Similarly, the feedback function f must also be chosen carefully, so as to cause neither instability, nor a slow response time. In Section IV-C, we describe the design of our feedback function, which yielded satisfactory results.

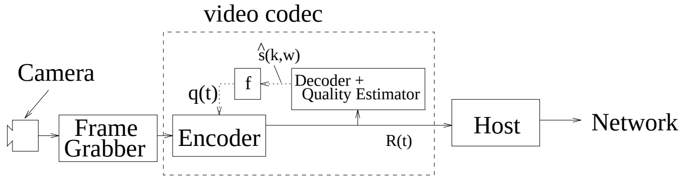


Figure 4: Block diagram of the sending station for Constant Quality VBR encoding

IV. Evaluation of the Constant Bit Rate, Open-Loop VBR, and Constant Quality VBR Encoding Schemes

In our evaluation, we have used three different video sequences, namely, *"miss america," "susie,"* and *"table tennis." miss america* and *susie* are typical video conferencing type sequences, with head and shoulders, containing relatively small amount of motion. Between the two, *susie* has more details, and contains somewhat more motion. *table tennis* is a typical sports type sequence, with lots of rapid movement, a zoom-out, and two scene cuts. Each sequence is 5 seconds long. We have encoded the sequences at 30 frames per second using a software H.261 encoder, developed by the Portable Video Research Group (PVRG) at Stanford University [10].

A. CBR Encoding

We have encoded the three video sequences at various values of V and B, to determine the effect of those two parameters on quality and delay. In Figure 5, we plot $\hat{s}$ versus B/V for the three video sequences for V=1536 kb/s. As the figure shows, when B/V is greater than about 100 ms, $\hat{s}$ is fairly insensitive to the buffer size. When B/V is less than about 100 ms, $\hat{s}$ gets smaller when the buffer size is chosen to be smaller. When B/V is chosen to be around 30-50 ms, buffer overflows occur, thus causing a sharp drop in quality. Similar results were observed for all the rates that we considered (ranging from 300 kb/s to 1.5 Mb/s).

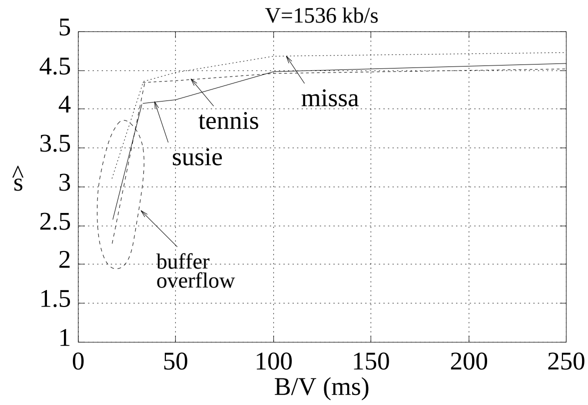


Figure 5: CBR: $\hat{s}$ versus B/V.

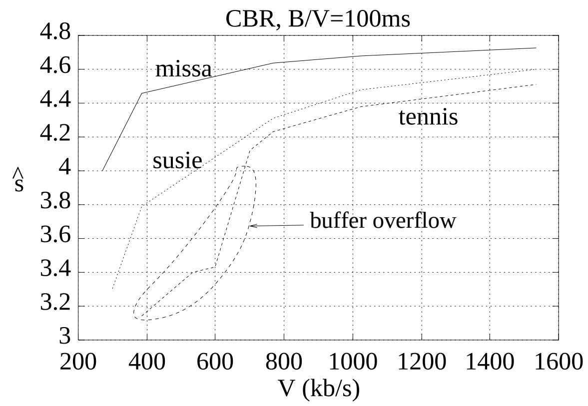


Figure 6: CBR: $\hat{s}$ versus V.

