A Video Encoding System With Conditional Picture-Element Replenishment

By F. W. MOUNTS

(Manuscript received January 29, 1969)

This paper describes an experimental method for encoding television signals which takes advantage of the frame-to-frame correlation to reduce transmission bit rate. The technique encodes only those elements that change between successive frames instead of encoding every element of every frame. We have demonstrated the method in real-time using the head-and-shoulder view of a person in animated conversation as the picture source, such as is likely to be encountered in a visual communication system. An average transmission rate of one bit per picture element gives quality comparable with standard eight-bit PCM transmission.

I. INTRODUCTION

Most known methods for efficiently transmitting pictures over communication circuits exploit point-to-point correlation along a scanning line. In particular, point-to-point predictive quantizing systems have been successful, but it is known that there is more correlation between television picture elements in the frame-to-frame time dimension than there is between adjacent elements in a single frame. This is especially true when using the head-and-shoulder view of a person as the picture source, such as is likely to be encountered in a visual communication system. Now that devices for storing large amounts of data and integrated circuits for complex digital processing are available, it is not only possible to take advantage of this fact in picture coding, but it is also economically attractive.

In this paper, we describe one method of encoding television signals which takes advantage of the frame-to-frame correlation to reduce the transmission bandwidth. We also describe the experimental facility and the results of initial experiments.

We want to emphasize that the picture source for these experiments is a head-and-shoulder view of a person carrying on an animated conversation in real time. Thus, the results are not particularly relevant to stationary graphics or to commercial television where frames are switched, panned, and zoomed. Our experiments have been confined to noninterlaced, 60 frames per second television pictures.

II. GENERAL DESCRIPTION

The technique encodes for transmission only those elements that change between frames instead of encoding every element of every frame. This method has been described previously by R. D. Kell, A. J. Seyler, and T. C. Damen.[1,2,3] Seyler has published statistics of frame-to-frame differences for commercial television signals[4] and has proposed coding methods[5] that are based on this information. E. R. Kretzmer has investigated the correlation between successive frames of motion-picture films showing that there is redundancy that may be exploited.[6] When using video-telephone-like signals with moderate motion in the scene, we find, on the average, that less than one-tenth of the elements change between frames by an amount which exceeds 1 percent of the peak signal. We regard such 1 percent changes as being significant.

We shall describe a complete transmission system that makes use of frame-to-frame redundancy to gain encoding efficiency. The technique which we call "conditional replenishment" is found to be particularly useful for the pictures encountered in visual communication systems. The conditional replenishment system uses a memory to store a reference picture, and only those elements of the picture that have changed significantly between frames are updated (or replenished). Only the picture information necessary to update the reference picture need be transmitted. At the receiver, this information is used to update a similar stored reference picture which is intended to track the one stored at the transmitter.

In order for the receiver to correctly update the picture elements, two pieces of information must be conveyed to the receiver—the new value and the position of the element to be replenished. Because this information occurs at a random rate, buffers are used to redistribute and present the information to the transmission channel at a uniform bit rate. In order to regulate the average replenishment rate to match the channel capacity, the threshold (which determines whether or not a significant change in the picture information has occurred) is varied as a function of the amount of information stored in the buffer.

This method of encoding requires that only the information pertaining to the active region of the picture format be transmitted. The receiver reinserts the horizontal and vertical blanking interval within the reconstructed video information.

III. TRANSMITTER

Figure 1 shows the operations performed by the transmitter. The video signal from the camera is band limited, sampled, and digitized into eight-bit PCM. A selector switch is provided which either conveys new information to the input of the reference frame memory whenever a significant difference is detected or alternatively recirculates the information presently stored in the frame memory. The frame memory consists of delay lines and has sufficient capacity to store one complete frame of video information—each sample encoded as eight-bit PCM.

New information from the camera is compared with the reference picture stored in the frame memory by a subtractor circuit which yields the absolute difference between the new sample of information and the reference value corresponding to the same picture element. During each sample period, the control logic makes a decision, depending upon the magnitude of the difference signal, as to whether a significant difference between the signal values exists. If the difference is significant, the output of the control logic operates the selector switch to strobe the new signal value into the frame memory.

