Copyright © 1973 American Telephone and Telegraph Company
THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 52, No. 1, January, 1973
Printed in U.S.A.

A Frame-to-Frame *Picturephone*® Coder For Signals Containing Differential Quantizing Noise

By D. J. CONNOR, B. G. HASKELL, and F. W. MOUNTS

(Manuscript received July 18, 1972)

The frame-to-frame coder described in Ref. 1 used an 8-bit PCM signal for input. If, instead, the signal is obtained by digitally integrating the output of an element difference coder, the quantization noise may be misinterpreted as motion, and cause unnecessary transmission. In the particular example of the Phase I coder,[2] the quantization noise loads the frame codec to the extent that it produces an unacceptable picture.

In this paper, a frame-to-frame coder for Picturephone® *signals is described which is capable of coding the digital output of a Phase I codec for transmission over a 2-megabit/second channel. Improved methods are used to segment the noisy picture into moving areas and background areas. The moving areas are then transmitted using a number of data reduction techniques. During periods of slow movement, clusters of frame-to-frame differences in the moving area are transmitted. For moderate movement, frame differences are sent only in every other field, the moving areas of intervening fields being transmitted by a conditional field interpolation technique. For rapid movement, 2:1 horizontal subsampling is used, and, finally, during violent motion when the buffer fills, frame repeating is used.*

The picture quality obtained from a laboratory simulation of this system is believed to be satisfactory even for a very active subject. With small amounts of motion the subjective quality is actually improved because the visibility of the quantizing noise from the Phase I codec is reduced by the inherent frame repeating action of the coder.

I. INTRODUCTION AND SUMMARY

In Ref. 1 an 8-bit-per-picture element (pel) *Picturephone*-type signal is coded using only 2 megabits/second (Mb/s). Clusters of significant frame differences are transmitted using a double-length code (four-bit and six-bit) for the frame differences and eight-bit addresses for the clusters. During periods of moderate movement, every other significant frame difference along the line is transmitted, the intervening elements being obtained by linear interpolation. If, during violent motion, the buffer fills, frame repeating is used.

In this system a frame difference was deemed significant if its magnitude exceeded some threshold value ($T = 4, 5, 6, 7$) which depended on the buffer fullness. Two exceptions to this criterion were made, however: (*i*) if a significant change was surrounded on both sides by two insignificant changes, then the change was deemed insignificant, and (*ii*) if two clusters of significant changes were separated by three or less insignificant changes, then the clusters were joined by relabeling the intervening changes as significant.

For maximum flexibility of the *Picturephone* transmission system, it is desirable that an interframe coder be able to accept as an input a signal that has previously been coded by an intraframe coder, such as an element difference coder. Such a signal will have a significantly higher level of quantization noise than an 8-bit PCM signal. The Phase I codec[2] is an example of an element difference coder. Since the quantization noise from this coder has been carefully shaped for minimum visibility, the signal it produces probably contains the highest noise level of any signal likely to be encountered by an interframe coder. Designing an interframe coder to work with such a signal thus reveals many of the problems involved in working with signals having realistic noise levels.

If the input signal contains element differential quantizing noise, the system in Ref. 1 does not perform well at all. An inordinate number of sizable frame-to-frame differences arise due to the quantizing noise, and in the case of the Phase I codec, acceptable video transmission at 2Mb/s is impossible. Raising the threshold of significance reduces the number of background frame differences which are transmitted, but it also reduces the number of subjectively important frame differences in the moving area which are sent. Unacceptable picture quality results.

Averaged over a small region in space and time, the frame differences due to quantizing noise differ in many ways from the subjectively important frame differences due to movement. For example, frame differences due to movement are correlated spatially, whereas frame differences due to quantizing noise are not.

These properties have been exploited to give a method for segmenting the picture into moving areas and stationary areas.[3] The moving area as defined by the segmenter tends to be slightly larger than the actual moving area, but it has been found that this is necessary if a subjectively acceptable picture is to be obtained.

The number of picture elements which must be transmitted using the noisy input and this segmenter is much larger than with the 8-bit input and the segmenting criterion of Ref. 1. Thus, even with a good segmenter, the data rate is larger than 2 Mb/s using only the data reduction techniques of Ref. 1. Other means of data compression are required if a 2-Mb/s rate is to be obtained.

Two techniques are proposed. First, since the segmenting criterion used here requires that all picture elements in the moving area be transmitted, a large number of zero frame differences are sent, i.e., the average transmitted frame difference is much smaller than in Ref. 1. Under these circumstances, variable word length codes can be used to good advantage. Using a variable word length code optimized for moderate motion, only about two bits per frame difference are required on the average. Using this same code during periods of active motion requires about three bits per frame difference on the average.