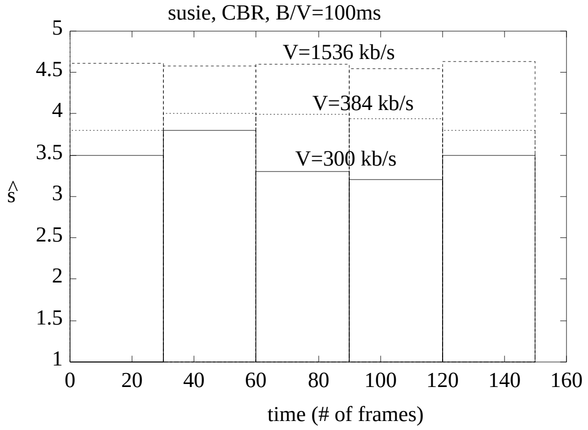
In Figure 6, we plot $\hat{s}$ versus V for the three video sequences, for B/V=100 ms (different buffer sizes yield similar results). As the figure indicates, *miss america* sequence reaches a plateau of quality at about 400 kb/s, whereas *susie* and *table tennis* sequences reach a similar plateau at about 800–1000 kb/s. It is important to note that for lower rates, there is a significant difference in quality from sequence to sequence; as the rate increases, the differences in quality diminish. The differences in quality can be explained by the contents of the sequences. In general, if a sequence contains more motion, then more bits are produced for a given q; thus, in CBR encoding, larger q values have to be used for sequences that contain more motion. Furthermore, the effects of quantization are more apparent when the original video image has large features, and sharp edges. The *miss america* sequence contains very small motion, and the edges in it are not very sharp; the *susie* sequence contains somewhat more motion, and it has larger features with sharper edges; and the *table tennis* sequence contains rapid motion (particularly during the zoom-out period), and large features at the beginning, when the camera is zoomed in on the paddle. Therefore, for any given rate, encoding *miss america* results in the highest quality, followed by *susie*, and then *table tennis*. Also note that there is a sharp drop in quality in the *table tennis* sequence at rates less than about 700 kb/s due to the fact that at such rates, buffer overflows occur for that particular sequence during the zoom-out period. This is because the zoom-out results in a large amount of difference between successive frames, thus producing a large number of bits even when $q = 31$.

We have also investigated the variation of quality within the sequences. In Figure 7, we plot $\hat{s}$(measured in 1 second intervals) as a function of time for *susie* and *table tennis* sequences. As the plot indicates, at V=1.5 Mb/s, quality is fairly constant within the scenes. As we get to lower rates, the quality starts to vary more in time. This is especially apparent in the *table tennis* sequence. In this sequence, during the first second, the camera is zooming in on the paddle; thus, large features are shown, for which the quantization results in a large amount of distortion. During second 2, the zoom-out begins, and as the features get smaller, the quality increases. The zoom-out continues during second 3, and eventually the rate buffer gets full, thereby causing some quality degradation. At the end of second 3, a scene cut occurs, and during seconds 4 and 5, there is no camera movement or zooming; thus, there is less motion, and the quality is better. For the *miss america* sequence, the quality remained fairly constant at all times for all the data rates that we considered, as expected from the consistent content of this sequence.

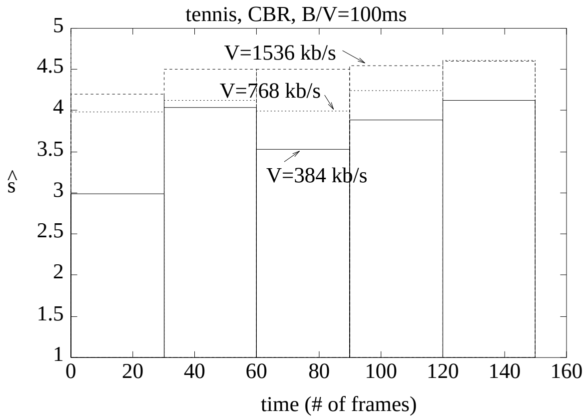
B. Open-Loop VBR Encoding

In Figure 8 we plot $\hat{s}$ versus q_0 for the three video sequences. As the figure indicates, especially at larger values of q_0, quality varies by a significant amount between *susie* and the other two sequences; thus, the quality for OL-VBR is a function of the scene content. The reason for the lower quality of the *susie* sequence is because it contains larger features with sharp edges; thus, quantization effects are more apparent.