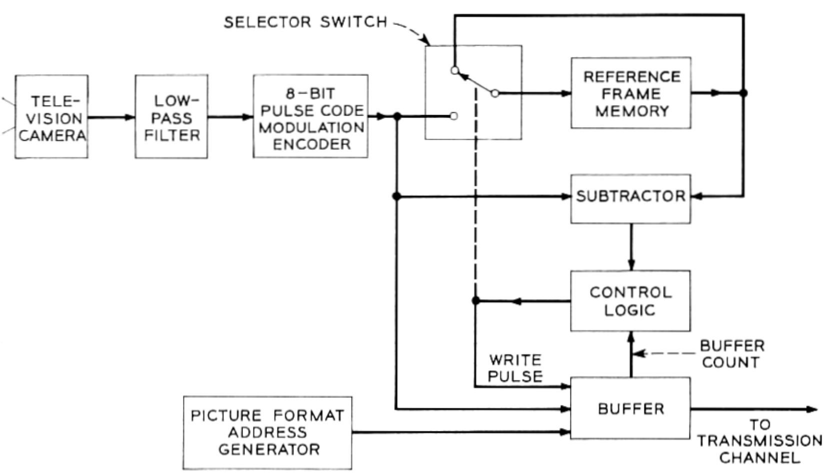

Fig. 1 — Conditional replenishment transmitter terminal.

If the difference is insignificant, the signal value stored in the frame memory is recirculated. In addition to replenishing new information in the frame memory, the control logic also causes the new signal value, accompanied by its address, to be stored in a buffer. The buffer store matches the varying data rate to the constant bit rate of the transmission channel. The information stored in the buffer is read at a constant rate, first-in, first-out.

In the implementation of the experimental system, the amplitude information is expressed as eight-bit PCM with an additional seven bits being used to identify the position information—a total of 15 bits comprising each word transmitted to the receiver. Seven bits in the address is sufficient only to give the horizontal position along the active region of a scanning line. Ambiguity in the vertical position is avoided by always sending the first active sample of each line whether it changes or not. A unique code word defines the first active element of the frame.

In order to force the average replenishment rate to match the channel capacity, the significant change threshold is varied as a function of the amount of information stored in the buffer. This may be accomplished by the control logic characteristic shown in Fig. 2. We express the absolute value of the frame-to-frame difference signal, derived by the subtractor circuit, along the ordinate with a range of 0 to 255 discrete levels. The number of replenished elements stored in the buffer is expressed along the abscissa and may range from 0 to M—the maximum capacity of the buffer. The staircase curve represents the threshold corresponding to each buffer state. The area above the curve represents a significant change in picture information where the control logic forces replenishment. The shaded area below and to the right of the curve represents an insignificant change where the control logic causes the information stored in the frame memory to be retained.

Three properties of this control function should be noted:

(*i*) As the subject becomes more active, causing an increased number of samples to be stored in the buffer, the value of the significant change threshold is increased to permit only the more significant changes in the picture to be replenished. As the subject becomes less active, causing fewer elements to be stored in the buffer, the threshold is decreased in value permitting the less significant changes to be corrected.

(*ii*) It is desirable to keep some minimum amount of data in the buffer at all times so that data is always available for transmission,

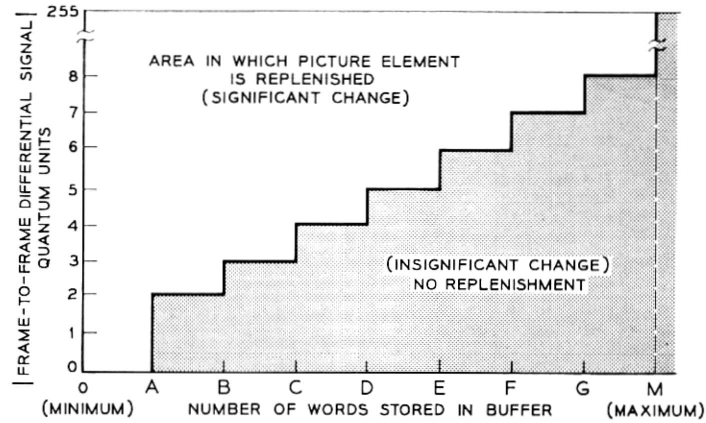

Fig. 2 — Control logic characteristic.

