Selective Recovery of Video Packet Loss Using Error Concealment

MASAHIRO WADA

Abstract—**A new efficient recovery method using error concealment is proposed for video packet loss in fast packet switching networks. In the packet switching network, packet loss is unavoidable and a correction method is required for video packet transmission which takes characteristics of video signals and video coding into account. In the method, the receiver detects the damaged picture area caused by packet loss from the structured picture data received, makes error concealments, notifies the transmitter, and continues decoding. The transmitter, having received the notice, calculates the affected picture area in the local-decoded picture and continues encoding without using this affected area. In this selective recovery method, video signals are not stopped even if a long propagation delay exists, no additional information is transmitted, and conventional coding algorithms can be applied. Also, the proposed method is suitable for multipoint communication. The relation tables of picture block propagation required at the transmitter consist of small amounts of bits. Simulation results show the affected picture area is localized for a considerable time, and concealed picture quality is sufficient.**

I. Introduction

FAST packet switching is considered the major candidate for building broadband ISDN [1] where video communications will be a main service. Fast packet switching can efficiently handle variable bit rate signals, and is suitable for bit-rate-reduced video signals [2].

Since fast packet switching achieves high utilization efficiency of circuits by statistical multiplexing, unavoidable packet loss occurs which causes serious errors in low-bit-rate video transmission where interframe coding is adopted. A method is therefore required for video packet transmission which recovers packet loss efficiently [3]. Unlike data communication, video communication is usually interactive and real-time critical; therefore, conventional retrieval protocol like HDLC is inapplicable when propagation delay exists.

This paper proposes a packet loss recovery method for low-bit-rate video transmission in the packet switching network which uses error concealment and selective correction. The method has higher efficiency than a conventional demand refresh which forces the encoding mode of a whole frame into intraframe. It also allows free-running of video signals, transmits no additional information for error recovery, and is applicable to conventional coding algorithms. Futhermore, the proposed method is suitable for multipoint communication.

Manuscript received August 11, 1988; revised November 1, 1988.
The author is with KDD Kamihukuoka Research and Development Laboratories, Kamifukuoka-shi, Saitama-ken 356, Japan.
IEEE Log Number 8927585.

In the following sections, Section II describes some requirements which restrict coding structures, Section III explains the selective recovery method, Section IV discusses transmission data structure, and Section V gives some simulation results which show the method is effective.

II. Restrictions on Coding

The proposed method requires the following restrictions on coding. These conditions seem adequate for the recent coding algorithms used for low-bit-rate video communications.

1) Localization of coded signals: Coding is carried out independently in one frame by each picture area called a picture block. Loss of some picture blocks does not affect other blocks in the same frame, i.e., decoding of the frame can be continued after loss of picture blocks.

2) Localization of packetized signals: Picture blocks are packed into one or several packets in a structured manner and the influence of packet loss is limited to one or several packets, i.e., the received bit streams are decodable after a packet loss.

For restriction 1), most picture coding algorithms currently used for low and medium bit rates are adopting interframe/intrablock adaptive prediction based on pixel blocks [4], [6], and this trend will increase due to progress in the study of orthogonal transforms such as DCT and vector quantization [7]. Frame-by-frame propagation of packet loss influence basically depends on the range of motion compensation adopted.

Restriction 2) can easily be realized under the condition that restriction 1) is guaranteed. Details of the packing structure of coded data will be described in Section IV.

With these restrictions, the damage of packet loss is localized within a certain area of the picture frame.

III. Selective Recovery Method

A. Basic Procedure at the Receiver

In this method, a receiver terminal operates according to the following procedures.

1) Packet loss is detected by a network or the receiver terminal. In the latter case a sequential number will be added to each packet.

2) Damaged picture block(s) is (are) specified by calculating the received data structure.

3) Notice of packet loss with the address(es) of damaged picture block(s) is sent back to the transmitter.

0733-8716/89/0600-0807$01.00 © 1989 IEEE

4) Error concealment is carried out over damaged picture block(s). For example, error concealment can be done by replacing the damaged picture block(s) with the block(s) in the previous frame.

5) Decoding is continued.

Calculation of damaged picture block(s) depends on the transmitted data structure and is discussed in Section IV.

B. Basic Procedure at the Transmitter

There are two possible procedures at the transmitter as follows.

Method A:

1) Notice of packet loss is received with the address(es) of damaged picture block(s).