Using the new segmenter and variable word length coding of frame differences, transmission below 2 Mb/s is easily accomplished during periods of slow movement. When motion becomes a little more rapid, however, the 2-Mb/s rate is surpassed, and another data compression technique must be used. Two-to-one horizontal subsampling generally results in subjectively unacceptable picture quality because the movement is too slow to hide the resolution loss. Thus, a conditional field interpolation technique[4,5] is used as the second method of data rate reduction.

With this technique, frame differences in the moving area are transmitted only during every other field. Each pel in the moving area of the intervening fields is obtained at the receiver by a four-way average of vertically adjacent picture elements in the two fields adjacent to the one being coded. However, if the four-way average is in error by an amount larger than some prescribed threshold, then a quantized correction value must be sent to maintain acceptable picture quality.[4]

The receiver as described above would still have to be told which picture elements in the intervening field are in the moving area, and which are in the background. However, since movement is so highly correlated from field to field, we believe that this information can be extracted from the two fields adjacent to the one being interpolated.

With rapid motion, 2:1 horizontal subsampling can be employed. This is brought in under buffer control. When motion becomes violent and the buffer fills, then transmission ceases for one frame period and the previous frame is repeated.

Using the data compression techniques described above, a laboratory simulation was constructed to test the important aspects of a 2-Mb/s frame-to-frame codec that is capable of coding the digital output of a Phase I codec. A simplified block diagram of the simulation is shown in Fig. 1. A digital signal identical to the output of the frame-to-frame codec was passed through another digital Phase I codec without degrading the picture noticeably. The system described is capable of accommodating about the same amount of movement as that in Ref. 1, with a picture quality comparable to that of the Phase I codec.

The Phase I codec was designed, of course, without any thought of frame-to-frame coding. It is not surprising, therefore, that many difficulties arise when frame-to-frame coding techniques are applied to the output of a Phase I codec. Changes in the Phase I coder to reduce the quantization noise would not only result in a simpler interframe coder, but could also lead to a data rate less than the 2 Mb/s obtained here. How much less will have to await further study.

The next four sections describe in more detail the operation of the frame-to-frame coder. The last section describes the simulation.

II. SEGMENTING THE PICTURE INTO "MOVING" AND "STATIONARY" AREAS

An essential preliminary to the development of the coder described in this paper was the development of methods for detecting or segmenting the moving area in a video signal which has already been corrupted by noise due to an in-frame coding operation. A full description of the work done on this problem will be given in subsequent papers. In this

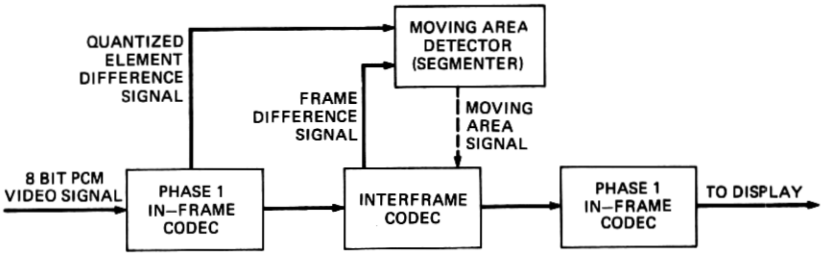

Fig. 1—A simplified block diagram of the simulation showing the signals used and produced by the segmenter.

section, we will simply state the various properties of the video signal and the coding noise which can be exploited in detecting the moving area. Following that, we give a description of the actual segmenter that was developed for use in the system described in this paper.

In order to separate the frame-to-frame brightness changes caused by movement from those caused by noise from an element difference quantizer, advantage can be taken of certain distinguishing properties. The most important property of the movement-generated frame differences is that they are spatially correlated. Two properties of the noise are important:

(*i*) It is almost entirely uncorrelated spatially;
(*ii*) The magnitudes of individual noise spikes are equal to the spacing of the representative levels used in the element difference quantizer.

The second property of the noise results from the fact that in stationary areas a small noise perturbation from one frame to the next can cause a change in the representative level used to encode a particular element difference. This change will be to an adjacent representative level in the quantizing scale, and, consequently, the resultant frame difference will be equal to the spacing of those levels. The more widely spaced outer levels of the companded quantizing scale are used to encode detailed areas and contrasty vertical edges. Thus, the frame difference noise is greatest in these regions.

Finally, a useful property of moving areas is that they are spatially and temporally contiguous. In other words, if a pel is in the moving area, it is highly probable that the spatially adjacent pels and the same pel in the next frame are in the moving area.

The signals employed by the segmenter in detecting the moving area are indicated in Fig. 1. A block diagram of the processing of the quantized element difference signal and the frame difference signal is given in Fig. 2. The frame difference signal undergoes two separate spatial filtering operations which increase the signal-to-noise ratio for the spatially correlated frame differences caused by movement. Filter A is designed to enhance the frame difference signal associated with moving edges and particularly with vertical edges moving horizontally. This signal is characterized by high horizontal spatial frequencies and lower vertical spatial frequencies. By averaging the frame difference signal from adjacent lines, these low vertical frequencies are enhanced relative to the spectrally flat frame difference noise.