especially in readiness for the vertical blanking interval when data leaves the buffer but none enters. To ensure the buffer does not empty, the significant change threshold is lowered to zero whenever the buffer count falls below a chosen amount.

(*iii*) Whenever the number of samples stored in the buffer is equivalent to the capacity of the buffer, all replenishment is stopped—independent of the frame-to-frame difference. This causes picture breakup as shown in Fig. 5b.

IV. RECEIVER

Figure 3 shows a buffer placed at the receiver to store the received picture information until it can be strobed, in the proper time sequence, into the receiver's frame memory. A transfer of new information from the buffer to the frame memory occurs whenever the output of the picture-format address generator agrees with the address information of the picture element to be read from the buffer. This agreement is determined by the address comparison circuit which operates the selector switch to enable the new amplitude information to flow from the buffer into the frame memory. The buffer readout then advances to the next element. When the addresses do not coincide, the information stored in the frame memory recirculates and the readout cell of the buffer is held fixed. The information stored in the frame memory, when decoded, provides the video information for visual display.

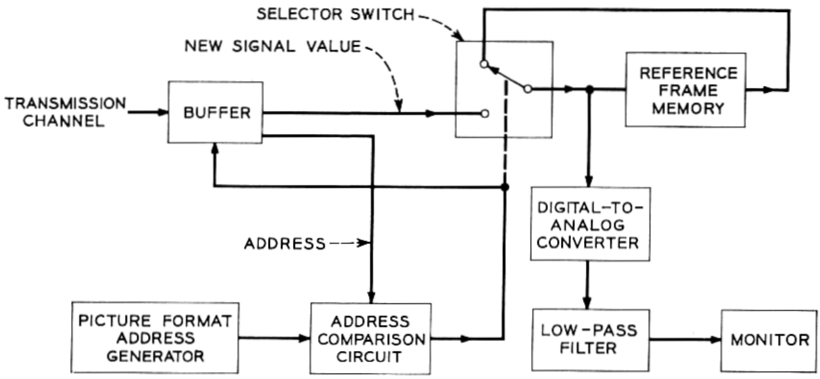

Fig. 3 — Conditional replenishment receiver terminal.

V. EXPERIMENTAL SYSTEM

In order to evaluate this method of encoding video information in real time, only the equipment for the transmitter terminal was assembled as shown in Fig. 4. The information stored in the transmitter's reference frame memory is decoded to recover the video information for visual display. The functional blocks are the same as described for the transmitter except for the buffer which is replaced

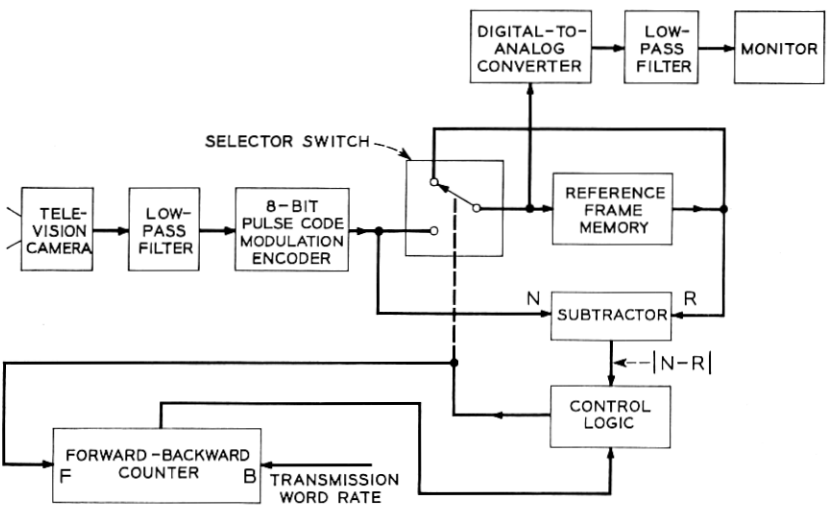

Fig. 4 — Conditional replenishment test terminal.