2) The affected picture area in local-decoded frames is calculated from the point of the damaged block(s) up to the currently encoded frame as shown in Fig. 1.

3) Encoding is continued without using the affected picture area.

Method B:

1) Notice of packet loss is received with the address(es) of damaged picture block(s).

2) The same error concealment as is performed at the receiver is carried out in the local-decoded frame stored, as shown in Fig. 2.

3) Local decoding is reexecuted from the point of the concealed block(s) up to the currently encoded block using the transmitted data stored so that the local-decoded picture at the transmitter becomes the same as the decoded picture at the receiver, as shown in Fig. 3, where local-decoded picture data stored in memory and concealing data give the initial value for decoding.

4) Encoding is normally continued throughout the above sequence.

In Method A, the amount of calculation is small, only the propagation tree of the affected picture blocks is required to be stored for a depth Td (Td is the response delay, i.e., equals the propagation delay of the transmission link plus the operation delay of packet loss detection), multipoint communication is possible, and overhead due to coding efficiency deterioration will be small when the affected picture area is localized. Moreover, as encoding without using the affected picture area does not necessarily mean intrablock coding, the coding efficiency deterioration itself will also be small.

In Method B, the encoding operation is not interrupted, the local-decoding process for recovery can run concurrently, and efficiency loss of encoding will be negligible. However, local-decoded pictures and transmitted data have to be stored for the above-mentioned depth Td at the transmitter, and multipoint communication is difficult.

From these comparisons, here we choose Method A, but due to its high efficiency and absence of effect on coding operation, Method B is still attractive when Td is small (in extreme cases, when the concealed picture area does not yet affect the encoding, it is required only to replace the corresponding area in a local-decoded frame by concealing data).

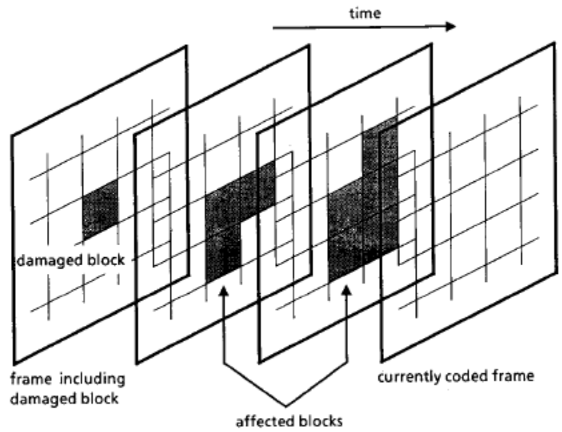


Fig. 1. Illustration of Method A.

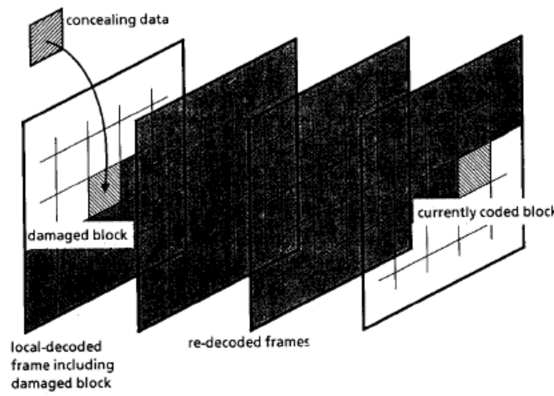


Fig. 2. Illustration of Method B.

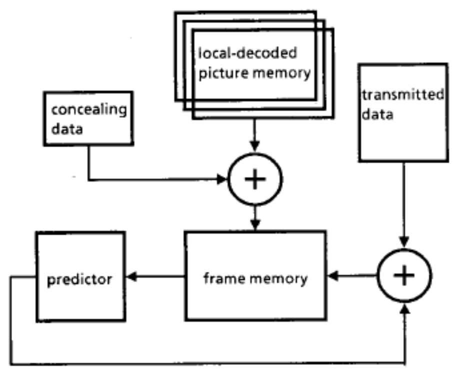


Fig. 3. Redecoding algorithm of Method B.

C. Propagation Tree

A propagation tree is used at the transmitter to specify the affected area in the local-decoded pictures. Fig. 4 shows an example of the tree where one block is damaged in frame i and three, four, and five blocks are affected in frame $\underline{i+1}$, $\underline{i+2}$, and $\underline{i+3}$, respectively.