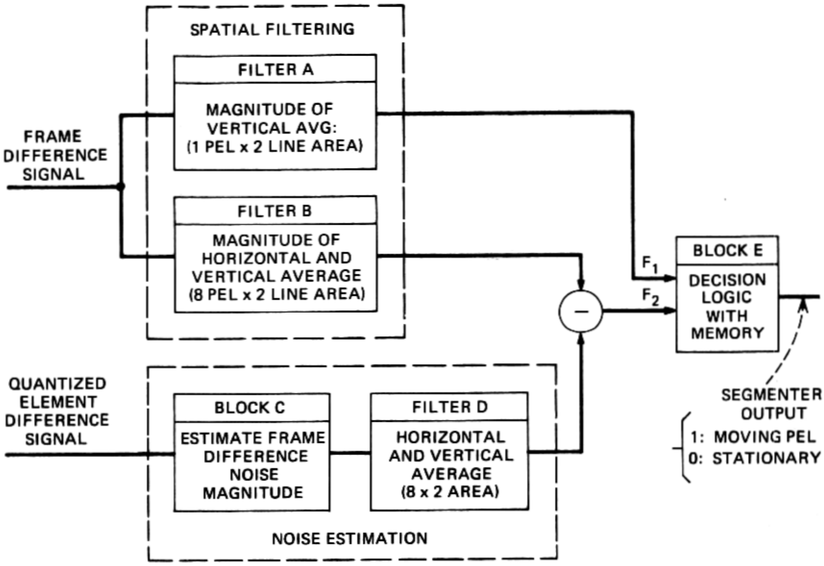

Fig. 2—A simplified block diagram of the segmenter showing the spatial filtering and noise estimation processes.

Filter B is designed to enhance the frame difference signal associated with the movement of relatively flat areas. This signal has most of its energy at low spatial frequencies. By averaging the frame difference signal in an 8-pel-by-2-line area, an increased S/N ratio is obtained. After the averaging operation, the signals from both filters are rectified since the frame difference signal can be of either sign.

Although filter B enhances the movement-generated frame differences in relatively flat areas, it is found that in highly detailed, stationary areas its output commonly exceeds the output arising in slowly moving, flat areas, such as hair. Thus, simple threshold detection is no good. However, it is possible to compensate the output of filter B for these detail-dependent variations in the frame difference noise by subtracting a filtered estimate of the magnitude of the noise signal.

As mentioned above, individual frame differences caused by quantizing noise are equal to the spacing between representative levels of the element-difference quantizing scale. Thus, in blocks C and D in Fig. 2, the filtered estimate of the noise signal is derived from the quantized element-difference signal by generating at the output of block C a non-negative signal that is proportional to the spacing between the input representative level and the adjacent smaller level in the element-difference quantizing scale. (Because the probability distribution of element differences is monotonic and peaks at zero, the most probable transition between representative levels due to a noise perturbation is from an outer level to the adjacent smaller level.) The estimated frame-difference, noise-magnitude assignment for the representative levels is modified for the four inner levels of the 16-level quantizing scale of the Phase I codec as shown in Table I, which gives the output versus input for block C. This modification reflects the fact that noise frame-differences are relatively small in flat areas of the picture. Experimentally it was found that flat, stationary areas could be more easily distinguished from flat, moving areas if no noise compensation was used in these regions. Thus, the estimated frame-difference noise magnitude for the four inner levels is set to zero.

The filtered and noise-compensated frame-difference signals serve as inputs to the decision logic of block E. This logic takes advantage of the fact that moving areas tend to be contiguous both spatially and temporally. Thus, if movement is occurring at a particular pel, there is a high probability that movement is occurring at pels that are spatially and temporally adjacent. Consequently, the philosophy for the design of the decision logic was to use a high decision threshold for the detection of movement in regions of the picture which were previously stationary, and a lower threshold in regions where movement had recently been detected.

A block diagram of the decision logic is given in Fig. 3. (For simplicity, a number of delays required to keep the binary signals in register have not been shown.) The filtered and noise-compensated frame-difference signals serve as inputs to this logic. They are first converted to binary signals by threshold operations having the following transfer characteristics,

$$B_i = 1 \quad \text{if} \quad F \geqq T_i$$

$$B_i = 0 \quad \text{if} \quad F < T_i$$

where F is the input, T_i the threshold, and B_i the corresponding binary output signal. A control signal from the interframe coder that indicates the amount of movement by measuring the buffer fullness is used to raise the thresholds T_1 and T_3 during periods of fast motion.[1] Movement detection is easy in this situation, and the segmenting accuracy can be increased.

