A Statistical Feedforward/Feedback Buffer Control for the Transmission of Digital Video Signals Compressed by DCT-Based Intrafield Coding

Seongkwon Jo and Yong Hoon Lee

Abstract— **A new buffer-control policy for intrafield coding of video signals is presented. This method employs two statistical bit rate predictors—feedback and feedforward. It uses a feedback predictor in stationary portions of an image sequence and a feedforward predictor at scene changes. It is shown that this buffer-control policy is reasonably simple to implement and can effectively control output bit rate even at scene changes.**

I. Introduction

The use of variable bit rate (VBR) coders for compressing video signals in a constant bit rate transmission environment requires a transmission buffer and buffer-control policies that adaptively control the output bit rate of the encoder. Buffer control directly influences picture quality, and has been investigated in many digital video compression systems [1]–[8]. Conventional buffer-control methods in [1]–[4] are based on observations of buffer occupancy; they control the bit rate in a feedback manner. More sophisticated feedback buffer-control methods, which have been recently proposed in [5] and [6], exploit appropriate statistical information about the encoder output along with buffer occupancy observations. It has been observed that statistical approaches provide some improvement in picture quality. One drawback of all feedback methods is their inability to control output rate at scene changes. In an effort to overcome this difficulty, feedforward buffer-control policies have been introduced in [7] and [8]. These methods, however, are considerably more difficult to implement than feedback methods, because they must produce several output bit streams having different bit rates.

In this correspondence, we introduce a new buffer-control method for intrafield coding of video signals. This method incorporates a feedback approach together with a simple feedforward approach and is called *statistical feedforward/feedback* buffer control. It employs two predictors—feedback and feedforward—for output bit rate prediction. The feedback prediction is used in stationary portions of an image sequence and the feedforward prediction whenever scene changes occur. Both predictors exploit the statistics of the encoder inputs and outputs; the feedback predictor is essentially the same as the one in [6], and the feedforward predictor uses the statistical relation between input picture complexity and output bit rate. The computational load required for the feedforward prediction is somewhat heavier than that for the feedback prediction. We shall show that this proposed policy is reasonably simple to implement and can effectively control the bit rate even at scene changes.

The buffer-control policy presented below uses the JPEG encoder [9] which is applied to intrafield coding of video signals. It is assumed that the rate is controlled once for each field.

Manuscript received February 9, 1994; revised July 24, 1995. This work was supported by GoldStar Co. The associate editor coordinating the review of this work and approving it for publication was Dr. Thomas L. Marzetta.

The authors are with the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Taejon, Korea (e-mail: yohlee@ee.kaist.ac.kr).

Publisher Item Identifier S 1057-7149(96)01787-3.

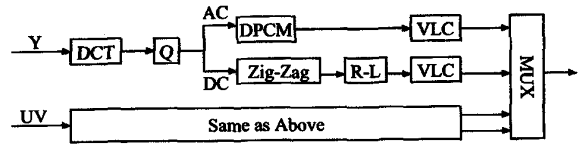

Fig. 1. JPEG baseline algorithm.

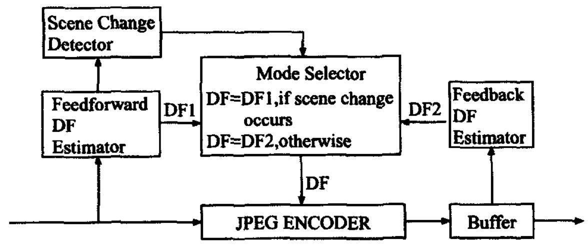

Fig. 2. Statistical feedforward/feedback buffer-control system.

II. Intrafield Coding Using the JPEG Baseline Algorithm

Fig. 1 illustrates the JPEG baseline algorithm proposed by ISO for still picture coding. This algorithm partitions the input picture (luminance and chrominance components) into 8×8 blocks, performs 2-D discrete cosine transform (DCT) on each block, and independently quantizes the DCT coefficients using a set of uniform quantizers. The quantization step size applied to each coefficient is determined from the contents of a 64–element quantization table presented in the system (two such tables for luminance and chrominance components are given). The quantized DC coefficients of DCT are encoded by interblock DPCM. Prediction errors are coded using a variable-length coding (VLC) technique. Quantized AC coefficients are converted into a 1-D array through zig-zag scanning, run-length (R-L) coded, and then encoded by VLC. Four Huffman tables are given in the JPEG system for VLC of luminance DC, AC, chrominance DC, and AC coefficients.