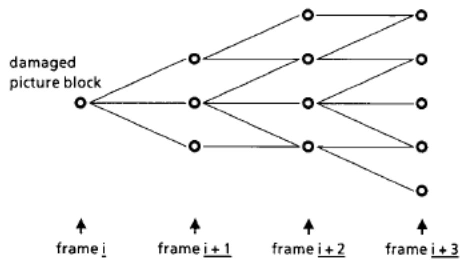


Fig. 4. An example of block propagation tree.

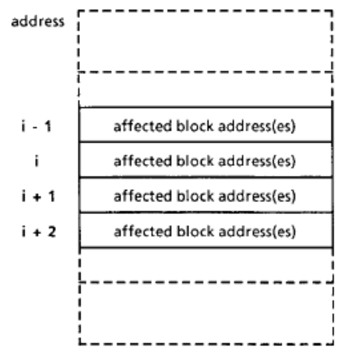


Fig. 5. Structure of block relation table.

In order to calculate the tree, relation tables shown in Fig. 5 are made for each frame simultaneously in the encoding operation. Each address of the table corresponds to the picture block address in the frame, and contains the address(es) of block(s) which used (a) pixel(s) in the addressed block for encoding.

One address may contain several blocks, and the maximum number is basically determined by the range of motion compensation adopted. The total number of blocks contained in a table, however, is at most four times as much as the number of blocks in one frame when an orthogonal block lattice is used.

When packet loss is notified, the table for the frame having the corresponding picture area is first searched by the informed address(es) of damaged block(s), and the address(es) of the affected block(s) is (are) selected. This procedure is repeated, and the propagation tree is thereby constructed.

IV. Transmitted Data Structure

A. Premises

In a fast packet switching network, transmitted signals are divided into smaller units named cells and then multiplexed [1]. The cell bit length is usually fixed for stable synchronization, although a variable length cell is more efficient. Here, therefore, the cell length is assumed fixed. On the other hand, a coded picture block has a variable bit length, and therefore efficient packing of each picture block has to be considered in this case. Here, we present an example of the transmitted data structure which satisfies the requirements mentioned in Section II.

In order to simplify the discussion, the following premises are made.

1) When conditional replenishment is adopted, nothing is transmitted on nonreplenished picture blocks. This is effective for increasing coding efficiency.

2) Packet loss is detected by the network or terminal by procedures outside the scope of this paper.

3) Coded data are sequentially decodable. Unique words are not used for efficiency.

These premises lay few constraints on coding algorithms.

B. Data Structure [5]

Two cases have to be considered, i.e., that the bit length of a picture block is shorter than a packet length, and that it is longer.

Fig. 6 shows the transmitted data structure. When a picture block length is shorter than a packet length, one packet can contain several blocks as is shown in Fig. 6(a), but a picture block is prevented from spreading over two packets. The first block in a packet has an absolute address (AA) for localization of packet loss influence, while (an)other block(s) in the same packet has(ve) a relative address (RA) for addressing efficiency. Valid information in a packet is terminated by an end-of-information code (EOI). This code can be used to reduce the packaging delay when the block bit length remains small.

When a picture block length is longer than a packet length, one picture block is divided into several packets as is shown in Fig. 6(b), where an absolute address (AA) is attached at the head of the block information in the first packet while a continuation flag is attached at the same area in other packet(s). It may be prevented from including the next picture block in the last packet in order to limit packet loss influence.

In this data structure, tradeoff between packaging efficiency and error resilience should be considered. If error resilience is considerable, one packet should contain a smaller number of blocks. Picture block size should also be determined from this point.

C. Operation of the Receiver

Based on the transmission data structure mentioned above, the receiver terminal operates according to the state transition diagram shown in Fig. 7. In the figure, three normal states and two temporary states exist.

"Idle" means no packet is under processing, and when a packet is received with a correct sequential number (event CPR), it transfers to a "decoding" state with absolute block address and an incrementing packet sequence number. The picture block information in the received packet is then decoded.

