Adaptation of the MPEG Video-Coding Algorithm to Network Applications

Masahisa Kawashima, *Student Member, IEEE*, Cheng-Tie Chen, *Member, IEEE*, Fure-Ching Jeng, *Member, IEEE*, and Sharad Singhal, *Member, IEEE*

Abstract—**The Motion Picture Experts Group (MPEG) video-coding algorithm is regarded as a promising coding algorithm for coding full-motion video. However, since MPEG was originally designed for storage applications, some problems must be solved before the algorithm can be applied to interactive services. Due to the use of periodic intraframe coding and bidirectional interframe prediction, the end-to-end delay of the MPEG algorithm is much larger than that of the H.261 algorithm. In packet video transmission, the large peak in bit rate caused by periodic intraframe coding may lower performance of statistical multiplexing. In this paper, real-time video transmission using the Hybrid Extended MPEG (Bellcore's proposal to ISO/MPEG) is considered. First, the end-to-end delay of the Hybrid Extended MPEG algorithm is analyzed. Then several schemes to reduce the delay are considered and compared with regular coding schemes in terms of image quality, end-to-end delay and performance of statistical multiplexing. Error resilience of the presented schemes is also tested by simulations assuming cell loss. It is shown that the presented schemes improved the end-to-end delay and performance of statistical multiplexing significantly.**

I. Introduction

THE Motion Picture Experts Group (MPEG) video-coding algorithm [1] is regarded as a promising compression algorithm for full-motion video. However, since MPEG was originally designed for storage applications, some problems must be solved before applying the algorithm to interactive services.

In MPEG, frames are coded in one of three coding modes—intraframe coding frame (I frame), predictive coding frame (P frame), or bidirectionally predictive coding frame (B frame). Usually I frames are coded using many more bits than P or B frames. When a network supports a constant bit rate (CBR) channel, a large buffer is required between the coder and the network to absorb the difference in data amount among frames with different coding modes. Therefore, the buffering delay is much larger than that of the H.261 [2] algorithm. When a network supports a variable bit rate (VBR) channel, this buffer is not required. However, in this case I frames create an unwanted peak in bit rate, which is usually periodic. The ratio of I frame to B frame is typically around 12 in simulation model 3 (SM3) [3], and thus this peak can create a periodic congestion of the network. Furthermore, when B frames are used, a delay of several frames called the frame reordering delay is introduced as a result of the difference between the encoding order and the display order of frames.

In this paper, real-time video transmission using the Hybrid Extended MPEG (Bellcore's proposal to ISO/MPEG) [4] is considered. The major problem in applying this algorithm to real-time interactive applications is delay. First, the end-to-end delay of the Hybrid Extended MPEG algorithm is analyzed. Then, several schemes to lower the delay are presented. In order to reduce the frame reordering delay, three different types of interframe prediction are considered. These three interframe predictions are compared with typical MPEG coding in terms of the decoded picture quality, the end-to-end delay, and performance of the statistical multiplexing in packet video transmission, by simulations. Elimination of periodic I frames is also considered to reduce the necessary buffer. Instead of using I frames, some macro blocks in P frames are coded by intraframe coding. Although this implementation sacrifices the random accessibility, such a function is usually not needed in real-time interactive applications. However, removing I frames raises a number of questions. One question is about error resilience because it is possible for errors to expand to neighboring areas and propagate forever due to motion compensation. Another question may be, "What happens if a receiver starts decoding video in the middle of a sequence?" This situation is referred to as delayed access. Simulation results demonstrating error resilience and delayed access are also presented.

II. Overview of the Hybrid Extended MPEG Algorithm

The Hybrid Extended MPEG algorithm is designed to code interlaced video signals at bit rates from 3 Mb/s up to 10 Mb/s. The input video signal is assumed to be in the CCIR601 format, where one field has 240 lines and one line has 720 luminance pixels. The chrominance signals are downsampled in preprocessing so that one chrominance pixel corresponds to four luminance pixels. The algorithm is similar to SM3 [3]. However, three new features are included to improve performance. First is the

Manuscript received April 1, 1992; revised November 5, 1992. Paper was recommended by Associate Editor Dimitris Anastassiou.

M. Kawashima was with Bellcore, Morristown, NJ. He is now with Waseda University, Tokyo, Japan.

C.-T. Chen, F.-C. Jeng, and S. Singhal are with Bellcore, Red Bank, NJ 07701.

IEEE Log Number 9209197.

hybrid field motion compensation (MC)/frame discrete cosine transform (DCT) coding where the coder performs MC on fields but performs DCT and quantizations on frames by merging even and odd fields together. The second feature is a new strategy for bit allocation and rate control [5]. The coder assigns a quantizer scale (quantization step size) to each macro block proportionally to the local spatial activity of the coded area after it observes the scene content of an entire frame. The third feature consists of frame adaptive variable-length code (VLC) tables based on the JPEG standard. The VLC table is transmitted for every frame because empirically, the bit count saved by using the adaptive VLC table is larger than the bit count consumed to transmit the VLC table at bit rates higher than 3 Mb/s. After quantizing all the DCT coefficients in one frame, the VLC table is generated according to the statistics of quantized coefficients.

Frames are coded in one of three modes, as shown in Fig. 1. I frames are coded using intraframe coding. P frames are coded using interframe prediction from the previous P (or I) frame. B frames are coded using bidirectional interframe prediction from the previous and the next P (or I) frames. Note that the encoding order of frames is different from the display order, because, in order to use the bidirectional interframe prediction, B frames have to be coded after the P frame that follows them in the display order. In Fig. 1, frames are arranged in the display order, while each frame is numbered with the encoding order. We define a group of pictures (GOP) as the interval from the I frame to the B frame before the next I frame in the encoding order.

