THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 1, January, 1972
Printed in U.S.A.

Buffer and Channel Sharing by Several Interframe *Picturephone*® Coders

By B. G. HASKELL

(Manuscript received August 5, 1971)

Simulations are described which test the feasibility of several conditional replenishment type Picturephone *coders sharing the same transmission channel. Channel sharing takes advantage of the fact that many different users are rarely in rapid motion at the same time. Thus, when the data from several sources is averaged together prior to transmission, it is much more uniform than that from a single source. Since the peaks in the averaged data are smaller than with a single source, less buffering is required, and the channel rate required for transmission is much closer to the average data rate generated. Results indicate that by combining twelve sources prior to transmission, a 2:1 reduction in the bit rate of the system described by Candy, et al.,*[1] *can be obtained. This means that twelve* Picturephone *sources could share a 12 megabits per second (one way) transmission facility.*

I. INTRODUCTION AND SUMMARY

In Ref. 1 a frame-to-frame coding system for *Picturephone* signals is described which operates at 2 Mb/s and uses a number of techniques to reduce the bit rate required for transmission. For the most part, only picture elements that change significantly are transmitted in each frame period.[2] These elements are addressed in clusters and transmitted as frame differences; and since information is generated at an irregular rate, the data is buffered prior to transmission. When the buffer starts to fill, indicating active motion, only every other changed picture element is transmitted,[3,4] the unsampled elements being replaced by the average of their neighbors. When the buffer fills completely, replenishment is stopped for one frame period allowing the buffer to empty before resuming transmission.

Two-to-one subsampling in the changing parts of the picture has been shown to be subjectively invisible during rapid motion.[1,4] Frame repeating on the other hand, is quite visible, but since it is rarely used (only during periods of camera panning or violent motion) it is not too objectionable.

The rate at which data is generated by the system is very irregular due to random spacing of changed elements within the field and due to variations in movement on the part of the subject. It is practical to provide enough buffering to smooth the data over a field time. However, it is not practical to smooth the data between peaks in human movement because of the expense involved and because of the transmission delay which such a large buffer would introduce.[1,5] Thus, the channel capacity which must be provided to accommodate active movement is much larger than the long-term average data rate generated by the system, and unless some sort of stuffing information is generated,[1] such as forced replenishment, the transmission channel is idle much of the time.

An obvious method to utilize the channel more efficiently is to transmit the data from several interframe coders over a shared channel, allocating more capacity and buffer space to those sources which happen to require it at a given moment and less to those which are producing relatively little data. Since separate *Picturephone* conversations can be assumed to be independent, the combined data on a per source basis should be much less peaked than that from a single source, or in statistical terms, if the data from N_S independent, identically distributed sources is averaged, then the variance is reduced by a factor of N_S (the mean is unchanged). This will not only reduce the required channel capacity, but it will also reduce the required buffering.

A number of schemes have been proposed.[6–8] Most of them are variations of the system shown in Fig. 1. Unbuffered data from the various sources is first stored in individual prebuffers, then transferred to the principal buffer via the multiplexing switch, and finally transmitted over a single high-capacity data channel to the receiver where the inverse operation takes place. The switch may rotate either sequentially or nonsequentially and either at a constant rate or at a variable rate. In any event, some source labeling information must be sent so that at the receiver a given block of data is identified with the correct source.

If the individual sources are far removed from the high-capacity channel and its associated multiplexing switch, then it may be desirable to buffer the data as in Ref. 1 prior to transmission to the prebuffers of Fig. 1. In this case each prebuffer should be preceded by a device which removes the stuffing information from the incoming constant-rate data stream before it is fed to the prebuffer.

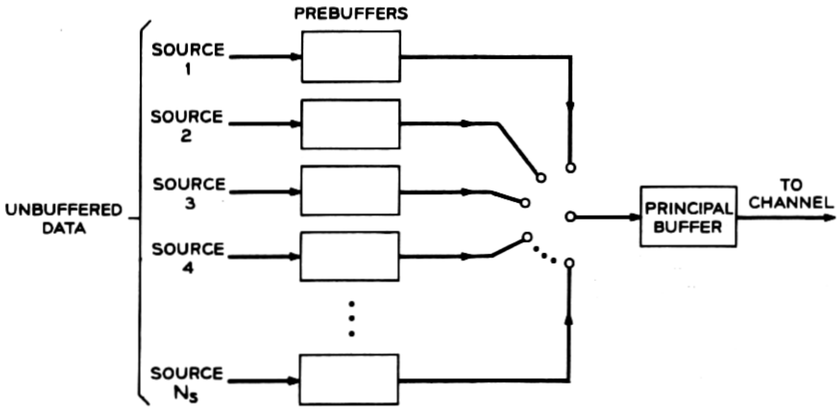

Fig. 1—Channel and buffer sharing system for sources which generate data at an irregular rate.