In this state, if decoding of one block is terminated (event EOB), the received block address is renewed and it transfers to "next address check." If decoding of one

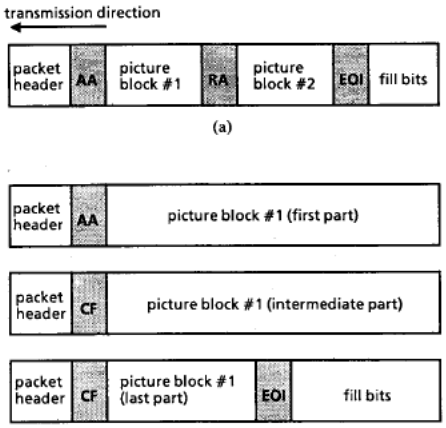
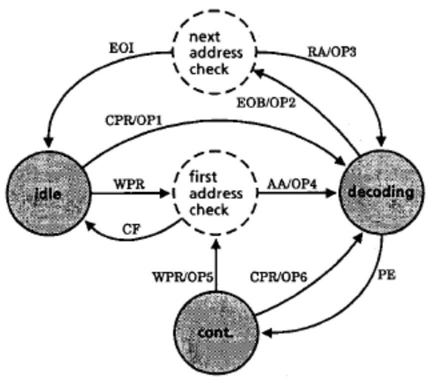


Fig. 6. Structure of data packing. (a) Picture block length < packet length. (b) Picture block length > packet length.



trigger events	
CPR	correct packet reception
WPR	wrong packet reception
EOB	end of block
EOI	end of information
AA	absolute address
RA	relative address
PE	packet end
CF	continuation flag

operations	
OP1	get absolute block address increment packet sequence number
OP2	renew received block address
OP3	get relative block address
OP4	error concealment notice of damaged picture block(s) reset packet sequence number get absolute block address
OP5	discard stored information
OP6	increment packet sequence number

Fig. 7. State transition diagram for packet receiving.

packet is finished before EOB (event PE), a state transition occurs to "cont."

In the temporary state "next address check," an address field after the processed block is examined and if a relative address is detected (event RA), decoding of the next block in the same packet restarts. In case that EOI code is detected (event EOI), it transfers back to the "idle" state.

In the "cont." state, when the next packet is received with a correct sequential number (event CPR), the packet sequence number is incremented and a state transition occurs to "decoding." Decoding of the next part of the same picture block is then continued.

When a packet loss has occurred and the next packet is received with a wrong sequential number (event WPR) in the "idle" state, the first address field in the packet is checked. If an absolute address is detected (event AA), this means synchronization of picture blocks is resumed, and error concealment over damaged picture block(s), notification of damaged block(s), and reset of the packet sequence number are carried out. After these operations, the state transfers to "decoding," and decoding of the picture block with the absolute address starts. If a continuation flag is detected (event CF), however, no operation takes place and the state remains in "idle."

When the same event WPR occurs in state "cont.," some of the information of the decoded picture block stored in memory is discarded, and the same procedures follow.

Damaged picture blocks which the receiver estimates may contain noncoded blocks which have not been transmitted. If the error concealment operation is the same as that of the noncoding, the overheads of the transmission and transmitter operation required for recovery can be reduced.

V. Simulation and Discussion

This section evaluates the proposed recovery method by simulation experiments using three kinds of video signals from videoconferencing/videophone sequences, "Salesman," "Miss America," and "Claire," which are used for CCITT standardization. Picture frame size is 352 pels $\times$ 288 lines, each pel consists of Y, $R-Y$, $B-Y$, each having 8 bits, and the frame rate is 10 frames/s. The coding algorithm used here is based on motion-compensated interframe prediction and discrete cosine transform (DCT) with conditional replenishment [4]. Motion compensation range reaches ± 7 pels/frame in both directions, and the picture block size is 16 pels $\times$ 16 lines (i.e., 396 blocks/frame). The step size of the quantizer is fixed, and the SN ratio of the coded picture becomes stable. The relation between the step size and SN ratio depends on used pictures. Table I shows characteristics of coding results averaged for 32 frames.

A. Characteristics of Propagation Tree

In the proposed method, the influence of damaged picture block(s) is propagated through consecutive frames, so the speed at which this influence spreads greatly affects the coding efficiency, because these affected picture blocks cannot be used for encoding. Fig. 8 shows the number of affected blocks in subsequent frames after loss of one picture block. These results are averages of coded (replenished) picture blocks for 16 frames. The number of affected blocks increases moderately, and the rate of increase falls as the frame depth advances.