When the JPEG algorithm is applied to compress an image, the elements of the quantization tables are divided by a constant, which will be referred to as the *division factor* (DF) and then used as step sizes of the uniform quantizers. Note that DF and step sizes are inversely proportional to each other. Thus increasing DF results in an increase in the output bit rate, and *vice versa*. Throughout this paper the bit rate of the JPEG encoder will be controlled by adjusting DF. We shall use the same DF value for the luminance and chrominance quantization tables.

III. The Statistical Feedforward/Feedback Buffer Control

A system block diagram of the statistical feedforward/feedback buffer control is shown in Fig. 2. For a given field, two DF values are produced from the feedforward and feedback DF estimators, and one of them is chosen by examining whether scene change occurs or not. Details for each block are described below.

A. The Feedback DF Estimator

DF value is estimated following the procedure proposed in [6]. Such estimation requires a statistical bit rate curve called the $DF -$

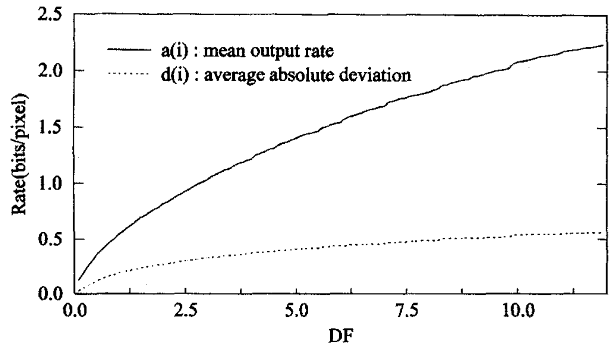

Fig. 3. DF-Rate curve showing the mean output rate and average absolute deviation of the output rate of a field as functions of DF.

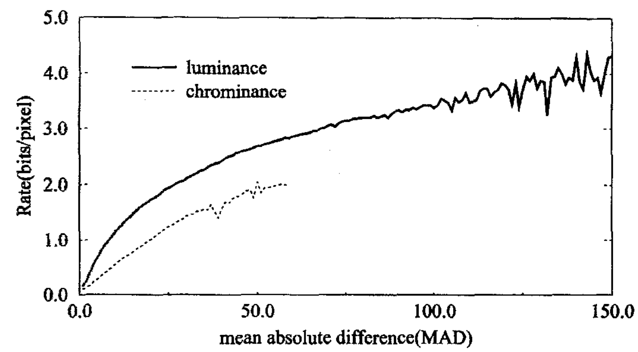

Fig. 4. Average output bit rate of an 8×8 block as a function of MAD.

$Rate$ curve, showing the mean and average absolute deviation of the output bit rate in bits per pixel (b/pixel) as functions of DF. Fig. 3 illustrates a $DF - Rate$ curve obtained by compressing eight CCIR 601 test images (Ping-Pong, Flower Garden, Suzie, Cheer Girls, Mobile & Calendar, Tempest, Football, and Bicycles) using the JPEG baseline algorithm. Here the DF values are increased from 0.1 to 12.0 with a 0.1 increment. The output bit rate can be predicted using the DF-$Rate$ curve. Let $n_k(m)$ be the bit rate of the k_{th} field when $DF = m$. If the value of $n_k(m)$ is given, then for any p, $p \neq m$, $n_k(p)$ is estimated as follows:

$$\hat{n}_k(p) = d(p)\frac{n_k(m) - a(m)}{d(m)} + a(p) \quad (1)$$

where $a(i)$ and $d(i)$, respectively, are the mean rate and the average absolute deviation of the rate corresponding to $DF = i$. In stationary portions of an image sequence $n_k(p) \approx n_{k+1}(p)$. Therefore, we predict $n_{k+1}(p)$ by using $\hat{n}_k(p)$. The feedback DF estimator predicts $n_{k+1}(p)$ for all p, $0.1 \leq p \leq 12$, and set $DF = p_0$ if $n_{k+1}(p_0)$ is closest to the transmission (target) bit rate. The estimation procedure is summarized as follows.

1) From the buffer, obtain bit rate $n_k(m)$ and the corresponding DF value m associated with the compressed field at time k.
2) Using (1), predict $n_{k+1}(p)$ for all p, $0.1 \leq p \leq 12.0$ (p is increased from 0.1 with an increment of 0.1).
3) Set DF=p_0 for some p_0 if $n_{k+1}(p_0)$ is closest to the transmission bit rate.