As shown in Fig. 2, one frame is composed of two fields. Motion estimation and compensation are performed on fields using a block of 16 × 16 pixels as the unit block for motion compensation. DCT and following operations, including quantization and variable-length coding, are performed on frames. One frame is split into blocks of 16 × 16 pixels called macro blocks. A super-macro block is defined as a pair of vertically adjacent macro blocks. One super-macro block corresponds to a block of 16 × 16 pixels on fields, which is the unit block for motion compensation. We also define a slice as a group of super-macro blocks in one row.

The block diagram of the encoder is shown in Fig. 3. First, the input signal from a camera is preprocessed to downsample the chrominance signals and change the pixel order from the raster order to the 16 × 16 block-based order. Motion estimation and motion compensation are performed for each field independently. After merging the two motion-compensated fields, the macro block type (coding type) is selected for every macro block and a residual signal is computed. The residual signal undergoes DCT and quantization. Then quantized DCT coefficients are coded by VLC's. Fig. 4 shows the block diagram of the decoder. In the decoder, the processes are operated in the reverse order to generate the decoded picture. Finally, the decoded picture is postprocessed to upsample the chrominance signals and convert the field-mixed blocks to an interlaced raster.

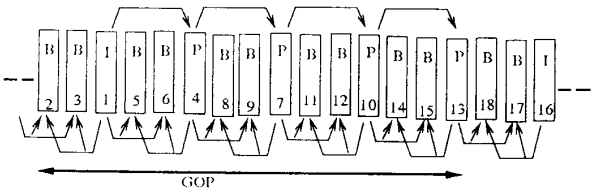


Fig. 1. Coding modes and GOP.

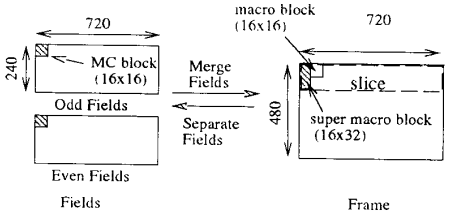


Fig. 2. Fields and frame.

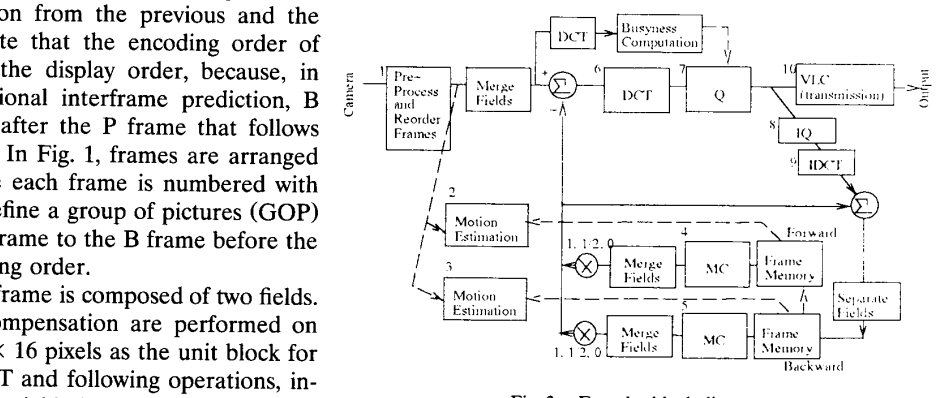


Fig. 3. Encoder block diagram.

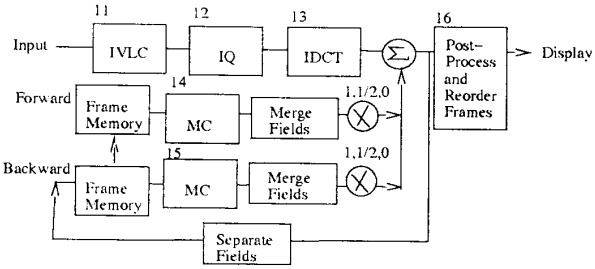


Fig. 4. Decoder block diagram.

III. Delay Analysis

In this section, the end-to-end delay of the Hybrid Extended MPEG is analyzed [4].

A. Encoding Delay

Fig. 5 shows the timing chart of the processes performed in the Hybrid Extended MPEG encoder. The arrows indicate the duration of each process. The num-

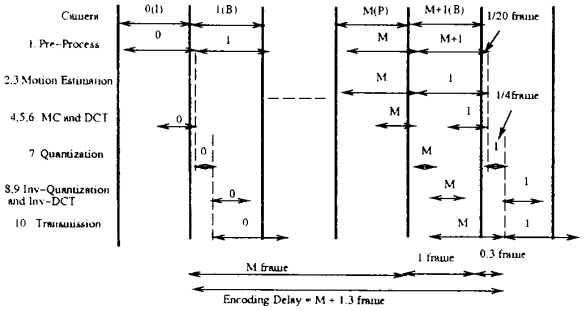


Fig. 5. Encoder timing chart.

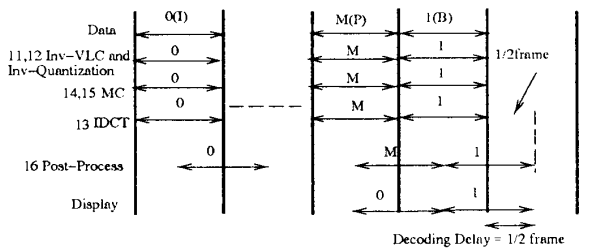


Fig. 6. Decoder timing chart.

bers attached to the arrows represent the frame number of the processed frame. Each process is labeled with the block number from the encoder block diagram in Fig. 3. We assume that M—1 B frames are inserted between two adjacent P frames [3].