In order to best describe the operation of the decision logic, we will start with the block labeled "Binary Threshold Logic with Hysteresis." This block will be referred to as an N out of M (N/M) device after

TABLE I—REPRESENTATIVE LEVELS OF PHASE I CODEC QUANTIZER AND CORRESPONDING ESTIMATES OF FRAME DIFFERENCE NOISE MAGNITUDE

Quantized Element Difference	Estimate of Frame Difference Noise Magnitude
±2/256	0/256
±6	0
±14	4
±30	8
±46	8
±62	8
±78	8
±94	8

Limb and Pease.[6] A block diagram of this device is given in Fig. 4. The accumulator in the N/M device keeps a count of the number of ones in the 8-by-3 block of 24 pels adjacent to the pel of interest as shown in Fig. 5. (Thus, M is 24.) If the output of the accumulator is greater than or equal to the threshold $N_1 = 9$, the output flip-flop is set; and segmenter output function B_5 becomes a one to indicate moving area. In keeping with the design philosophy mentioned above, the flip-flop can only be reset by having the output of the accumulator drop below the lower threshold $N_2 = 4$. Note that by setting N_1 equal to nine, the signal B_3, which indicates the occurrence of flat area movement on the present line can never by itself cause the flip-flop to

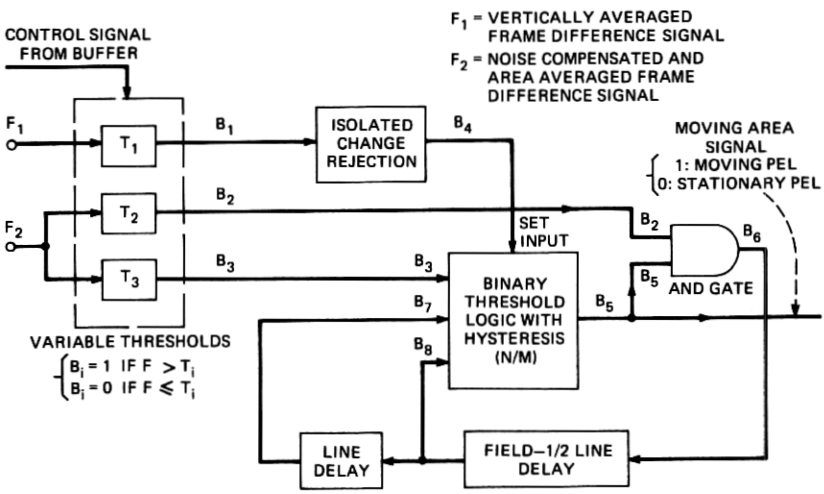

Fig. 3—Decision logic. The N/M device processes binary signals from the present and previous fields to produce the moving area signal.

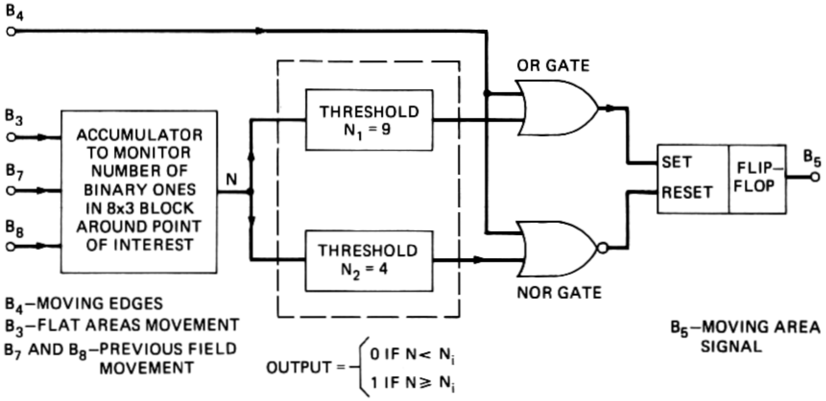

Fig. 4—N/M device. Two thresholds are applied to the output of an accumulator that keeps a count of the number of binary ones in an 8-pel-by-3-line area around the point of interest. These thresholds control the state of the output flip-flop along with the moving edge signal B_4.

be set. Initially, the only way the flip–flop can be set is for a one to occur in the signal B_4. Since this function indicates movement of edges, edge movement must be detected before flat area movement. However, once edge movement is detected, the flip-flop is set and the lower threshold N_2 determines whether adjacent pels on the same line will be designated as moving. In addition, referring to Fig. 3, if B_2, which is a more sensitive but noisier indicator of flat area movement than B_3, is a one when the flip-flop is set, B_6 will be a one. Hence, in keeping with the design philosophy, the value of N_1 for the spatially and temporally adjacent pels in the next field will be effectively lowered by the appearance of these ones in B_7 and B_6. As a result of the interactions described above, the N/M device tends to fill in moving areas, and to designate areas as moving for a short while after they become stationary.