In order to simulate the system of Fig. 1, one hour of unbuffered data from a single source was stored on a digital disk. The data included all types of activity in the proportion in which they could be expected to occur in actual *Picturephone* conversations. The simultaneous occurrence of N_S independent conversations was then simulated by starting each conversation at a different point on the digital disk. This method of simulation avoids the necessity of constructing a reliable statistical model for the data; however, it does require a large amount of random access storage.

Controlling the coding mode of a single source according to the number of bits in a buffer as is done in Ref. 1 reduces the peak data rate considerably, but it does not result in the minimum long-term average data rate, a more important parameter where channel sharing is concerned. Furthermore, if the prebuffers were the only ones available, they would probably not be large enough to perform this task effectively. Thus, in the simulations the number of significant changes per field was used to control the transitions between full sampling in the changed area, sampling every other changed picture element (2:1 subsampling), and sampling every fourth changed picture element (4:1 subsampling).

Sections II and III describe the method of acquisition of the data and various statistical results obtained from it. Section IV describes simulations of single interframe coders using different techniques to control the coding mode according to the amount of activity in the scene. Required buffer size versus channel rate curves are obtained under the assumption that the buffer is at least large enough to smooth the data over a field period. Simulation results for shared-buffer/shared-channel systems are described in Section V.

The simulations indicate that there is indeed considerable advantage in several interframe coders sharing a buffer and a channel. Using the techniques of Ref. 1, a channel rate of 2.0 Mb/s (one way) is required to transmit the data from a single *Picturephone* source. However, with a moderate amount of buffering (about the same as that of a single frame-to-frame coder[1]), combining five sources prior to transmission reduces the required one-way channel rate to 1.2 Mb/s on a per source basis, i.e., five sources over a 6-Mb/s transmission facility. With 12 sources per shared channel a transmission rate of 1.0 Mb/s per source can be achieved with relatively little buffering. Significantly more savings in channel rate is obtained only by sharing a very much larger number of sources. As the number of sources per shared channel increases, the bit rate per source approaches 0.6 Mb/s asymptotically.

II. ACQUISITION OF DATA

Frame-to-frame television picture coders are, at present, only in the experimental stage. Thus, it is not possible to test in real time the feasibility of sharing transmission channels. Computer simulations can be carried out, however, by storing a long sample of unbuffered data from a single source on a digital disk and simulating the simultaneous occurrence of N_S *Picturephone* conversations by starting each conversation at a different point on the digital disk. As long as the starting points are not close together the conversations simulated in this manner will be independent.

The apparatus used to generate the data was essentially the same as in Ref. 1. The picture was scanned with 271 lines at 30 frames per second (two interlaced fields per frame) producing a nominally 1-MHz signal. This was sampled at about 2 MHz and coded as 8-bit PCM ($0 \cdots 255$) so that all processing could be done digitally. A picture element was deemed to have changed significantly if the magnitude of its frame-to-frame difference was greater than four out of a possible 255. However, two exceptions to this criterion were made: (*i*) if a significant change was preceded and followed 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 elements as significant changes.

For each field, two numbers were recorded: the number of significant changes and the number of clusters requiring addressing. These were then stored on a digital disk via a Honeywell DDP-224 digital computer. Other data such as line sync words are generated, at least in these simulations, at a constant rate and, therefore, do not require storage. The bit rate for each field can then be computed by assigning four bits to each significant change, 12 bits to each cluster requiring addressing, and then adding the bits which occur at a constant rate.

Storing only two numbers per field makes it impossible to simulate phenomena that occur because of the irregular distribution of data within fields. It also requires that in all simulations buffers must be large enough to smooth these irregularities over a field time. More will be said about this point in later sections.

It was intended that these data should represent as much as possible those that would result in practice. To accomplish this goal, product-trial video tapes of actual *Picturephone* conversations were obtained and used as a guide as to the amount of motion which could be expected. A subject was placed in front of a TV camera whose output was being coded by the interframe coder, and, at the same time, the product-trial video tape was played through a monitor which was visible to the subject. The subject was then instructed to mimic as much as possible the motion displayed on the monitor.