Differences of slopes can be explained by the ratio of motion-compensated (MC) blocks among all coded

TABLE I
CHARACTERISTICS OF CODING RESULTS

sequence name	Salesman			Miss America			Claire		
quantizer step size (×1/256)	8	12	16	8	12	16	8	12	16
luminance SN ratio (dB)	38.8	36.3	34.6	41.3	40.0	39.0	43.6	41.6	40.0
coded blocks/all blocks ratio	0.98	0.82	0.65	0.97	0.67	0.53	0.53	0.37	0.32
MC blocks/all blocks ratio	0.22	0.17	0.14	0.33	0.32	0.29	0.11	0.11	0.12
MC blocks/coded blocks ratio	0.22	0.21	0.21	0.34	0.48	0.54	0.21	0.30	0.37
coded bitrate (kbit/s)	665	333	200	378	182	102	204	122	86

MC : motion-compensated
coded : replenished

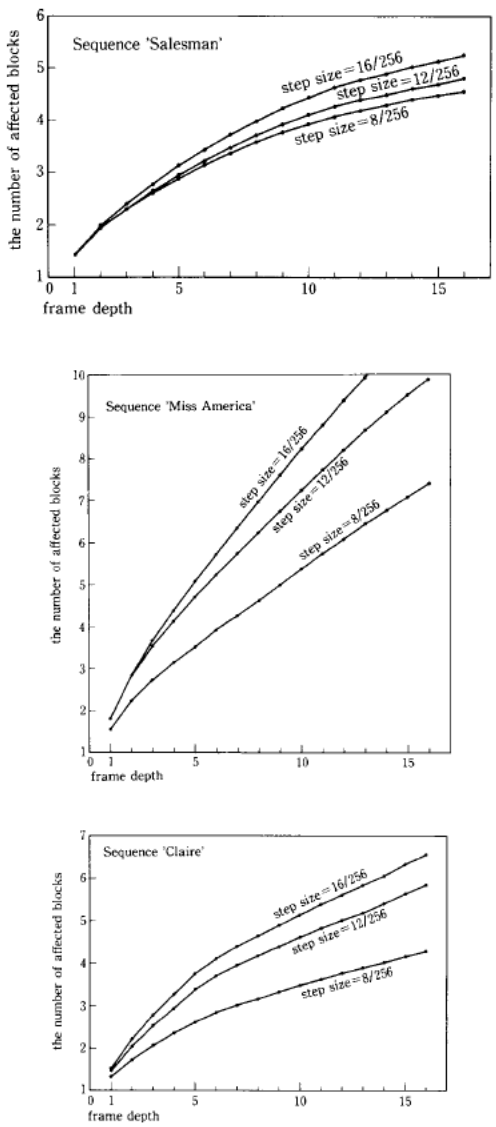


Fig. 8. The number of affected blocks after loss of one picture block.

blocks, i.e., if this ratio is high, the number of blocks which is used for encoding one block becomes high. For the "Miss America" and "Claire" sequences, as is shown in Table I, this MC/coded ratio also increases as the step size of the quantizer increases because the number of MC blocks does not vary greatly, while the percentage of coded blocks decreases. The slopes of the graphs therefore increase according to the step size. On the other hand, this MC/coded ratio does not vary for the "Salesman" sequence, and therefore the slopes of the graphs for this sequence are not very different.

The reason that the number of affected blocks does not increase exponentially is that the picture blocks in the next frame which are affected by the neighboring picture blocks in the previous frame overlap to some extent. Actually, the number of affected blocks is limited to 10 per 396 after 13 frames (1.3 s) even for the most severe case.

As is described in Section IV, one packet can contain several picture blocks, so a loss of consecutive picture blocks may occur. Fig. 9 shows the number of affected blocks in subsequent frames after loss of consecutive picture blocks. These results are averages of coded (replenished) picture blocks for 24 frames, and the quantizer step size is fixed to $12/256$ (this means pel values range from 0 to 255 and the step size for coefficients in the transformed domain is fixed to 12). The slopes of the graphs vary very little for a different initial number of blocks, and all graphs are in a parallel shift relation. The ratio of increase per initial block is almost constant in the same sequence, and it varies from 1 to 2 among sequences. The reason why there is no exponential increase is the same as above.

These experimental results and considerations show the proposed method is efficient even if the response delay is long or one packet contains several blocks.