B. The Feedforward DF Estimator

The feedforward prediction of the bit rate is based on the fact that the amount of data produced by the JPEG encoder with a fixed DF value is roughly proportional to picture complexity. A simple measure used here for examining the complexity of each 8×8 block is the mean absolute difference (MAD) defined as

$$MAD = \frac{1}{7 \times 8}\left(\sum_{i=0}^{7}\sum_{j=1}^{7}|f(i,j) - f(i,j-1)| + \sum_{i=1}^{7}\sum_{j=0}^{7}|f(i-1,j) - f(i,j)|\right) \quad (2)$$

where $f(i,j)$ is the gray level of the ith row, jth column within an 8×8 block. The feedforward predictor utilizes the statistical relation between the rate and MAD. To obtain the statistics, evaluate the bit rate and MAD for every block of the eight test images (Ping-Pong, Flower Garden, Suzie, Cheer Girls, Mobile & Calendar, Tempest, Football, and Bicycles) and then average the bit rates of the blocks having identical MAD values. The JPEG algorithm with $DF = 2.0$ is used here, and the luminance and chrominance components are treated independently. The results are shown in Fig. 4. It is seen that the output rate of a block is roughly proportional to the MAD.

To compress a given field, the feedforward DF estimator evaluates the MAD's of all blocks and predicts the bit rate of each block using Fig. 4. To predict the bit rate of the field for $DF = 2.0$, average the predicted bit rates of all blocks. Here, luminance and chrominance rates are predicted independently, and then added to get the field bit rate. The rates for $DF \neq 2.0$ are predicted using (1) and Fig. 3. Specifically, in (1) we assume $n_k(m)$, $m = 2$ is the predicted field bit rate for $DF = 2.0$ and calculate $\hat{n}_k(p)$. Finally, the DF is selected as in the case of the feedback DF estimation. The feedforward DF estimation is summarized as follows.

1) For a given field, evaluate the MAD's of all 8×8 blocks and predict the bit rate of each block using Fig. 4.
2) Predict the bit rate for $DF = 2.0$ by averaging the predicted rates of all blocks in Step 1.
3) Using (1) and the curve in Fig. 3, predict bit rates for all DF values between 0.1 and 12, $DF \neq 2.0$.
4) Select the DF value corresponding to the predicted bit rate closest to the transmission rate.

Due to Steps 1 and 2, the feedforward DF estimation requires more computation compared to feedback estimation, but the additional computational load is not heavy; the implementation of the feedforward estimation is considerably simpler than those in [7] and [8].

C. Scene Change Detector

A scene change is claimed at the (k+1)th field if the absolute difference between the feedforward prediction values at the (k+1)th and kth fields for $DF = 2.0$ is greater than the threshold value. In our simulation, the threshold is set at 10% of the transmission bit rate.

IV. Simulation Results

The image sequence used for examining the performance of the proposed buffer-control policy consists of four different types of CCIR 601 images. Specifically, the sequence has ten Flower Garden fields followed by ten Suzie, ten Ping-Pong, and ten Football fields; it contains a total of 40 fields with three scene changes.

The JPEG algorithm is applied to the image sequence under the assumption that the transmission bit rate is 0.8 b/pixel. The DF value of the JPEG algorithm is adjusted once for each field by using the proposed buffer-control policy and the statistics in Figs. 3 and 4. For comparison, the statistical feedback buffer-control method employing only the feedback DF estimator—after removing the feedforward DF estimator from Fig. 2—is also applied.

Fig. 5 shows the evolution of output bit rates for JPEG encoders employing the feedforward/feedback and feedback buffer-control

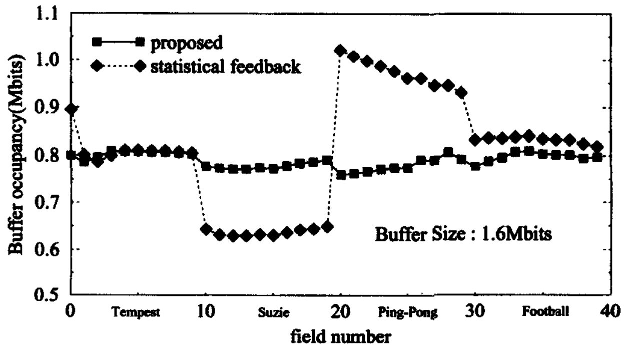

Fig. 5. Evolution of output bit rates.

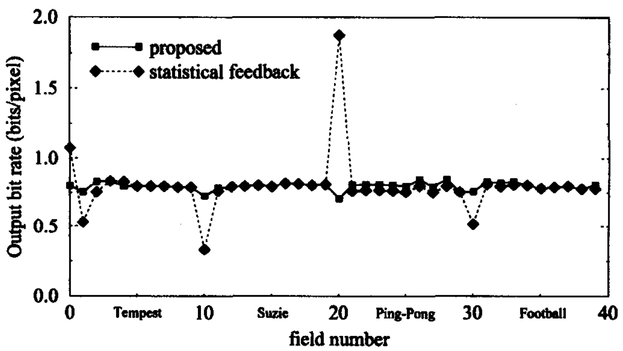

Fig. 6. Evolution of buffer occupancy.

methods. It may be seen that both methods perform well in stationary portions of the image sequence. At scene changes, however, only the proposed feedforward/feedback method controls the bit rate, which the feedback method cannot. The buffer occupancy observation shown in Fig. 6 indicates that the probability of buffer overflow and underflow associated with this proposed method is considerably lower than that associated with the feedback method.