In Figure 9, we plot the total number of bits required to encode each sequence as a function q_0. As shown in the



(a) susie



(b) table tennis

Figure 7: CBR: $\hat{s}$ versus time

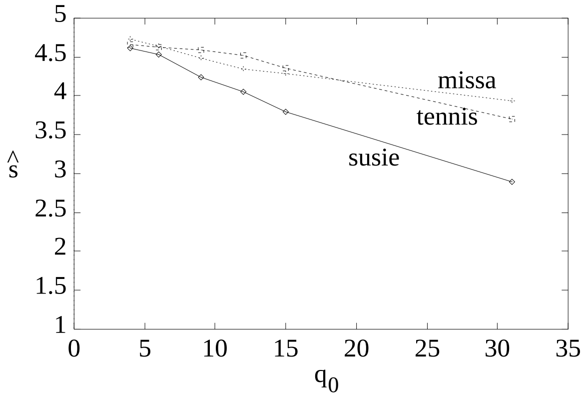


Figure 8: OL-VBR: $\hat{s}$ versus q_0.

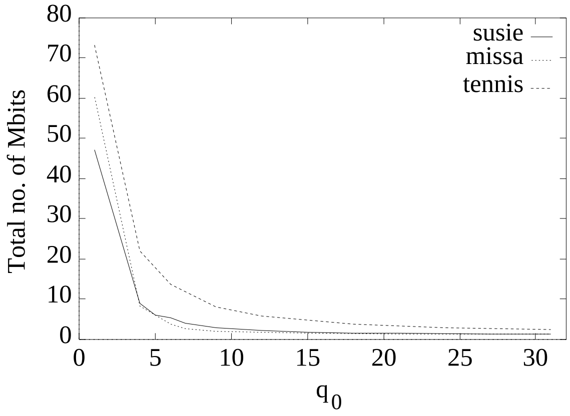


Figure 9: OL-VBR: Total number of bits versus q_0.

figure, the *table tennis* sequence is encoded using more bits compared to the other two sequences for any given q_0. This is because of the greater amount of motion in the *table tennis* sequence compared to the other two sequences.

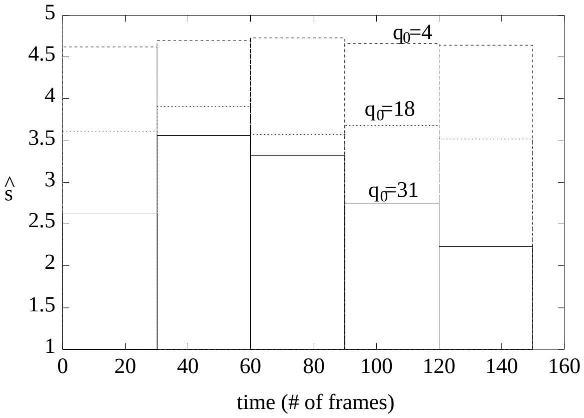
Finally, to illustrate the variation of quality within the sequences, in Figure 10 we plot $\hat{s}$(measured in 1 second intervals) as a function of time for *susie* and *table tennis* sequences. As shown in the figure, although the quality is fairly constant within the scenes for small values of q_0, it becomes highly variable as q_0 gets larger. For the *table tennis* sequence, the low quality during second 1 can be again attributed to the large features shown during that time. For the *susie* sequence, the increase in quality during seconds 2 and 3 can be attributed to the fact that at that period, there is a rapid head movement in the scene, and as a result, the original scene is blurred significantly. As mentioned above, when the edges in the original image are not sharp, the effects of quantization are less apparent. For the *miss america* sequence, as expected, we have observed that the quality again remained fairly constant at all times for all values of q_0.