Given the above description of the N/M device, the functions and choice of design variables for the various other blocks in Fig. 3 become

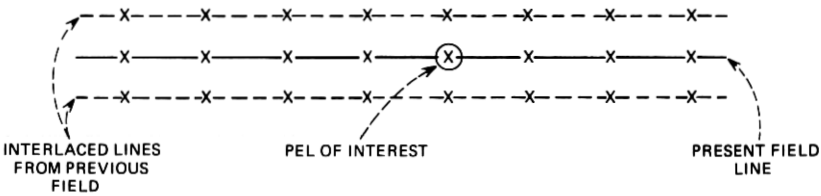

Fig. 5—Arrangement of the 8-by-3 block of pels monitored by the N/M device.

evident. The threshold T_1 is set relatively high ($\sim$10 on an 8-bit PCM scale of 256 levels) to insure that a binary one in the function B_1 is indeed caused by the movement of an edge. This function undergoes further processing so that isolated ones (no other ones within two pels horizontally in either direction) arising from noise spikes are set to zero.[1] Similarly, the threshold T_3 is set relatively high ($\sim$4/256) to insure that the condition $B_3 = 1$ corresponds to movement in flat areas. The threshold T_2 on the other hand can be set lower ($\sim$2/256), since it causes ones to occur in B_6 only if the segmenter output, B_5, is a one. However, it eliminates from B_6 most of the "fill-in" pels generated by the N/M device. This process stabilizes the feedback loop around this device.

If the thresholds T_1 to T_3 are fixed, they must be set quite low in order to detect very slow motion. Given the level of quantization noise from a Phase I coder, such low thresholds inevitably lead to the inclusion of some background points in the moving area. By using the control signal from the buffer, the thresholds can be made speed dependent. For even moderate motion, the segmenting is then virtually ideal.

III. VARIABLE WORD-LENGTH CODING OF FRAME DIFFERENCES

In Ref. 1, the 9-bit frame differences ($-255 \cdots 0 \cdots +255$) were quantized into 64 levels. Since the Phase I coder gives an effective 6-bit signal (6 bits with the seventh bit alternately 0 and 1 along the line), only frame differences that are multiples of 4/256 can occur. This set of frame differences is sufficiently coarsely quantized for efficient transmission.

Also, in Ref. 1 it was very much easier to separate the subjectively important frame differences from those few due to camera and system noise. In the system described here, where a Phase I signal is used as an input, once the moving area has been identified, all frame differences in it must be transmitted since it is not possible to tell which are due to movement and which are due to quantizing noise. Within the moving area, as defined by the segmenter, many zero frame differences do occur. However, since they are randomly interspersed among the nonzero frame differences, it is much more efficient to transmit them than it would be to delete them and address the remaining nonzero frame differences.[1]

This causes the average magnitude of transmitted frame differences to be considerably smaller than in Ref. 1 where an 8-bit input is used.

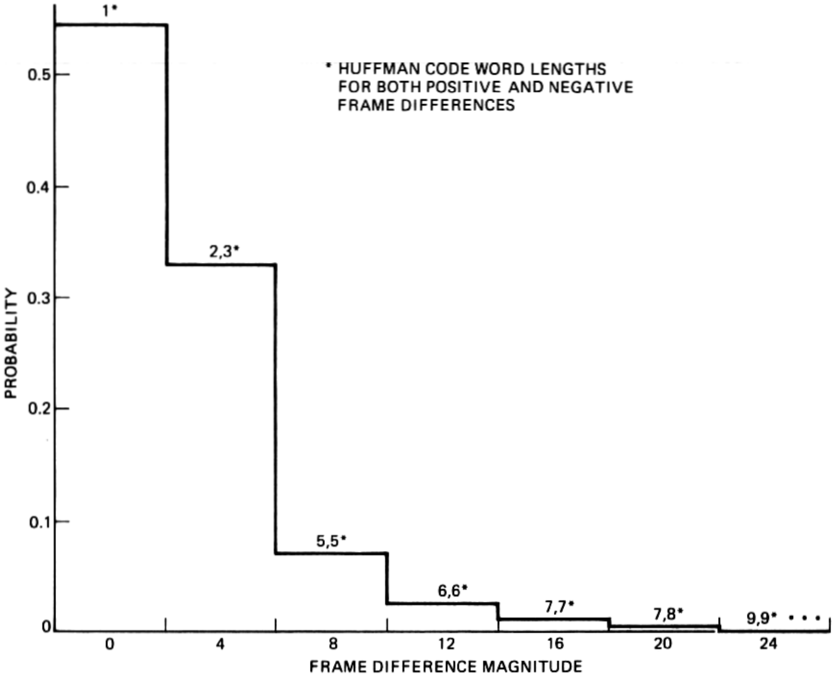

Fig. 6—Typical histogram of moving area frame differences during moderate motion. Huffman code word lengths are shown for each level.

Thus, a more complex variable word-length code can be used to good advantage in reducing the average number of bits required to transmit a frame difference. Preliminary measurements indicate that with a good variable word-length code, less than two bits per frame difference are required on the average during periods of slow movement. During moderate movement, a little more than two bits per frame difference are required; and during rapid movement, about three bits are needed.