As described in the previous section, an input signal is preprocessed to downsample the chrominance signals and convert the pixel order into the 16×16 block-based order. To downsample the chrominance signals, an 11-tap filter is used. Since each macro block is 16 lines high and 5 more lines are needed to apply the 11-tap filter, a delay of 21 lines, which is approximately 0.05 (21/480) frame, forms the preprocessing delay.

The encoder incurs a half-frame delay before DCT computation can start because DCT's are computed on a frame basis. We assume that DCT's for the entire picture are finished within a half-frame interval. The quantization cannot be started until the entire picture has been transformed, because quantizer scales are assigned after observing the content of the entire picture. Also, transmission of the quantized coefficients cannot be started until quantization for the entire picture is finished because a frame-adaptive VLC table must be generated that depends on the statistics of the quantized coefficients. Therefore, one-frame interval and time involving quantization should be counted as delay. Assuming that quantization and VLC table generation take a quarter-frame time, the encoder delay becomes 1.25 frames.

When bidirectional interframe prediction is used, an additional M-frame delay is imposed due to reordering of frames. So the total encoding delay, including preprocessing, is $M + 1.3$ frames.

B. Decoding Delay

Fig. 6 shows a timing chart of the processes performed in the decoder. Each process is labeled with the corresponding block number in the decoder block diagram in Fig. 4. The postprocessor takes the decoded field-mixed signal in the block-oriented order and converts it into an interlaced signal in the raster order. A half-frame delay is necessary in the postprocessor to convert the signal into the interlaced one; this is the only delay in the decoding process.

C. End-to-End Delay

In CBR transmission, the end-to-end delay is the sum of the encoding delay, the buffering delay and the decoding delay. Therefore, the end-to-end delay is $M + 1.8$ frames plus the buffering delay, assuming that propagation delay in a network is negligible in CBR transmission.

In packet-video transmission (VBR transmission), the packetization delay should be considered instead of the buffering delay. We assume that the payload in each cell is 48 bytes and that at least one cell is transmitted for each slice. So the worst-case packetization delay is 1/15 frames, i.e., one slice time interval. In VBR transmission, the propagation delay is not negligible because cells may undergo queueing several times during transmission. The end-to-end propagation delay has to be considered in relation to the cell-loss ratio and the utilization of the network. This trade-off is analyzed in Section V by some simulations.

IV. Schemes to Lower the Delay

When we consider regular coding by the Hybrid Extended MPEG, which assumes $M = 3$ and a buffer size of six frames, the end-to-end delay is 10.8 frames ($3 + 1.8 + 6 = 10.8$), which corresponds to 360 ms. This delay is larger than the tolerable delay for interactive applications, which is assumed to be about 150–300 ms. In this section, several schemes to lower the delay are considered.

A. Reduction of Bidirectional Frames

Fig. 7 shows three different types of interframe prediction, which reduces the use of bidirectional prediction frames and thus reduces the frame-reordering delay. In the first scheme (I), only one B frame is inserted between two P frames. In the second and third schemes, no B frames are used. In the second scheme (II), a new picture type called P′ frame is defined. P′ frames are coded without backward interframe prediction as if they were P frames but are not used as references for coding succeeding frames. P′ frames can be regarded as B frames that do not have backward interframe prediction. The coder allocates fewer bits to P′ frames and more bits to P frames. Since the P frames are used as reference for coding the succeeding frames, the efficiency of interframe prediction is improved by assigning more bits to the P frames. Although intuitively this may result in periodic quality degradation, such periodic quality degradation is not perceived empirically when M is small. Moreover, as will be shown in Section V, the signal-to-noise ratios (SNR's) for the P′ frames are better than those obtained from the

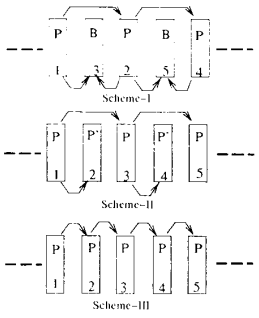


Fig. 7. Schemes to reduce the frame-reordering delay.

conventional interframe coding scheme, due to improved interframe prediction efficiency, especially for low-scene-activity frames. In the third scheme (III), only P frames are used. We refer to the regular Hybrid Extended MPEG algorithm, where two B frames are inserted between P frames and I frames are inserted for every fifteen frames, as scheme O, which is shown in Fig. 1

In CBR transmission, the end-to-end delay for scheme I is 3.8 frames plus the buffering delay because $M = 2$. For schemes II and III, the end-to-end delay is 1.8 frames plus the buffering delay because frames do not have to be reordered. Scheme III requires a shorter buffer because each P frame results in almost equal bits.

B. Replacement I Frames with I Columns

In MPEG, I frames are inserted periodically to enable random access to any segment in a scene. But the use of I frames requires a large buffer. Removing periodic I frames reduces the necessary buffer size in CBR transmission or reduces the peak-to-mean ratio in bit rate in VBR transmission. However, if the data is not refreshed by intraframe coding, errors—such as channel errors or computation mismatches in the inverse DCT—may propagate forever. To flush these errors periodically, some macro blocks are forced to be intraframe mode. As show in Fig. 8, each frame is split into fifteen columns, each of which is three macro blocks wide. Every frame, some columns are then selected as I columns in order and coded by intraframe coding.