Violent activity such as camera panning or flipping the mirror to display printed material was simulated by waving a large card in front of the camera. Such behavior occasionally caused changes in nearly 90 percent of the picture area. Simultaneous movement by more than one person in the scene (this rarely occurred) was simulated by the subject moving closer to the camera to more nearly approximate the area in motion.

Data were recorded for one hour, using two subjects; the scene for both the simulation and the product trial was a standard office with average lighting. The dress of the subjects was somewhat more colorful than that of the individuals in the video tape, however. But since this will tend to make the results conservative, if anything, it was not considered to be a drawback.

III. STATISTICAL CHARACTERISTICS OF THE DATA

The long-term average characteristics of movement by the typical single *Picturephone* user are shown in Fig. 2 where the probability of the number of changes per field exceeding n versus n is plotted, i.e., $1 -$ probability distribution function of number changes per field.

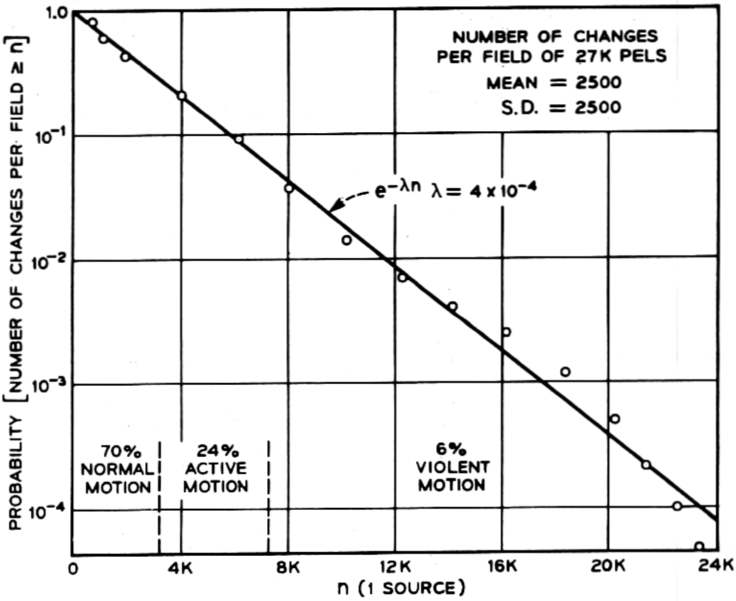

Fig. 2—Probability that the number of changes per field exceeds the threshold value n. Points lie very close to the exponential $e^{-\lambda n}$ with $\lambda = 4 \times 10^{-4}$.

The curve is very close to the exponential $e^{-\lambda n}$ (for $\lambda = 4 \times 10^{-4}$) over much of its range. The exponential shape agrees with results obtained by Seyler[9] for certain types of commercial television material, but, as reported by Pease,[10] the curve tends to be much lower (i.e., a smaller proportion of the picture changing on the average).

Inactive conversation (less than 11 percent of the picture changing) comprises about 70 percent of the data. Active motion (between 11 percent and 25 percent of the picture changing) comprises about 24 percent of the data, and violent motion (more than 25 percent of the picture changing) comprises the remaining six percent of the data. Thus, an interframe coder which was designed to accommodate active motion would fully use the required channel capacity only about 6 percent of the time. For the remainder of the time, stuffing information would have to be transmitted to maintain synchronization within the system.

Figures 3 to 6 show the curves corresponding to Fig. 2 when data streams from more than one source are averaged together prior to transmission. Simultaneous occurrence of several *Picturephone* con-