C. Constant Quality VBR Encoding

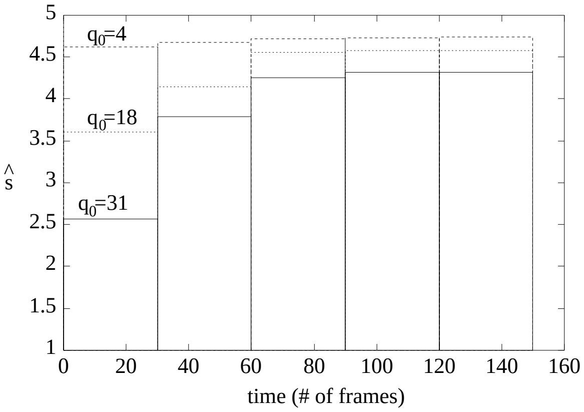
To achieve constant quality, we have designed a feedback function of the PID (Proportional, Integral, Derivative) type [11] since this type of feedback function is known to be effective for a wide range of systems. If we define $e(k) = \hat{s}(k,w) - \hat{s}_{target}$, and we let T be the sampling period of the system (in our case, time between successive frames), then, the PID feedback function is given by:

$$\begin{aligned} q(k+1) &= K_p e(k) + K_p \frac{T}{T_I} \sum_{i=1}^{k} e(i) + q(0) + \\ & K_p \frac{T_D}{T}[e(k) - e(k-1)]. \end{aligned}$$

In this equation, the first term is proportional to the error, the second term is proportional to the integral of the error, and the last term is proportional to the derivative of the error; hence the name PID control. The coefficients K_p, T_I, and T_D, as well as the estimation interval w, are our design variables. Since it is not practical to determine the exact dynamic model for the video encoder, we employ Ziegler-Nichols PID tuning using the stability limit method [11] to determine the values for K_p, T_I, and T_D. The method works as follows. The system is first controlled using only proportional control. The gain, K_p, is increased until continuous oscillations result, at which point the gain, K_u, and the oscillation period, P_u, are recorded. The PID gains are then determined as follows: $K_p = 0.6K_u$, $T_I = P_u/2$, and $T_D = P_u/8$.



(a) susie

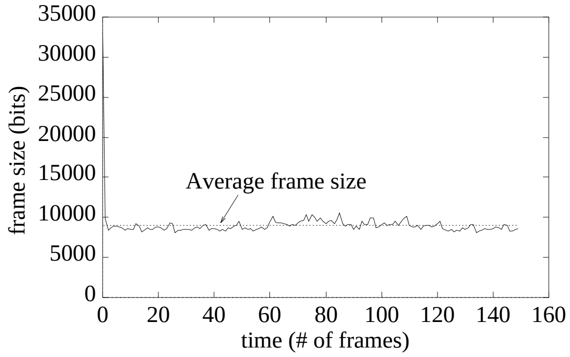


(b) table tennis

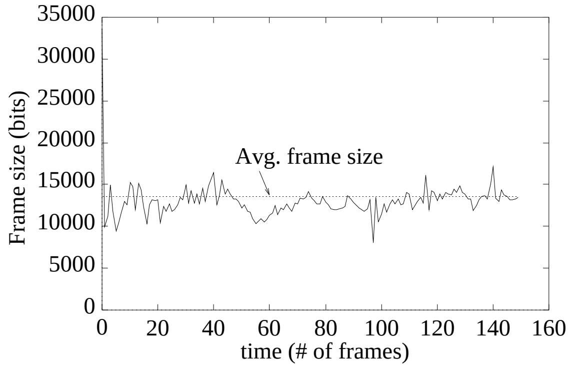
Figure 10: OL-VBR: $\hat{s}$ versus time