Figure 6 shows a typical histogram of the magnitude of the frame differences in the moving area during moderate motion. Also shown are the Huffman code word lengths corresponding to this distribution. The average word length per frame difference is 2.05 bits.

IV. CONDITIONAL FIELD INTERPOLATION

During very low-speed movement, variable word-length coding of frame differences in the moving area is sufficient to code at a rate below 2 Mb/s. Unfortunately, the speed at which the bit rate rises

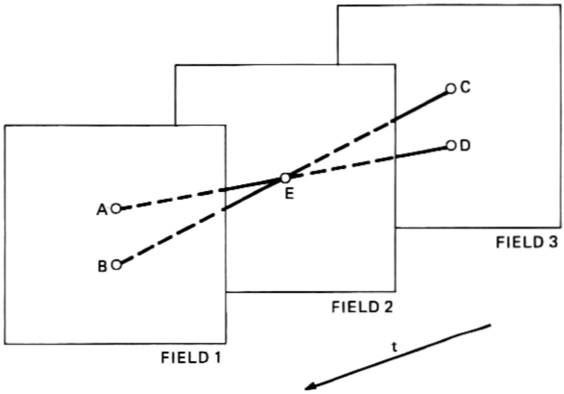

Fig. 7.—Four-way vertical averaging. Fields 1 and 3 are sent via frame differences in the moving area. Information about moving area pels (E) in field 2 is sent only if the interpolation error $|E-(A+B+C+D)/4|$ exceeds a threshold.

above 2 Mb/s is still too slow to hide the resolution loss incurred by 2:1 horizontal subsampling. Thus, another data compression technique is used.

With conditional field interpolation (called conditional vertical subsampling in Ref. 4) only every other field is transmitted by sending frame differences in the moving area. The moving area pels in the intervening fields are obtained from a 4-way average of vertically adjacent pels in the two adjacent fields. In Fig. 7, fields 1 and 3 have been transmitted via frame differences in the moving area, and pel E is to be sent via conditional field interpolation. Pels A and C are directly above E, and pels B and D are directly below E. The 4-way average $(A + B + C + D)/4$ is computed and used as a prediction of E. If the interpolation error does not exceed some prescribed threshold value, then nothing is sent, and the 4-way average is used in place of E. If the interpolation error does exceed the threshold, then a quantized correction value is transmitted.

Since the receiver treats background area in the interpolated fields differently than it does moving area, it must be told which picture elements are in the moving area and vice-versa. Preliminary measurements indicate that addresses for the moving area of the interpolated fields could probably be transmitted using less than 0.1 Mb/s. Alternatively, the moving area of the interpolated fields might be satisfactorily obtained from the union of the moving areas in the two adjacent uninterpolated fields. This would not require any additional information to be transmitted.

In order to determine whether or not the field interpolation error was acceptable, threshold values between 7 and 15 out of 255 were used. These values gave acceptable to marginally acceptable picture quality, and a data rate which was drastically reduced compared with sending frame differences.

V. BLOCK DIAGRAM

Figure 8 shows a block diagram of the system. (The segmenting operation is shown in detail in Figs. 2 to 4.) During very slow movement, every field is transmitted by sending frame differences (B′ − D) in

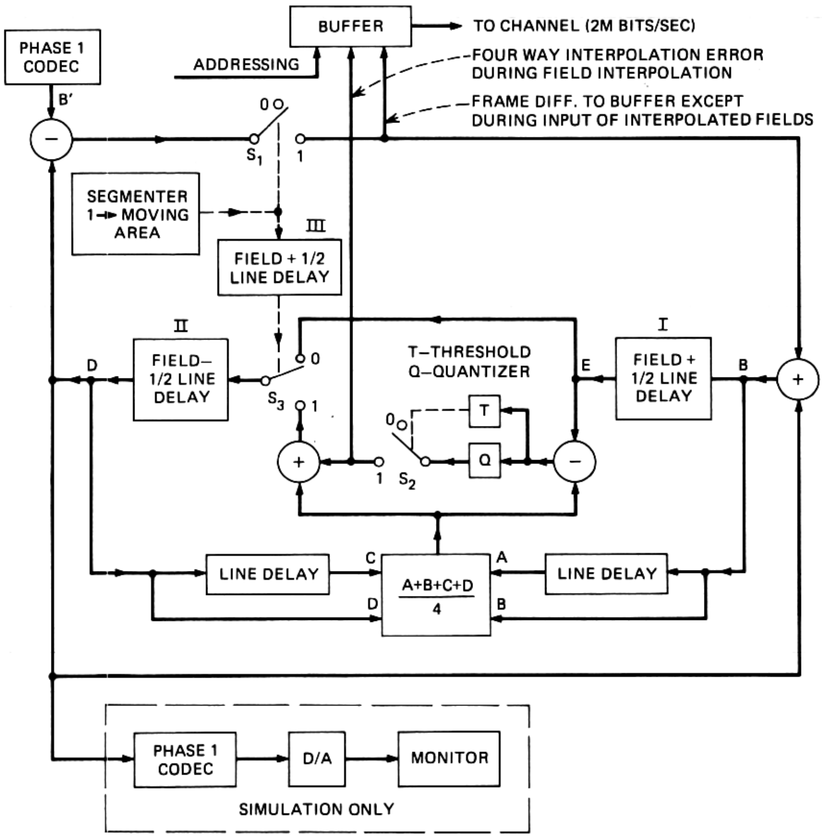

Fig. 8.—Frame-to-frame coder for *Picturephone*® signals with Phase I quantizing noise. During field interpolation, information from two fields is fed to the buffer simultaneously.