V. Conclusions

A statistical feedforward/feedback buffer-control method employing two statistical bit rate predictors, feedforward and feedback, has been introduced. It has been shown that the proposed method is reasonably simple to implement and can control the bit rate even at scene changes.

This buffer-control policy can be directly applied to some other DCT-based intrafield or intraframe coding techniques such as the ones in [6] and [7]. Futhermore, it is expected that this statistical feedforward/feedback control concept can be applied to various image-compression systems not based on DCT. Research in this direction is being pursued.

References

[1] Z. L. Budrikis, J. L. Hullett, and D. Q. Phiet, "Transient-mode buffer stores for nonuniform code TV," *IEEE Trans. Commun.*, vol. COM-19, pp. 913–922, Dec. 1971.

[2] S. K. Goyal and J. B. O'Neal, "Entropy coded differential pulse-code modulation systems for television," *IEEE Trans. Commun.*, vol. COM-23, pp. 660–666, June 1975.

[3] B. G. Haskell and R. L. Schmidt, "A low-bit-rate interframe coder for videotelephone," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1475–1495, Oct. 1975.

[4] H. Murakami, S. Matsumoto, Y. Hatori, and H. Yamamoto, "15/30 Mbits/s universal digital TV codec using a median adaptive predictive coding method," *IEEE Trans. Commun.*, vol. COM-35, pp. 637–645, June 1987.

[5] J. Zdepski, D. Raychaudhuri, and K. Joseph, "Statistically based buffer control policies for constant rate transmission of compressed digital video," *IEEE Trans. Commun.*, vol. 39, no. 6, pp. 947–957, June 1991.

[6] K. Tzou, "An intrafield DCT-based HDTV coding for ATM networks," *IEEE Trans. Circuits Syst. Video Technol,* vol. 1, no. 2, pp. 184–196, June 1991.

[7] M. Yoneda *et al.*, "An experimental digital VCR with new DCT-based bit-rate reduction system," *IEEE Trans. Consumer Electron.*, vol. 37, no. 3, pp. 276–282, Aug. 1991.

[8] A. Ortega, K. Ramchandran, and M. Vetterli, "Optimal buffer-constrained source quantization and fast approximations," in *Proc. IEEE ISCAS*, 1992, pp. 192–195.

[9] ISO/JPEG, "Digital compression and coding of continuous-tone still images," Draft Int. Std. DIS 10918-1, Aug. 1991.

Comments on "On the Invertibility of Morphological Representation of Binary Images"

Ben K. Jang and Roland T. Chin

***Abstract*—In a recent paper, Charif-Chefchaouni and Schonfeld investigated the invertibility of a morphological representation of binary images and determined the necessary and sufficient conditions for its inverse. In this correspondence, we show that one of the derived necessary conditions is not valid. A counterexample is given to illustrate our observation.**

I. Introduction

In the above paper,[1] Charif-Chefchaouni and Schonfeld derived a number of necessary conditions for the exact reconstruction of a binary image from its morphological representation. One of these conditions is believed to be invalid, and a counterexample is given to demonstrate our observation.

For convenience, we shall use the same notations and definitions as described in the above paper,[1] to represent the basic morphological operations, which are dilation ($X \oplus B$), erosion ($X \ominus B$), opening [$\gamma_B(X)$], and closing [$\phi_B(X)$].

Consider a sequence of structuring elements $\{B(n): n \geq 0\}$ such that $(0, 0) \in B(n)$ and $B(n) \neq \{(0, 0)\}$ for all n. The sequence $A(n)$ is given by

$$A(n+1) = A(n) \oplus B(n) \quad \text{for} \quad n \geq 0$$

and

$$A(0) = \{(0, 0)\}. \tag{1.1}$$

Manuscript received June 6, 1995; revised June 9, 1995.

B. K. Jang is with the Advanced Imaging Research Laboratory, Kodak Health Imaging System Inc., Rochester, NY 14653-5740 USA.

R. T. Chin is with the Department of Computer Science, University of Wisconsin—Madison, Madison, WI 53706. He is now with Hong Kong University of Scioence and Technology, Kowloon, Hong Kong.

Publisher Item Identifier S 1057-7149(96)01792-7.

[1] M. Charif-Chefchaouni and D. Schonfeld, *IEEE Trans. Image Processing,* vol. 3, no. 6, pp. 847–849, Nov. 1994.