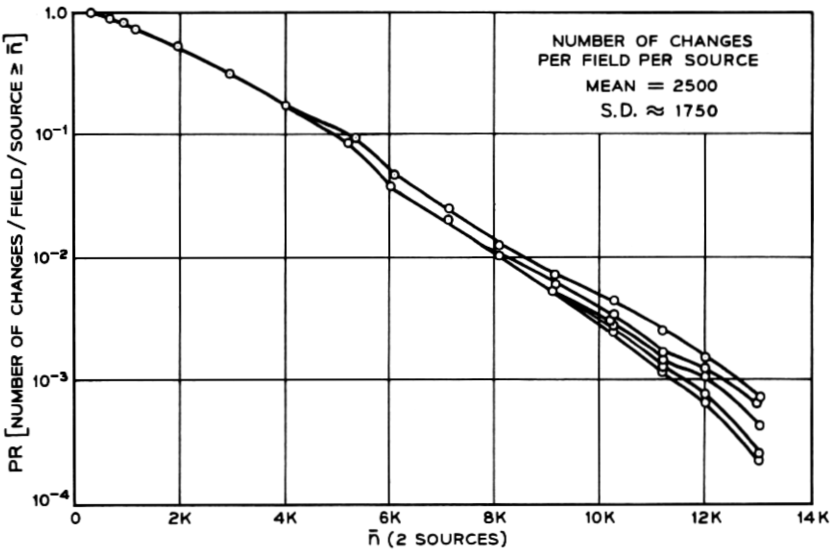

Fig. 3—Probability that the number of changes per field per source exceeds $\bar{n}$ when data from *two sources* is combined. Multiplicity of curves results from different starting positions in the data.

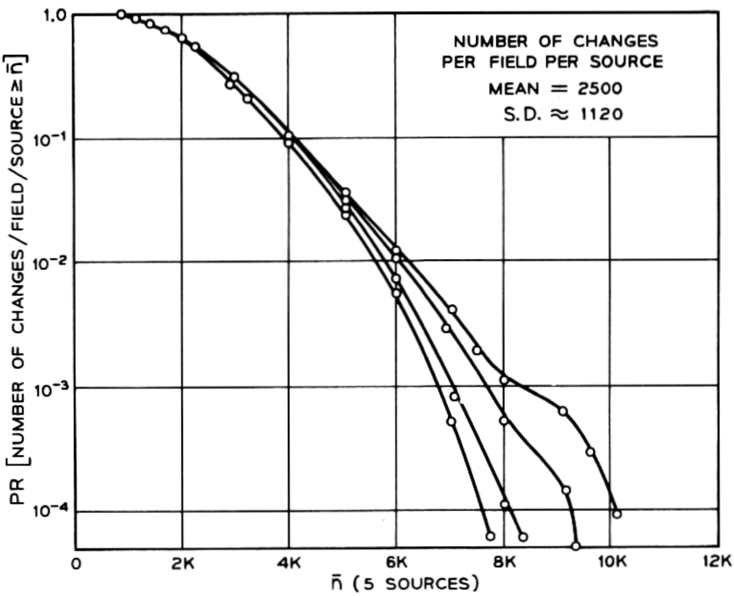

Fig. 4—Probability that the number of changes per field per source exceeds $\bar{n}$ when data from *five sources* is combined.

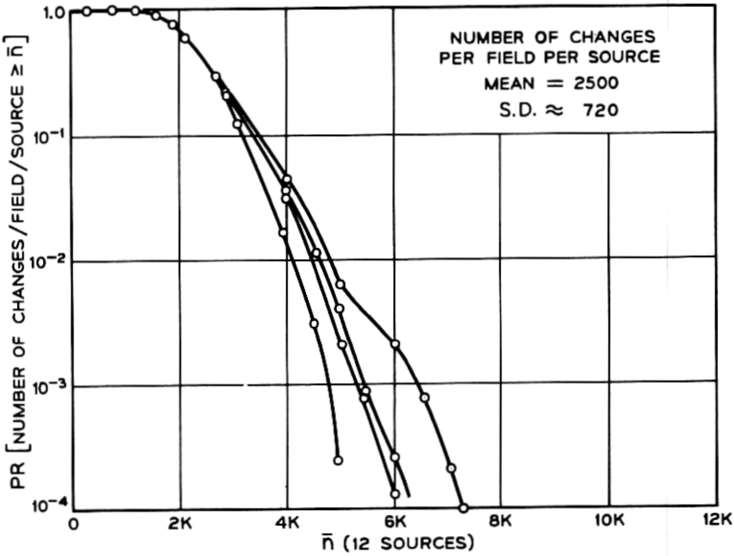

Fig. 5—Probability that the number of changes per field per source exceeds $\bar{n}$ when data from *twelve sources* is combined.