the moving area as defined by the segmenter. S_3 is in the 0 position to give an uninterrupted frame memory, and S_2 is in the 0 position so that no interpolation error information reaches the buffer. S_1 is controlled by the segmenter. When in the 0 position, the previous frame value D (see Fig. 7) is fed to delay I, and no frame difference is fed to the buffer. When in the 1 position, the new pel $B' = D + (B' - D)$ is fed to the delay, and a frame difference is fed to the buffer for coding, addressing, and transmission.

When movement becomes more rapid and the buffer fills beyond some prescribed threshold, only every other field is sent via frame differences in the moving area as outlined above. Mode switching occurs only at the end of a field. During input of a field which is to be interpolated, S_2 and S_3 are in the 0 positions allowing uninterpolated fields to enter delay II unchanged. S_1 is controlled by the segmenter as before; however, *no frame differences are fed to the buffer for transmission.* Coding and transmission of this field takes place at a later time. Thus, during input of interpolated fields no amplitude information is fed to the buffer. Addressing information needed to specify the moving area at the receiver could be sent at this time if it is found to be more efficient; however, this information could just as well be obtained from the output of delay III and sent later during the actual coding and transmission of the interpolated fields.

During input of uninterpolated fields, coding and transmission of frame differences in the moving area are carried out as usual by means of switch S_1. However, at the same time, coding and transmission of interpolated fields are also performed. When pel E in an interpolated field (see Fig. 7) emerges from delay I, pels A, B, C, and D are emerging from their respective delays as shown in Fig. 8. The output of delay III identifies E as either a background or a moving area pel.

If E is a background pel, S_2 and S_3 are switched to the 0 positions. E enters delay II and no information is fed to the buffer. If E is a moving area pel, then S_3 is switched to position 1, and S_2 is controlled by the threshold logic T. The threshold logic compares the magnitude of the interpolation error $[E - (A + B + C + D)/4]$ with a prescribed threshold. If the error is smaller than the threshold value T, then S_2 is opened (0 position), nothing is fed to the buffer for transmission, and the 4-way average enters delay II in place of E. If the interpolation error is too large, S_2 is closed (1 position), a quantized interpolation error generated by the quantizer Q is fed to the buffer for transmission, and the corrected interpolation value is fed to delay II in place of E.

A number of implementation aspects have not been discussed.

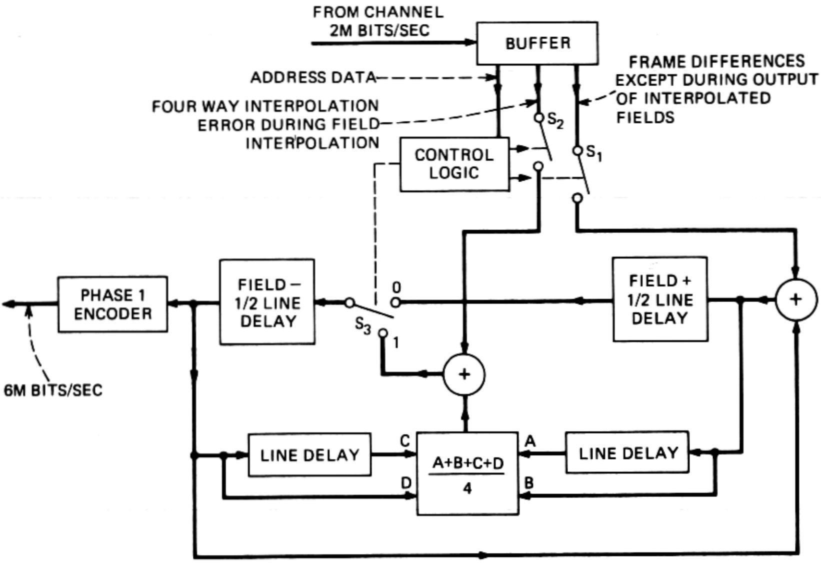

Fig. 9—Receiver configuration. Information is read from the buffer two fields at a time during field interpolation.

For example:

(*i*) All moving area picture element information fed to the buffer must, of course, be accompanied by addressing information, and efficient addressing may require that some of the switch control functions be modified, e.g., isolated point rejection, gap bridging (see Ref. 1).