by a forward-backward counter in order to obtain a count of the data that would have been stored in a buffer had one been used. The count is increased by "one" whenever a picture element is replenished and decreased by "one" each time a word is transmitted. The state of the counter provides feedback to the control logic representing a measure of buffer fullness. In this way, we ignore transmission error and we have eliminated the actual buffer for experimental purposes.

In the initial experiment, the simulated buffer was assumed to have sufficient capacity to hold data relating to as many elements as there are in two complete frames, a total of 51,240 words, each word comprised of 15 bits. In practice, this would result in an inherent one-way signal delay of one-half second. The backward count rate of the counter was set to be one-eighth of the transmission rate required to send the picture directly as eight-bit PCM. The head-and-shoulder view of a person, as might be used in a visual communication system, was used as the picture source. The video information generated by the camera was band limited to 0.75 MHz. The picture format was composed of 140 picture elements per line with 183 lines per frame sequentially scanned at a rate of 60 frames per second. The active region of the picture format was composed of 120 picture elements by 171 lines. The highlight luminance of the picture was 70-80 fL (24-27.4 cd/m^2) and the ambient illumination was 125 fc (1350 lm/m^2).

VI. RESULTS

Photographs of a single frame of video information are used to illustrate the effects of conditional replenishment. These photographs are not very effective in portraying picture quality subjectively since impairments are produced only in the presence of motion.

The following results have been obtained for the experimental system outlined above:

(*i*) When motion is moderate, the picture quality is nearly the same as for eight-bit PCM coding as shown in Fig. 5a.

(*ii*) As the motion of the subject becomes more rapid, the number of picture elements stored in the buffer increases. This causes the significant change threshold to be increased so that small changes in the picture are not reproduced. The reproduction becomes somewhat poorer since all picture elements are not represented with the same