This idea is applied to all three of the schemes described in Section IV. In schemes I and II, two out of the fifteen columns in P frames are selected as I Columns. In scheme III, one column is selected as an I column because the number of P frames during the same interval is double that in scheme I or scheme II. I columns are selected in turn so that each column is refreshed with equal frequency, as shown in Fig. 8. It is also possible to arrange intraforced macro blocks based on slices. In this case, however, the large peak in bit rate due to intraframe coding is still observed when a picture is transmitted slice by slice.

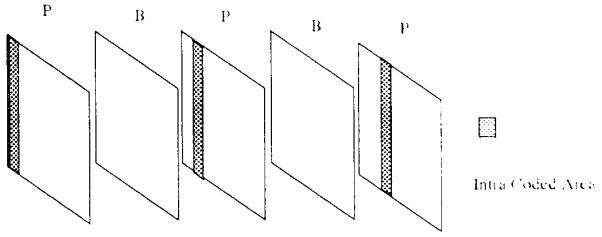


Fig. 8. Periodic refreshment by intracoding.

C. Measures Against Scene Changes

Scene changes cause large peaks comparable to I frames in the output rate. To enable transmission with a small buffer, some measures must be taken against scene changes. When more than half the macro blocks are selected as intracoded macro blocks in a P or P′ frame, we assume a scene change has taken place. In B frames, a scene change is assumed if a scene change is detected in the previously coded P frame. Different measures are taken for different coding modes at a scene change. In P frames, quantizer stepsizes are multiplied by a certain scale factor. In B frames, the residual signal is neglected, i.e., only the macro block type information and the motion vectors are transmitted. In P′ frames, all the macro blocks are forced to be fixed blocks; in other words, the coded frame is dropped and the frame is replaced with the previous frame at the decoder. Note that in P or B frames, a scene-change frame is coded with more distortion than regular frames, and in scheme II a scene change is delayed by one frame when it happens in the P′ frame. However, because the human visual system is not very sensitive to distortion at scene changes, the perceptual picture quality does not degrade.

V. Simulation Results

The MPEG test sequences Mobile and Calendar and Table Tennis were coded at 3.9 Mb/s using the Hybrid Extended MPEG algorithm with the four different schemes, including scheme O. Within the constraint of the coding bit rate, the rate control adjusted the quantizer scales to keep the ratio of quantizer step sizes between P (or I) frames and B (or P′) frames at 1:2 [5]. The necessary buffer size was determined heuristically from experiments. In scheme O, a six-frame-long buffer was used. In schemes I and II, a two-frame-long buffer was used. In scheme III, a one-frame-long buffer was used. Total end-to-end delay in CBR transmission is summarized in Table I.

A. SNR Comparison

Fig. 9 and Fig. 10 show signal-to-noise ratio (SNR) of the decoded pictures as a function of frame number for the four different schemes. Generally speaking, scheme O shows the best SNR performance, while scheme III shows the worst. In Figure 9, the scene activity is relatively low. In this sequence, the four schemes show perceptible differences in the quality of the decoded pictures. In Fig. 10, the scene activity is relatively high from frames 15–45

TABLE I
DELAY

Scheme	Delay in Frames: Frame Reordering	Delay in Frames: Buffering	Total
O	3	6	10.8(360 ms)
I	2	2	5.8(192 ms)
II	0	2	3.8(127 ms)
III	0	1	2.8(93 ms)

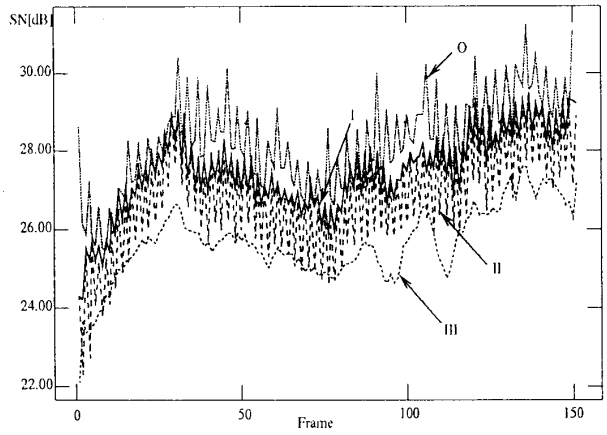


Fig. 9. SNR in Mobile and Calendar at 3.9 Mb/s.

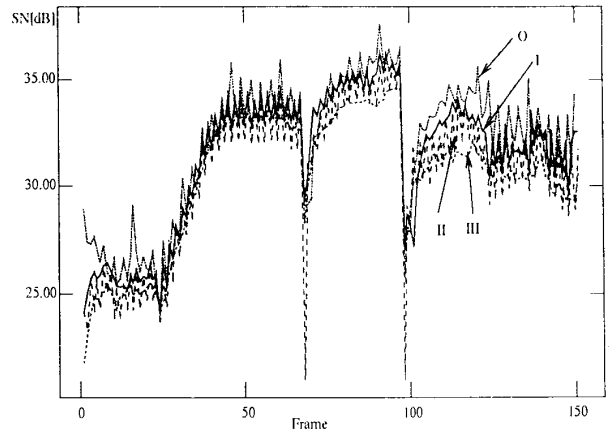


Fig. 10. SNR in Table Tennis at 3.9 Mb/s.

because of zooming. In this part, the four schemes do not show significant difference in quality. In scheme II, SNR drops dramatically around frames 68 and 98. This is because the rate-control strategy in scheme II forces all macro blocks to be fixed blocks and replaces the frame with the previous one at the scene change.