(*ii*) During field interpolation, information from two fields is fed to the buffer simultaneously. Thus, some multiplexing arrangement must be devised in order to implement the system as described. For example, a buffer might be provided for each field and the outputs switched.

(*iii*) The receiver configuration is very similar to that of the transmitter (see Fig. 9).

(*iv*) Two-to-one horizontal subsampling, and frame repeating have not been discussed here since they are covered elsewhere.[1,6]

VI. SIMULATION OF THE SYSTEM

A number of short cuts were taken to simulate the system described above. First, no coding, buffering or transmission of the data was undertaken. In the simulation, only the picture processing performed at the transmitter was undertaken. The picture which would have appeared at a receiver in the absence of transmission errors was equivalent to the output of field memory II in Fig. 8. As in Ref. 1, buffer control of the picture processing was obtained by using an analog integrator to keep track of the number of bits that would have been in a real buffer had one been built. Also, as in Ref. 1, a buffer size of 67,000 bits was chosen so that it would completely empty if the input of data were stopped for one frame period.

Second, the effect of the variable word-length coding was only partially simulated. Recall from Section III that with a good variable word-length code, pels in the moving area could be transmitted using, on the average, less than two bits per frame difference during periods of slow movement, approximately two bits during moderate movement, and about three bits during rapid movement. This was simulated by counting two bits per frame difference during periods of slow and moderate movement and four bits during rapid movement when 2:1 horizontal subsampling was employed.

During conditional field interpolation, the same bit assignment scheme was used to account for the transmission of interpolation errors. Although transmitted interpolation errors were not quantized in the simulation, preliminary results indicate that they can be quantized quite coarsely. Thus, a 2-bit, 4-bit assignment is not unreasonable.

Transmission of moving area addresses for the interpolated fields was not simulated. Preliminary measurements indicate that with rapid motion, the number of clusters requiring addressing is, on the average, about two per line. If 16 bits are used to address each cluster, then about 0.1 Mb/s would be required to transmit them. If, as was conjectured in Section IV, this moving area can be obtained adequately from the uninterpolated fields, then no extra information need be transmitted.

Finally, transmitted information from interpolated fields was delayed by a field period before being fed to the buffer simulator purely for reasons of expedience. This means that during most of conditional field interpolation, information from two fields does not enter the buffer simulator at the same time as is described in Section V. This should not affect the results very much since much less data is generated during interpolated fields than during uninterpolated fields. However, frame repeating due to buffer filling may occur slightly more often in the actual system than it did in the simulation if the same buffer size is used.

The acceptability of the pictures obtained using the simulation described above was determined mainly by comparison with pictures from the Phase I codec alone. This codec gives pictures that have moderate amounts of both granular noise and edge busyness throughout the picture. The frame-to-frame codec described above transmits information only about the moving area. Consequently, the Phase I codec noise in the background becomes stationary and, hence, much less noticeable. In this sense, the pictures are improved.

Some loss of quality is caused, however, by the use of subsampling. Under some conditions, a slight jerkiness in the movement being depicted is noticeable as the codec enters the vertical subsampling mode. Also, for very high-speed movement, a slight checkered pattern at contrasty edges is detectable. This is caused by the use of both horizontal and vertical subsampling.

On an overall basis, the picture quality produced by this 2-Mb/s codec is felt to be equal to the quality of the input Phase I codec signal.

VII. ACKNOWLEDGMENTS

Many helpful discussions were held with J. C. Candy, J. O. Limb, and R. F. W. Pease. We also acknowledge the technical support of Mrs. M. A. Franke and W. G. Scholes.

REFERENCES

1. Candy, J. C., Franke, Mrs. M. A., Haskell, B. G., and Mounts, F. W., "Transmitting Television as Clusters of Frame-to-Frame Differences," B.S.T.J., *50*, No. 6 (July-August 1971), pp. 1889–1917.
2. Abbott, R. P., "A Differential Pulse-Code-Modulation Coder for Videotelephony Using Four Bits Per Sample," IEEE Trans. Commun. Tech., *COM-19*, No. 6 (December 1971), pp. 907–912.
3. Connor, D. J., Limb, J. O., Pease, R. F. W., and Scholes, W. G., "Segmenting Noisy Television Pictures into Moving and Stationary Areas," unpublished work.
4. Pease, R. F. W., "Conditional Vertical Subsampling—A Technique to Assist in the Coding of Television Signals," B.S.T.J., *51*, No. 4 (April 1972), pp. 787–802.
5. Limb, J. O., and Pease, R. F. W., "A Simple Interframe Coder for Video Telephony," B.S.T.J., *50*, No. 6 (July-August 1971), pp. 1877–1888.
6. Pease, R. F. W., and Limb, J. O., "Exchange of Spatial and Temporal Resolution in Television Coding," B.S.T.J., *50*, No. 1 (January 1971), pp. 191–200.