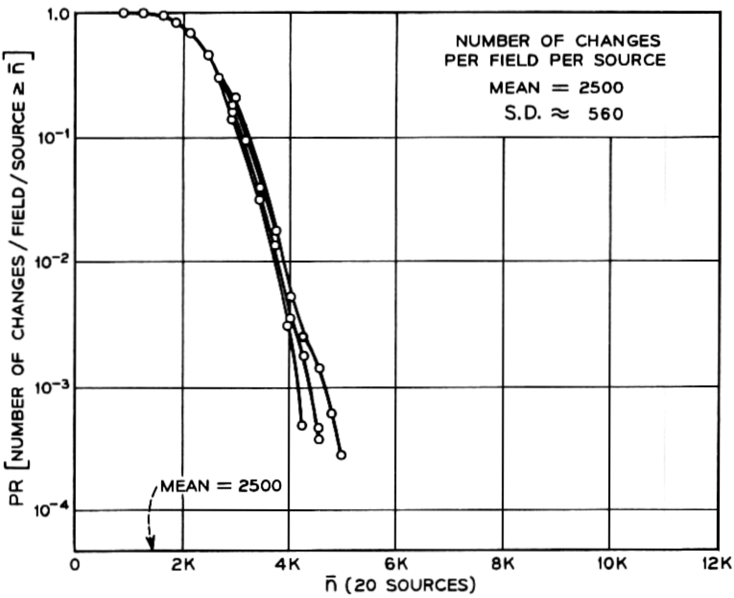

Fig. 6—Probability that the number of changes per field per source exceeds $\bar{n}$ when data from *twenty sources* is combined.

versations was simulated by starting each conversation at a different point on the digital disk. The multiplicity of curves in each figure results from various choices for these starting positions. The number of changes per field per message source is obtained by adding the number of changes per field contributed by each message source and dividing by N_S, the number of sources. The probability of this quantity exceeding $\bar{n}$ versus $\bar{n}$ is plotted in Figs. 3 to 6 for $N_S = 2, 5, 12$ and 20 respectively. As the number of sources increases, the variation of the data from the mean decreases indicating that the data are becoming more and more uniform.

When the starting positions of the separate conversations on the digital disk are changed, the results vary, especially in the tails of the distribution. This occurs because the tails are determined completely by the alignment of movement peaks in the data from each of the several sources, and when the relative positions of these peaks are changed the alignment will also change.

Figures 3 to 6, which were obtained experimentally, are very close to those obtained analytically using the exponential $\lambda e^{-\lambda n}$ as the probability density function of n for the single-source case. For more than 15 sources ($N_S \geqq 15$) the central limit theorem begins to come into effect and $\bar{n}$ is very nearly Gaussian with mean λ^{-1}, i.e., 2500 changes, and variance $(N_S\lambda^2)^{-1}$.

The large spread and consequent large possible deviation from the mean of the data from a single source is only one reason why buffering is difficult. The other reason is the high field-to-field correlation in the data due to the fact that motion in successive fields is very much the same.[1] Figure 7 shows the normalized autocorrelation function of the number of changes per field (this function is independent of N_S). According to the curve there is still a 50 percent correlation between fields which are as much as one second apart.

For a single source, transmission channels become saturated and buffers become filled only during periods of rapid motion. Thus, one is less interested in the long-time average correlation in the data as shown in Figure 7 and is more interested in how long the large peaks in the data rate are going to last. Figure 8 indicates the average durations of data peaks for a single source. If n, the number of changes per field, rises above some threshold value T during field i and does not go below T again until field j, then the duration of the peak is $(j - i)$ fields. From Fig. 8 it is seen that although periods of high data rates are quite rare, when one does occur chances are that it will last for at least several fields. In fact, for human conversations large peaks in the data can last for a

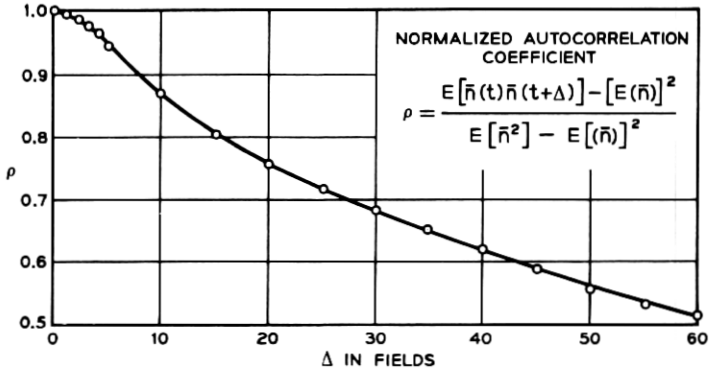

Fig. 7—Normalized autocorrelation coefficient of the number of changes per field. There is still a 50-percent correlation between fields which are 1 second apart.

few seconds, a period much too long to allow complete smoothing by a buffer because of the transmission delay that would be introduced.[1,5]