B. Rate Behavior

Fig. 11–14 show bitcount observed slice by slice in frames 30–60 in Mobile and Calendar. The dashed lines in the figures show the average bitcount, which is derived from the coding bit rate. In scheme O, the rate transits among three levels (i.e., I, P, and B) where the bitcounts in I frames are about six times larger than the average. In schemes I and II, the rate transits between two levels (i.e., P and B), where the bitcounts in P frames are about two times larger than the average. In scheme III, only P

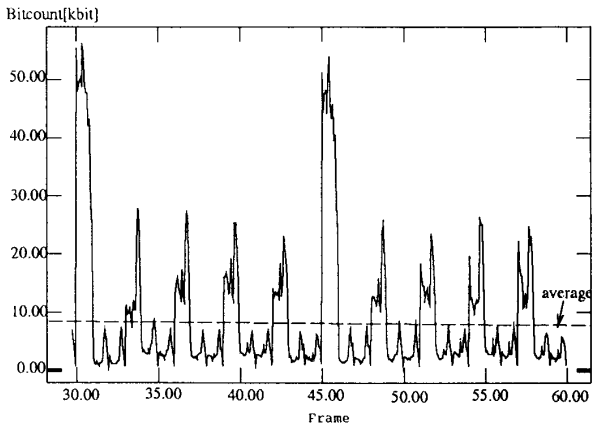


Fig. 11. Bitcount per slice (scheme O).

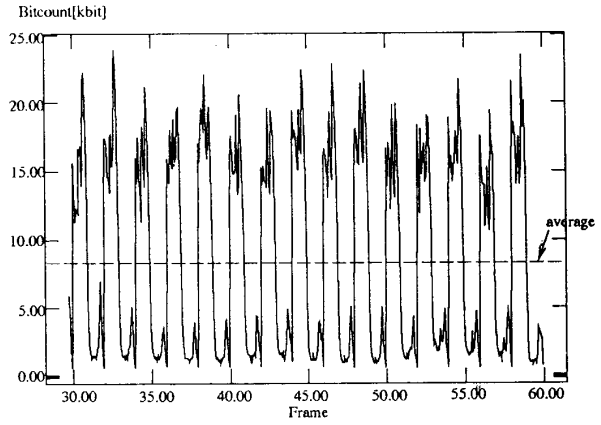


Fig. 12. Bitcount per slice (scheme I).

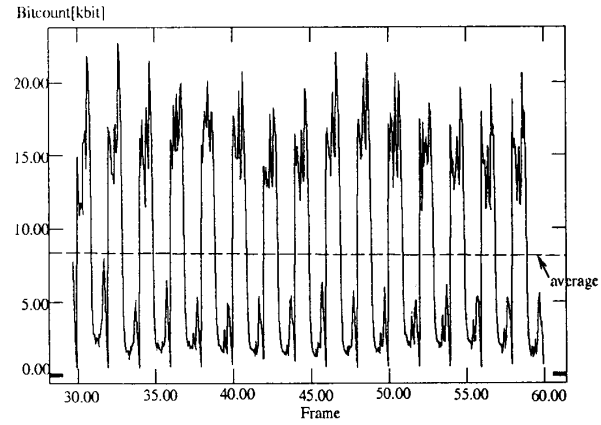


Fig. 13. Bitcount per slice (scheme II).

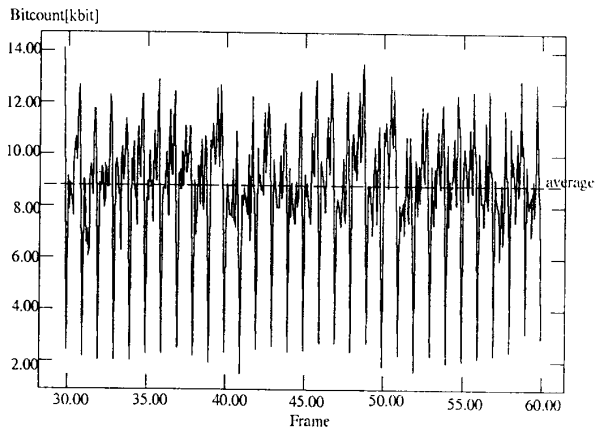


Fig. 14. Bitcount per slice (scheme III).

frames are used, and the ratio of the peak bitcount to the average is less than two.

C. Delay and Cell-Loss Ratio in Packet-Video Transmission

1) Simulations of Multiplexing: In packet-video transmission, the cell-loss ratio depends on the maximum allowed end-to-end delay and the utilization of the network. The larger the acceptable end-to-end delay, the smaller the cell-loss ratio for a given network loading. Here this trade-off is analyzed for the four schemes by simulations of multiplexing.

Simulations of multiplexing are carried out according to the model shown in Fig. 15. In Fig. 15, K represents the number of multiplexed sources and C represents the channel bit rate. Each source is assumed to be controlled with the leaky bucket control [6], where the target average bit rate of source i is represented by m_i. It is assumed that the size of the queueing buffer is infinite and cell loss is caused only by excess end-to-end queueing delay.

The channel utilization, which is represented by U, is defined as follows:

$$U = \frac{\sum_{i=1}^{K} m_i}{C}. \tag{1}$$

Since the rate and buffer control in CBR transmission can be regarded as a leaky bucket control, the rate observed slice by slice in Section V.B. is used to emulate each source. First a 10-second-long reference source model is made by concatenating the rate in Mobile and Calendar and Table Tennis together. Different sources are emulated by starting the reference model from a randomly selected slice. When a source reaches the end of the reference model, it jumps back to the beginning. Since the target average rate is 3.9 Mb/s in the reference model, Eq. (1) becomes

$$U = \frac{3.9K}{C}. \tag{2}$$

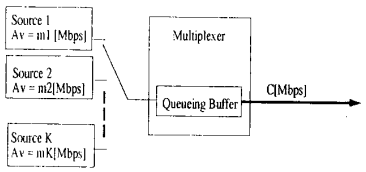


Fig. 15. Multiplexing model.