To determine the estimation interval w, we have experimented with various values ranging from 3 to 50, and determined that smaller estimation intervals lead to better stability and faster response time. Thus, we chose $w = 3$ frames, the smallest estimation interval for which $\hat{s}$ is meaningfully defined. For that choice of w, we have found that K_u=20, and $P_u = 4T$. Thus, $K_p = 12$, $T_I = 2T$, and $T_D = 0.5T$. Using this feedback function, we have observed that for the target quality values of 3.5, 4, and 4.5, for all three video sequences, our scheme matched very closely the target value at all times. The largest deviation we have observed was around 0.2, within the error margin of the quality measure.

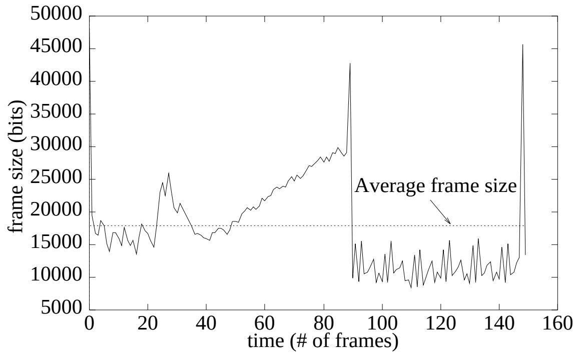
In Figure 11, we show the frame size versus time for the three video sequences, all constant-quality encoded at $\hat{s}_{target}$=4.0. As expected from the scene contents, the frame sizes in *miss america* and *susie* are fairly constant with respect to time. However, for the *table tennis* sequence, the frame size varies significantly in time. In particular, between frames 40 and 90, the effect of the zoom-out is very apparent; as the camera zooms out, the scene becomes more detailed, and the difference between the successive frames increases; thus, more bits are required to encode each successive frame. There is a scene cut at frame 89, and the scene between frames 90 and 147 contains less motion, thus requiring fewer bits to attain the same quality. Note that, when this sequence is encoded at a constant bit rate, there is less variation in the frame sizes, and those variations are short-term; thus the frame size fluctuates around the average value (determined by the rate V).



(a) miss america



(b) susie



(c) table tennis

Figure 11: CQ-VBR: Frame size versus time ($\hat{s}$=4.0).

D. Comparison of the Schemes

In order to compare the three encoding schemes, we have first encoded the three sequences using the CQ-VBR scheme for $\hat{s}_{target}$={4.0,4.5}, and recorded the number of bits resulting from encoding each sequence. We have then encoded the sequences using the CBR and OL-VBR schemes, choosing the encoding parameters (V and q_0, respectively) such that the resulting number of bits for each sequence was the same as in CQ-VBR encoding. In Table 1, for each sequence we show the number of bits and the quality for the three encoding schemes. As the table indicates, the quality of the

	miss america		susie		table tennis	
Quality Target for CQ-VBR	4.5	4.0	4.5	4.0	4.5	4.0
number of kilobits per sequence	1950	1350	4590	1950	5500	2680
Quality for CBR at the same no. of bits	4.5	4.0	4.4	3.8	4.4	3.4
Quality for OL-VBR at the same no. of bits	4.5	4.0	4.5	3.9	4.5	3.9

Table 1: Comparison of qualities for the CQ-VBR, CBR, and OL-VBR encoding schemes

		$\hat{s}$ for missa	$\hat{s}$ for susie	$\hat{s}$ for tennis	overall $\hat{s}$
Total size = 6 Mbits	CQ-VBR	4.0	4.0	4.0	**4.0**
	CBR	4.5	3.8	3.2	**3.4**
	OL-VBR	4.1	3.4	4.0	**3.7**

Table 2: Comparison of quality for the three schemes when the three sequences are concatenated

individual sequences were generally comparable for the three schemes when they are encoded using the same number of bits. This is expected given the sequence contents, since the variations within the sequences are relatively small (particularly for *miss america* and *susie*, as indicated by Figure 11).