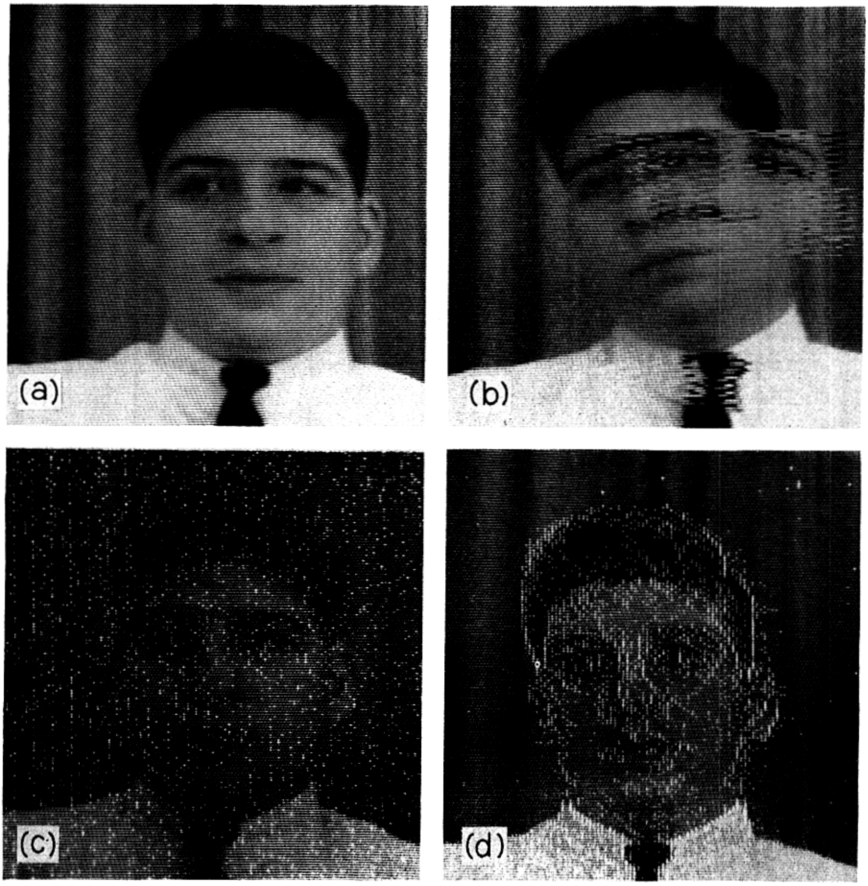

Fig. 5 — Single frames of video information processed in real time by the conditional replenishment system.

accuracy and the result resembles a scene viewed through a dirty window.

(*iii*) When the subject becomes very active, that is, when the picture contains sustained rapid motion covering a large part of the field of view, the buffer becomes saturated, allowing no more changes to be accepted. This condition is referred to as buffer overload and causes picture breakup as shown in Fig. 5b. Picture breakup is only momentary and a quick recovery takes place as soon as the subject slows down. This might be a serious defect except that there is still much that can be done with buffer and threshold strategy to reduce this effect.

Figures 5c and d show pictures depicting the output of the system with markers or flags purposely superimposed upon the picture whenever a picture element is replenished (they are not a defect in the system). When the subject is not moving, as shown in Fig. 5c, the points are replenished more or less at random since the peaks of noise exceed the low threshold. As soon as the subject moves, as shown in Fig. 5d, the changing elements of the picture take precedence and one can see that replenishment concentrates on the subject as one might expect. The background noise causes very few picture elements to be replenished.

By viewing the accumulation of markers representing replenished points, we observe that the picture is replenished very quickly around the moving subject and that it takes a much longer time to randomly replenish the other elements. Left to chance, there are a few parts that are not replenished for a long time. We demonstrated that it is more efficient to gradually update all picture elements according to a predetermined pattern, rather than to lower the threshold to permit noise to cause replenishment.

VII. SUMMARY

We have presented a method of encoding television signals taking advantage of frame-to-frame redundancy. Only the address and amplitude of elements that have changed significantly between successive frames are transmitted. Varying the significant change threshold value helps to match the average rate of replenishment to the capacity of the transmission channel. A buffer then smooths the data flow for transmission.

Conditional replenishment lends itself to many ways of efficiently encoding pictures for transmission. We think that the buffer capacity and transmission requirement can be considerably reduced over that demonstrated here.

VIII. ACKNOWLEDGMENTS

The writer is indebted to L. H. Enloe for his leadership and support of this project, to C. C. Cutler for his encouragement and many helpful discussions and suggestions, and to W. T. Wintringham for his inspiration and interest in this work.

I would also like to express my very sincere thanks to J. C. Candy for his help and advice, to J. B. Pestrichelli for assistance in setting up and testing the experimental equipment, to R. C. Brainard and E. S. Bednar for designing the frame memory, to R. L. Eilenberger for providing the television camera equipment, and to Sal Farina for his assistance.

REFERENCES

1. Kell, R. D., British Patent No. 341811, 1929.
2. Seyler, A. J., "Frame Run Coding of Television Signals: A New Method for Bandwidth Reduction." Research Laboratory Report No. 5064, Commonwealth of Australia, Postmaster General's Department, September 1959.
3. Damen, T. C., unpublished work.
4. Seyler, A. J., "Probability Distributions of Television Frame Differences," Proc. IREE (Australia), *26* (November 1965), p. 355.
5. Seyler, A. J., "The Coding of Visual Signals to Reduce Channel-Capacity Requirements," The Institution of Electrical Engineers Monograph No. 535E, July 1962.
6. Kretzmer, E. R., "Statistics of Television Signals," B.S.T.J., *31,* No. 4 (July 1952), pp. 751–763.