Multiplexings parametrized by K and U were simulated. In each simulation, data were observed slice by slice for 5 min. while changing the starting slice of each source every 5 s. Simulations were carried out independently for the four different schemes using the actual rate behavior in each scheme.

2) Queueing Delay: Fig. 16 shows the maximum queueing delay (the maximum queue size) as a function of the channel utilization when $K = 15$. As expected, scheme O shows the largest delay due to I frames, while scheme III has the lowest delay. Schemes I and II have almost equal delay performance.

Next, the distribution of the queueing delay was observed. The queueing delay was calculated for each cell, and then the delay was quantized with a granularity of 0.01 ms. In other words, delays within a range from $(k - 0.5)/100$ ms up to $(k + 0.5)/100$ ms are represented by k. Here, $P_1(k)$ is defined as the probability of the case where the quantized queueing delay is k. $P_1(k)$ was measured from the distribution of the queueing delay. Fig. 17 shows $P_1(k)$ when $K = 15$ and $U = 0.8$. Once again, scheme O shows the worst performance while scheme III is the best. Schemes I and II are again found to be roughly equivalent and somewhere in the middle.

3) Tradeoff Between Delay and Cell-Loss Ratio: Cell-loss ratio can be derived from the distribution of the queueing delay $P_1(k)$.

When cells undergo queueing N times during transmission, the distribution of the end-to-end queueing delay can be obtained by operating convolutions of $P_1(k)$ recursively as follows:

$$P_N(k) = \sum_{l=1}^{k} P_{N-1}(l) P_1(k-l). \tag{3}$$

The maximum limit on the end-to-end delay is related to the cell-loss ratio. Here, it is assumed that any delay but the queueing delay is negligible. When $k_{\max}$ is given as the limit of the end-to-end delay, the corresponding cell-loss ratio, represented by $P_L(k_{\max})$ is calculated from

$$P_L(k_{\max}) = 1 - \sum_{k=0}^{k_{\max}} P_N(k). \tag{4}$$

Fig. 18 shows $P_L(k_{\max})$ when $N = 4$, $K = 15$, and $U = 0.8$. Comparing the four schemes in Figs. 16–18, scheme O shows significantly worse performance than the others. The reason is that in scheme O each source transits among three levels, i.e., I frame, P frame, and B frame. We conclude that replacement of I frames with partial intraframe coding in P frames improves performance of

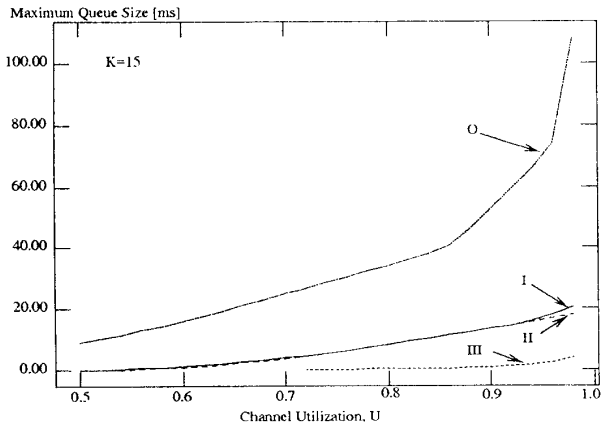


Fig. 16. The maximum queueing delay.

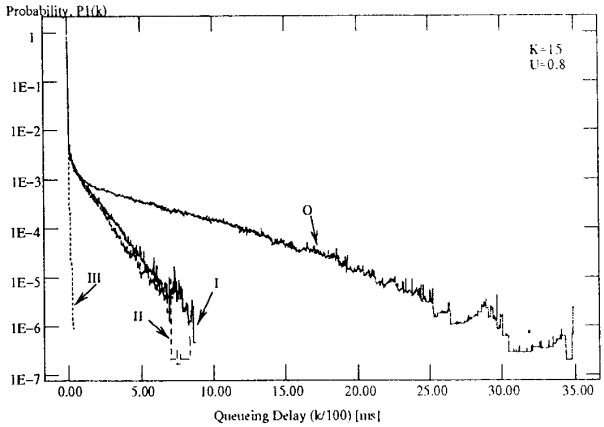


Fig. 17. Distribution of the queueing delay.

statistical multiplexing significantly in packet-video transmission. Schemes I and II show almost equal performance, though scheme II performs slightly better. This is because although each source transits between two levels (i.e., P frame and B frame or P frame and P′ frame), the ratio of the rate in P frames to that in P′ frames in scheme II is slightly lower than the ratio of the rate in P frames to that in B frames in scheme I.

D. Error Resilience

Since removal of I frames reduces robustness against errors, cell-loss resilience of the decoder was tested by simulations. In the following simulations, only scheme I was considered. Simulations were carried out by dropping 48-byte-long segments of the bit stream randomly at the ratio of 10^{-4}. In the simulation it was assumed that the decoder was not notified of any cell loss by the network, and cell loss was detected by the decoder itself when any of the following conditions were encountered:

- The synchronization words were not found when they were expected.
- The decoder reached an undefined codeword.

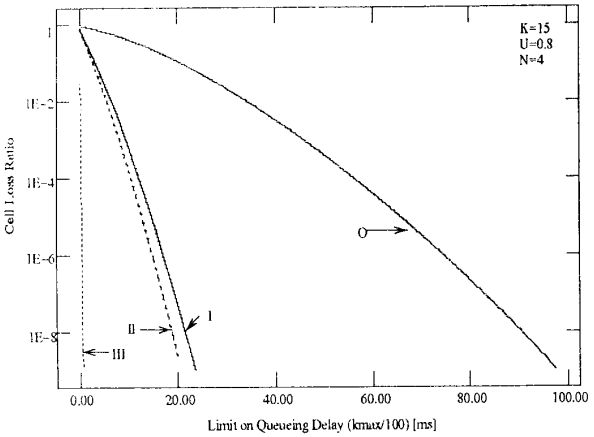


Fig. 18. Tradeoff between cell-loss ratio and the maximum delay.