Combining data from several sources prior to transmission reduces the impact of these peaks considerably. Figures 9 to 12 show the average durations of peaks in $\bar{n}$, the number of changes per field per source,

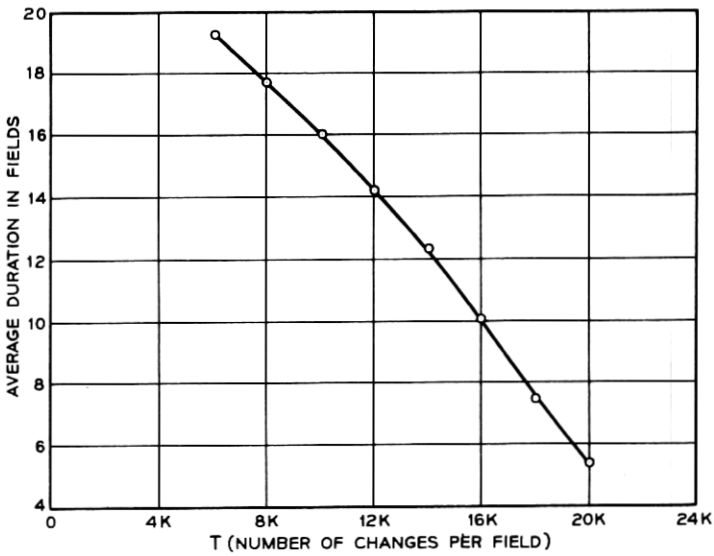

Fig. 8—Average duration of peaks in the number of changes per field. A peak starts when n rises above T and ends when n falls back below T.

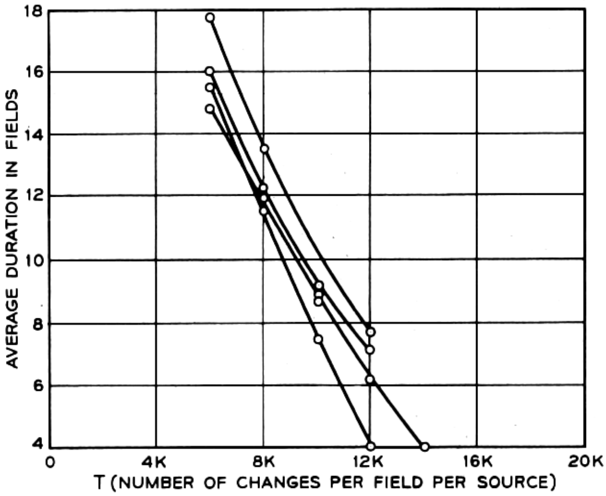

Fig. 9—Average duration of peaks in the number of changes per field per source when data from *two sources* is combined.

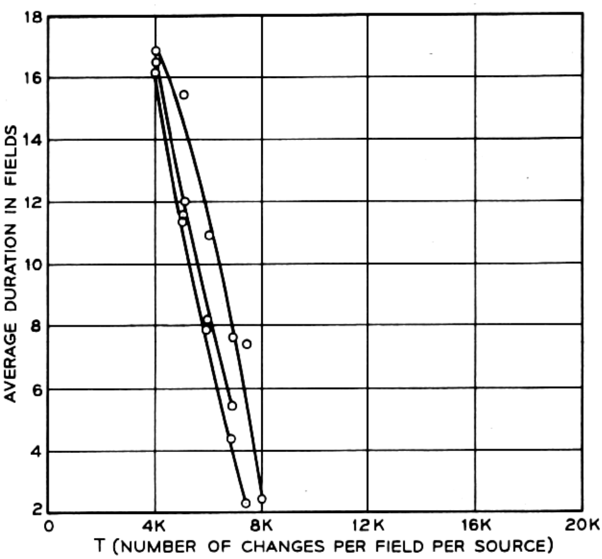

Fig. 10—Average duration of peaks in the number of changes per field per source when data from *five sources* is combined.