Table II shows the required overhead for error recovery by the proposed method and conventional demand refresh method, i.e., the ratio of bit increase per normally coded bits when error recovery is carried out in the fourth frame after packet loss. These values are averages for all coded (replenished) picture blocks in the frame where packet loss occurs. In the proposed method, a simple coding mode decision algorithm is adopted where blocks which are normally interframe coded using affected picture blocks

TABLE II
REQUIRED OVERHEAD RATIO FOR ERROR RECOVERY

sequence name	Salesman			Miss America			Claire		
quantizer step size (×1/256)	8	12	16	8	12	16	8	12	16
proposed method (1block lost)	0.01	0.01	0.02	0.01	0.03	0.05	0.04	0.08	0.10
proposed method (2 block lost)	0.01	0.02	0.02	0.02	0.04	0.07	0.06	0.11	0.15
conventional demand refresh	1.35	1.93	2.50	0.97	1.83	3.00	2.69	3.98	4.87

Error recovery is carried out in the 4th frame after packet loss

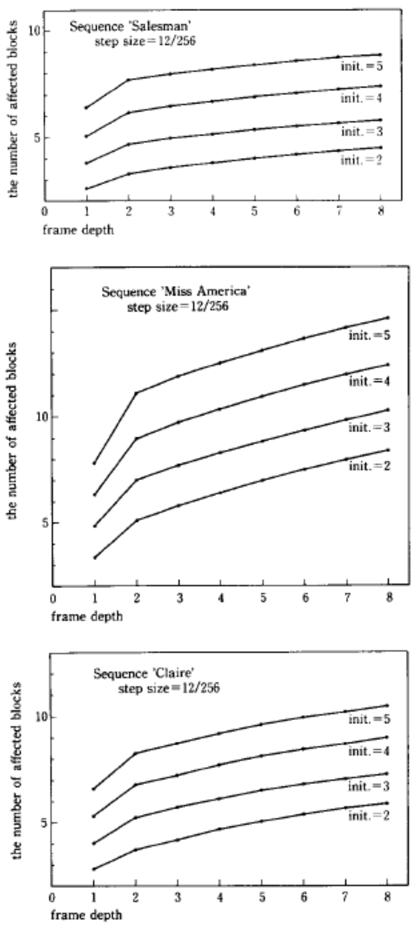


Fig. 9. The number of affected blocks after loss of consecutive picture blocks.

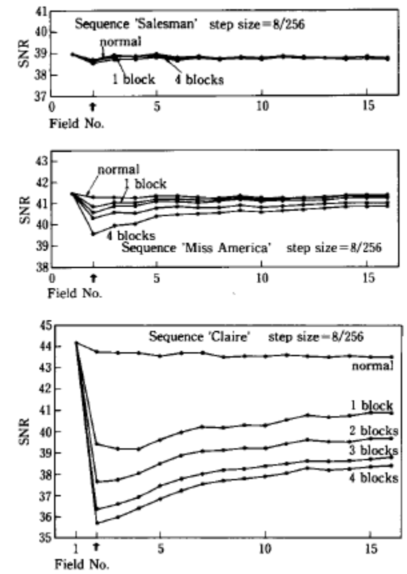


Fig. 10. SNR of decoded pictures after error concealment in the second frame.

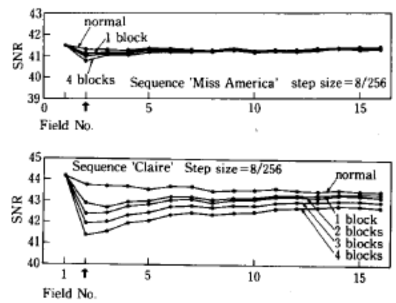


Fig. 11. SNR of decoded pictures after motion-compensated error concealment in the second frame.

are intrablock coded. The ratio of bit increase by the proposed method remains very low. It should be noted that this overhead is only required for one frame where error recovery coding is performed.

B. Picture Quality

Decoded pictures remain erroneous for several frames after error concealment has been made, and picture quality in this period therefore decides the validity of the method. Picture quality by error concealment, however, greatly depends on the area where a block loss occurs. For example, if a block loss occurs on the background, the influence on picture quality will be small, while if it occurs on the face of a person, it will damage the picture.

Here, we examined the picture quality deterioration in the worst case, i.e., when a block loss occurred on the eyes of persons and the quantizer step size is small as described later. Fig. 10 shows the SNR of decoded pictures after error concealment in the second frame. Blocks

Fig. 12. Decoded picture of the second frame.

in the previous frame were used for concealment and the number of concealed blocks is 1 to 4. SNR falls initially, but after several frames it grows better rather than worse. This means coding data which refresh picture blocks decrease mismatch between the local-decoded picture of the encoder and the picture of the decoder.