- The motion vector pointed to an area out of the picture frame.
- The sum of the run-length of zero and the current frequency order of DCT coefficient exceeded 63.

When a cell loss was detected, the error was concealed by replacing the affected slice with the corresponding one in the previous P frame.

Figs. 19 and 20 show the behavior of SNR in Mobile and Calendar and Table Tennis. It can be seen that usually the error propagation is terminated in about 15 frames. Note that protective measures such as segregation of the data with priorities were not taken because the objective of this simulation was to observe propagation of errors, though these measures could reduce degradation in the quality.

E. Delayed Access

Simulations on delayed access were carried out by letting the decoder start decoding the bit stream from frame 61 with the decoder frame memory filled with a default value. In the simulations, 128 was used as the default value. Figs. 21 and 22 show the behavior of SNR in case of delayed access. It can be seen that the SNR of the sequence reaches to the steady state in about 30 frames.

VI. Conclusion

In this paper, real-time transmission of video using the Hybrid Extended MPEG video coding was considered. We analyzed the delay of the Hybrid Extended MPEG. In order to reduce the frame reordering delay, three different types of interframe prediction were considered. One scheme had only one B frame between two P frames. In the second scheme, P′ frames which are similar to B frames but contain only forward prediction, were used. The third scheme had only P frames. In these three schemes, I frames were replaced with partial intraframe coding in P frames in order to reduce the necessary buffer size in CBR transmission.

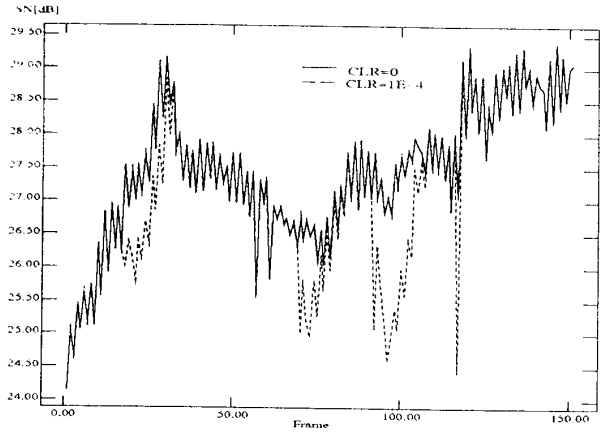


Fig. 19. SNR at CLR = 10^{-4} in Mobile and Calendar.

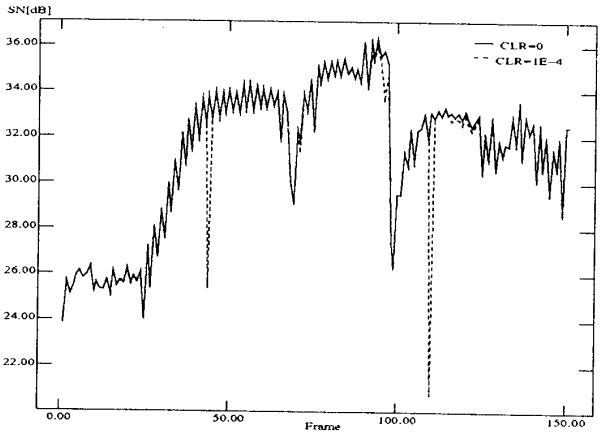


Fig. 20. SNR at CLR = 10^{-4} in Table Tennis.

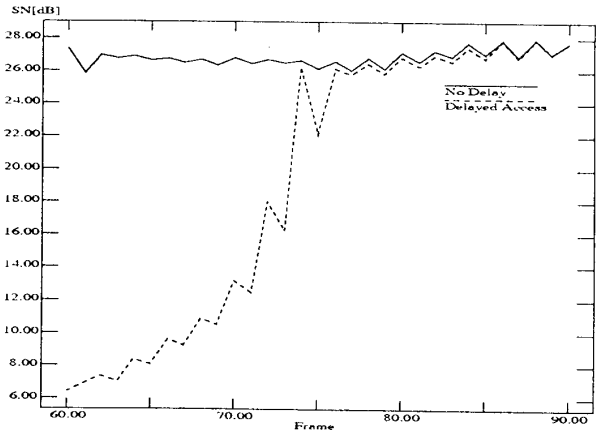


Fig. 21. SNR in Mobile and Calendar in case of delayed access.

The three schemes were compared with regular coding by the Hybrid Extended MPEG algorithm in terms of the quality of the decoded pictures, the end-to-end delay, and the performance of statistical multiplexing. From simulations it was shown that the difference in image quality

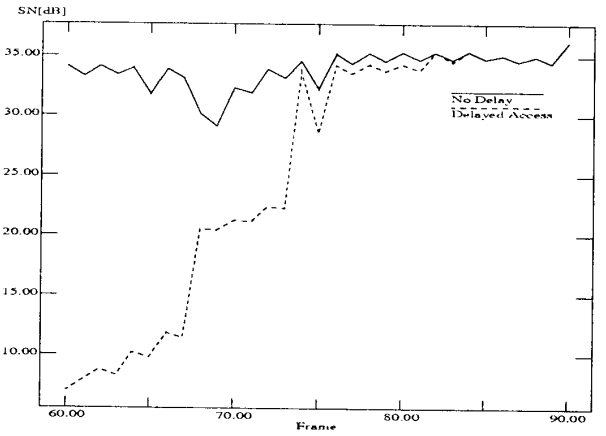


Fig. 22. SNR in Table Tennis in case of delayed access.

among the four schemes, including the regular coding, is not significant when the scene activity is relatively high. However the four schemes show perceptible difference when the scene activity is relatively low. Also, it was shown that elimination of I frames improves the performance of statistical multiplexing significantly.