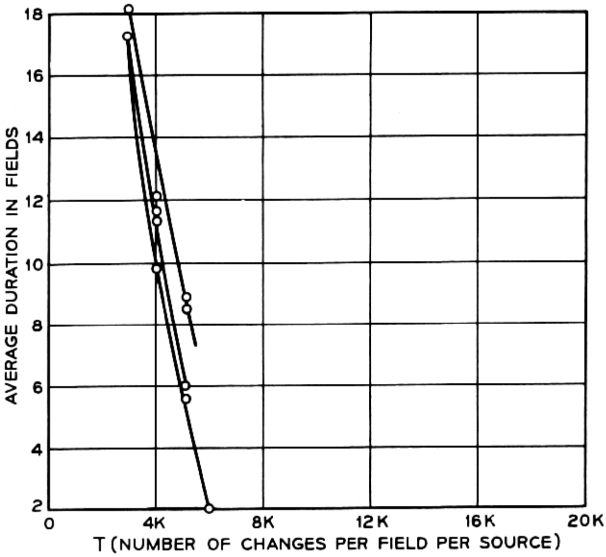

Fig. 11—Average duration of peaks in the number of changes per field per source when data from *twelve sources* is combined.

when $N_S = 2, 5, 12$, and 20 sources. As before, the multiplicity of curves represents changes in the digital disk starting positions. As N_S, the number of sources per shared channel increases, the magnitudes of large data peaks decrease due to the probable misalignment of peaks from the several data sources.

IV. SIMULATION OF SINGLE INTERFRAME CODERS

A number of simulations were carried out of frame-to-frame coders operating by themselves without channel or buffer sharing. These simulations differ from those of Ref. 5 in that the data used represent all types of motion in the relative proportions that could be expected in *Picturephone* transmission, whereas in Ref. 5 each segment of data represented a particular level of activity. Also, two stages of data-rate reduction were employed in these simulations, whereas in Ref. 5 only 2:1 subsampling was used.

Several features of real-time interframe coders could not be simulated using the data described in Sections II and III. *First*, since the number of changes and the number of clusters were recorded on a per-field basis, behavior within a field could not be simulated. Thus, it is always assumed that buffers are large enough to smooth the data over a field period. For a bad case suppose that all of the data for a busy field occurred in the bottom half of the picture. Then for a 2 Mb/s transmission rate ($\approx$33,000 bits per field), a buffer capable of storing 16,500 bits would be required to smooth over a field period. Even assuming a large buffer, however, small queue lengths cannot be represented accurately in systems where extra stuffing information is generated during periods when the buffer is empty.

Second, raising the threshold, which determines a significant frame difference during periods of violent motion, as was done in Ref. 1 requires feedback to the interframe coder; this was not employed during the collection of these data. However, the absence of this feature should not affect the results very much since during periods of high activity it is not a particularly effective technique for data reduction (see Fig. 16 in Ref. 1). Furthermore, use of a constant threshold in the simulations would tend to make the results conservative, if anything.

A third effect which could not be simulated adequately was switching between the various stages of data-rate reduction. Switching from 2:1 subsampling to full sampling, for example, requires that extra information be transmitted to update those picture elements which

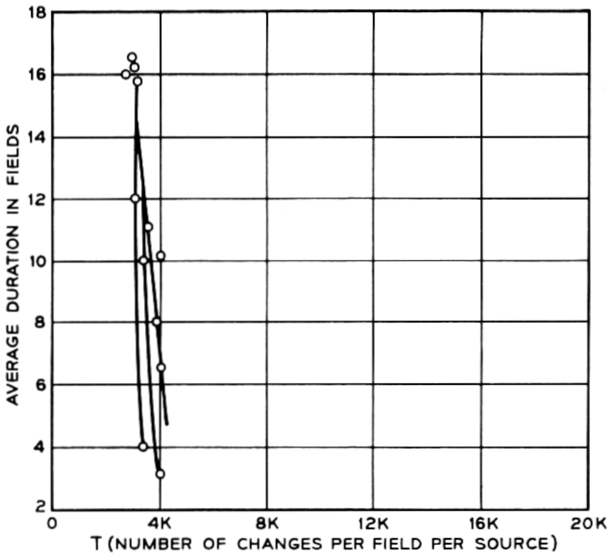

Fig. 12—Average duration of peaks in the number of changes per field per source when data from *twenty sources* is combined.