The difference between the sequences can be explained as follows. If the ratio of coded blocks is high, the concealed area is frequently refreshed and the mismatch is alleviated. Furthermore, if the normal SNR of the coded picture is low, the influence of concealment itself will be small. This can also explain the fact that the drop of SNR caused by block loss decreases as the quantizer step size grows.

The results of subjective estimation indicated that the picture quality achieved by the concealment method mentioned above was sufficient for the "Salesman" and "Miss America" sequences, but some block noise could be noticed for the "Claire" sequence which has very high SNR.

In order to reduce the deterioration of SNR and increase the picture quality, motion compensation is effective for error concealment. Fig. 11 shows the SNR of decoded pictures after motion-compensated error concealment in the second frame. Blocks in the previous frame shifted by the average motion vectors of neighboring blocks of the concealed area were used for concealment. The SNR is greatly improved by motion compensation.

Fig. 12 shows the normally decoded pictures and pictures of the frames where a picture block around the eyes is lost and concealed with motion compensation. Slight block noise can be seen in the frame, but this noise is hardly visible in the actual moving image. This is because coding data which refresh picture blocks immediately decrease the concealment noise.

VI. Conclusion

This paper proposes a new packet loss recovery method for low-bit-rate video transmission in packet switching networks which uses error concealment and selective correction. Merits of the method are summarized as follows.

1) The method has higher efficiency than a conventional demand refresh even when a long response delay of packet loss detection exists.

2) Video signals are not stopped even when a long propagation delay exists.

3) No additional information for error recovery has to be transmitted.

4) The method is applicable to conventional coding algorithms and suitable for multipoint communication.

Detailed analysis of the efficiency loss of coding and the efficiency of packing data into packets requires further study.

Acknowledgment

The author would like to thank Dr. K. Ono, Director of KDD Kamifukuoka R&D Laboratories, and Dr. H. Yamamoto, Director of KDD R&D Administration Department, for their encouragement and support of this study. He also thanks Dr. H. Murakami, Manager, and his colleagues in the Visual Communication Laboratory for providing many helpful discussions.

References

[1] P. Gonet, "Fast packet approach to integrated broadband networks," *Comput. Commun.*, vol. 9, Dec. 1986.

[2] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson, and J. D. Robbins, "Performance models of statistical multiplexing in packet video communications," *IEEE Trans. Commun.*, vol. COM-36, pp. 834–844, July 1988.

[3] F. Kishino and H. Yasuda, "Video packet coding for multiple video sources," in *Proc. IEEE COMSOC Int. Workshop Future Prospects Burst/Packetized Multimedia Commun.*, Nov. 1987, Ohsaka, Japan.

[4] S. Okubo, R. Nicol, B. Haskell, and S. Sabri, "Progress of CCITT standardization on n × 384 kbit/s video codec," in *Proc. GLOBECOM '87*, vol. 1, Tokyo, Japan, Nov. 1987, pp. 36–39.

[5] M. Wada and H. Yamaguchi, "Multimedia transmission method for 64 kbit/s integrated visual telecommunication" *Trans. IEICE*, vol. J70-B, pp. 451–460, Apr. 1987 (in Japanese).

[6] S. Ericsson, "Fixed and adaptive prediction for hybrid predictive/transform coding," *IEEE Trans. Commun.*, vol. COM-33, pp. 1291–1302, Dec. 1985.

[7] N. Nasrabadi and R. King, "Image coding using vector quantization: A review," *IEEE Trans. Commun.*, vol. COM-36, pp. 957–971, Aug. 1988.

Masahiro Wada was born in Yokohama, Japan, in 1956. He received the B.S. and M.S. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1979 and 1981, respectively.

He joined Kokusai Denshin Denwa (KDD) Company, Ltd., Tokyo, Japan, in 1981 and has been at the Research and Development Laboratories since 1983, where he has been engaged in research on the coding and communication systems of video signals. His current interests are the development and standardization of a videophone system, and study of packet video communication for the ATM network.

Mr. Wada is a member of the Institute of Electronics, Information, and Communications Engineers of Japan and the Institute of Television Engineers of Japan.