However, note that the total number of bits required to achieve the same quality target varies significantly among the sequences. Therefore, in order to maintain a given quality target for all types of scenes, CBR encoding has to use a high V at all times, thus using more bandwidth than needed for less complex scenes. A similar argument applies to the OL-VBR scheme. On the other hand, CQ-VBR encoding offers the ability to use no fewer and no more bits than that required to meet a target quality objective for a given scene. To illustrate this point, we have experimented with a 15-second sequence which was formed by concatenating *miss america*, *susie*, and *table tennis* sequences (in that order). We have encoded this sequence at a constant quality for $\hat{s}_{target}$=4.0, which resulted in a total of about 6 Mbits. We have then encoded the same sequence using the same total number of bits with the CBR and OL-VBR schemes (by appropriately choosing V and q_0, respectively). The resulting quality values are shown in Table 2, where we report the quality for the individual sub-sequences, as well as the overall 15-second sequence. As shown in the table, the variations in quality for the CBR scheme are significant (from 3.2 to 4.5), and the overall quality is only 3.4. For the OL-VBR scheme, there are also variations in quality, although, as expected, they are not as large as those in the CBR scheme. The overall quality for the OL-VBR scheme was 3.7, better than the overall quality for the CBR scheme, but not as good as the quality for the CQ-VBR scheme. As these results indicate, the CQ-VBR scheme will perform better compared to the other two schemes when the content includes scene changes over time.

V. Conclusions

We have proposed a Constant-Quality VBR encoding scheme, and shown that it achieves better quality than the existing schemes (i.e., CBR and OL-VBR) when the same total number of bits are used for encoding. This is particularly apparent when the scene content varies significantly in time (as would be the case with any given movie). Our evaluation with the three test sequences served to illustrate that there is a gain in using the CQ-VBR scheme. However, we believe that it is important to use a larger sample size to quantify this gain better; therefore, we are currently working on a similar evaluation using excerpts from movies which are of several minutes long.

References

[1] "Draft ANSI Standard Specification of Video Performance Terms and Definitions," ANSI T1A1.5/94-137 R1, Oct. 3, 1994.

[2] P. Pancha and M. El Zarki, "MPEG Coding for Variable Bit Rate Video Transmission," *IEEE Communications Magazine*, vol. 32, pp. 54–66, May 1994.

[3] H. S. Chin, J. W. Goodge, R. Griffiths, and D. J. Parish, "Statistics of Video Signals for Viewphone-Type Pictures," *IEEE JSAC*, vol. 7, no. 5, pp. 826–832, 1989.

[4] A. A. Webster, C. T. Jones, M. H. Pinson, S. D. Voran, and S. Wolf, "An Objective Video Quality Assessment System Based on Human Perception," in *SPIE Human Vision, Visual Processing, and Digital Display IV*, vol. 1913, (San Jose, CA), pp. 15–26, Feb. 1993.

[5] "CCIR Rec. 500-3, Method for the Subjective Assessment of the Quality of Television Pictures," CCIR, 1986, XVI'th Plenary Assembly, Volume XI, Part 1.

[6] A. K. Jain, *Fundamentals of Digital Image Processing.* Prentice Hall, Englewood Cliffs, NJ 07632, 1989.

[7] "Video CODEC for Audiovisual Services at p×64 kbit/s," ITU-T Rec. H.261, (Geneva, 1990).

[8] "MPEG: Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbits/s," International Organization for Standardization (ISO), Nov. 1991.

[9] "Description of the Reference Model 8," CCITT SG XV. Spec. Group on Coding for Visual Telephony, May 1989.

[10] The PVRG H.261 encoder is available via anonymous ftp from havefun.Stanford.EDU.

[11] G. F. Franklin, J. D. Powell, and M. L. Workman, *Digital Control of Dynamic Systems (second edition).* Addison-Wesley Publishing Company, 1990.