Propagation of errors, which may result from the replacement of I frames, was simulated assuming cell loss at the ratio of 10^{-4}. It was concluded that the propagation is terminated in about 15 frames. It was also shown by simulations that the image quality usually reaches the steady state in about 30 frames when the decoder starts decoding in the middle of a sequence.

Since the presented scheme does not take any protective measures against cell losses, measures to achieve more robustness is an issue for further study. One of these measures is rearrangement of the bit stream, with error correction cells added according to the significance of the data. Also, a further analysis of end-to-end delay and cell-loss ratio using a more detailed network model remains to be studied.

ACKNOWLEDGMENT

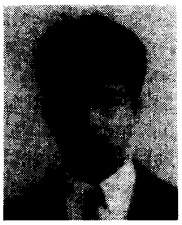
The authors would like to thank A. Wong of Bell Communications Research for her meaningful comments and efforts regarding MPEG-II work. Thanks are also given to M. Garret of Bell Communications Research for his many suggestions about this paper. Masahisa Kawashima would like to thank Prof. H. Tominaga of Waseda University and Dr. C. Judice of Bell Communications Research for their efforts to give him a chance to work on the presented research at Bellcore.

REFERENCES

[1] MPEG Video CD Editorial Committee, ISO-IEC/JTC1/SC2/WG11/MPEG90/176, "MPEG video committee draft," Dec. 1990.

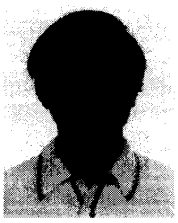
[2] *Draft revision of recommendation H.261*, Doc. 572 CCITT SG XV, Mar. 1990.

[3] MPEG Video SM Editorial Group, ISO-IEC/JTC1/SC2/WG11/MPEG90/041, "MPEG video simulation model three," July 1990.
[4] Bellcore and NTT, ISO-IEC/JTC1/SC2/WG11/MPEG91/204, "Hybrid extended MPEG coding algorithm," Nov. 1991.
[5] M. Kawashima, C. T. Chen, F. C. Jeng, and S. Singhal, "A new rate control strategy in the MPEG video coding algorithm," *Visual Commun. Image Rep.*, submitted for publication.
[6] COMPENDIUM ATM; Coding procedures, CCITT SGXV, AVC-13, Nov. 1990.



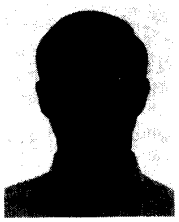
Masahisa Kawashima (S'92) received the B.S. and M.S. degrees from Waseda University, Tokyo, Japan in 1989 and 1991, respectively. He is currently working toward the Ph.D. degree at Waseda University.

From April 1991 to March 1992 he worked at Bellcore, Morristown, NJ as a resident visitor from Waseda University. His research interests are in video transmission over broadband networks and image processing/coding.



Cheng-Tie Chen (S'79–M'89) received the B.S.E.E. degree from the National Taiwan University, Taipei, R.O.C. in 1977, and the M.S.E. and Ph.D. degrees from the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, in 1981 and 1983, respectively.

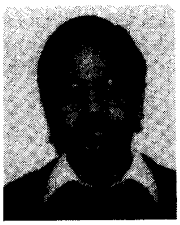
From 1983 to 1984, he was a Research Associate with the Department of Electrical Engineering, Princeton University, Princeton, NJ, where he was engaged in the research of non-Gaussian and sensor-array signal processings. From 1984 to 1987, he was with the Research Laboratories, Eastman Kodak Company, as a Senior Research Scientist. His research activities there were in the areas of image restoration and image coding/transmission. Since 1987, he has been with Bell Communications Research, Red Bank, NJ, as a Member of Technical Staff. His current research interests are in the areas of visual communications and multimedia applications. He was also an Adjunct Associate Professor with the Department of Electrical Engineering, University of Pennsylvania.



Fure-Ching Jeng (S'84–M'89) received the B.S.E.E. degree from the National Taiwan University, Taipei, R.O.C., in 1980, and the M.S. and Ph.D. degrees in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY, in 1985 and 1988, respectively.

From 1984 to 1988 he was with the ECSE Department at Rensselaer Polytechnic Institute as a Research Assistant. He is currently with Bell Communications Research in Morristown, NJ. His research interests include image segmentation, video compression, pattern recognition, and digital communications.

Dr. Jeng is a member of Sigma Xi.



Sharad Singhal (S'78–M'82) received the B.S. degree in electrical engineering from the Indian Institute of Technology, Kanpur, India, in 1977 and the M.S. and Ph.D. degrees from Yale University, New Haven, CT, in 1978 and 1982, respectively, both in electrical engineering.

In 1982 he joined the Acoustics Research Department at AT&T Bell Laboratories in Murry Hill, NJ, where he worked on speech coding. He joined Bell Communications Research in 1984, where he first was a Member of the Technical Staff, and then a District Manager of the Visual Communications Research Group. His research interests are in the areas of speech, audio and video coding, speech recognition, and digital signal processing. He is currently Director of the Speech Recognition Application Group.

Dr. Singhal is a member of the Acoustical Society of America.