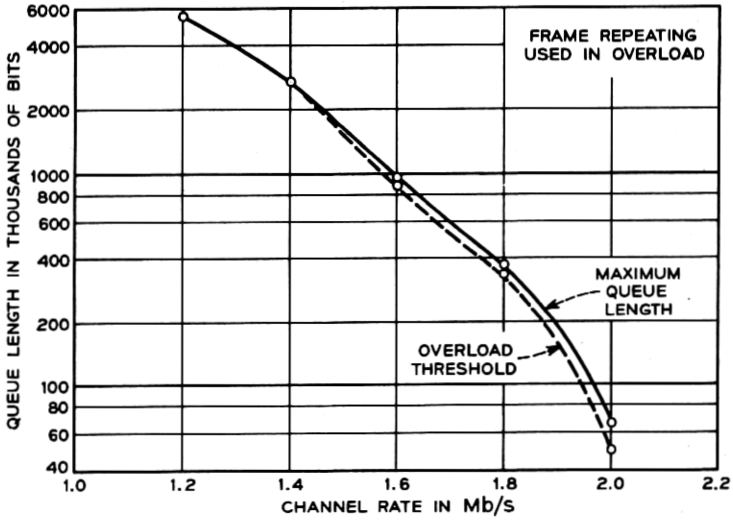

Fig. 13—Buffer requirements versus channel rate for a single source and 0.11-percent probability of overload. Overload detected via queue length at end of previous field. Frame repeating used in overload.

were previously being derived through interpolation. This switch need not take place in one frame-time, however. In a bad case where 60 to 70 percent of the picture changes during subsampling and the interpolation is completely unsuccessful, only about 600 extra picture elements per field need be transmitted to produce a full resolution picture in 0.5 seconds. When switching from 4:1 subsampling to 2:1 subsampling, only 300 extra picture elements per field need be transmitted. In view of these facts, it is felt that the switching itself can be carried out in practice in such a way as to have little effect on simulation results.

Initially, schemes similar to Ref. 1 were simulated where the number of bits in the buffer, i.e., queue length, was used to determine the usage of 2:1 subsampling or frame repeating. Twelve bits were assigned to each cluster that required addressing and four bits to each frame difference that was transmitted. When the buffer was more than 20 percent full, 2:1 subsampling in the changing area was switched in. When the queue length exceeded some higher threshold value, a whole frame was repeated thus requiring no transmission for the duration of that frame.

In the first simulation, the queue length at the end of the previous field was used to control subsampling and repeating in the present field. Using the channel rate and buffer size of Ref. 1, namely 2 Mb/s and 67,000 bits, overflow occurred during 0.11 percent of the frames, i.e., about 1 second in a 15-minute conversation. For other channel rates, Fig. 13 shows the buffer sizes and threshold values required to maintain a 0.11 percent probability of visible degradation due to buffer filling.

If instead, one controls the subsampling and repeating by the queue length requirements in the present field, the results are changed very little, as can be seen by comparing Figs. 13 and 14. This is due to the high correlation in the amount of movement in successive fields.

The effect on the picture of frame repeating during periods of violent movement is quite noticeable,[1] and it is speculated that this degradation might be made less objectionable if 4:1 subsampling were used instead of frame repeating.[1,4] A simulation was carried out to see what the effect on buffer requirements would be of switching to a 4:1 subsampling mode when the queue length at the end of the previous field was above some threshold. Figure 15 shows the buffer size and threshold versus channel rate required to maintain a 0.11 percent probability of buffer overload. The buffer sizes are significantly larger than those of Figure 13.

With 4:1 subsampling, however, a higher usage of this mode may be

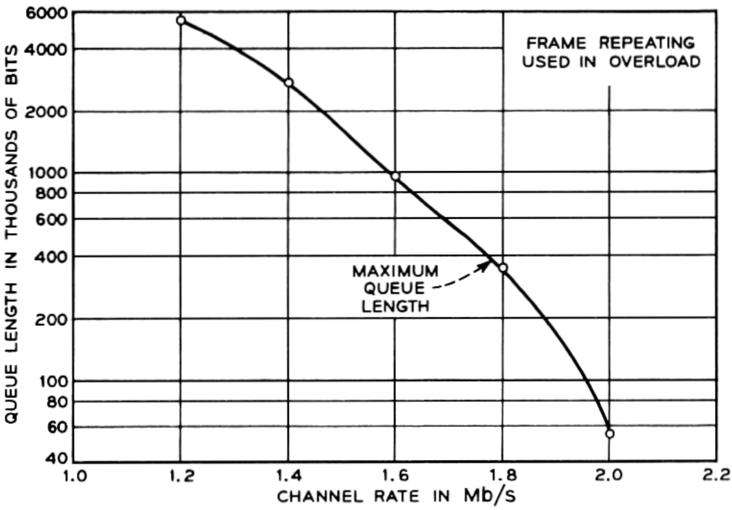

Fig. 14—Buffer requirements versus channel rate for a single source and 0.11-percent probability of overload. Overload detected via queue length in present field. (Curve is very close to that of Fig. 13.)

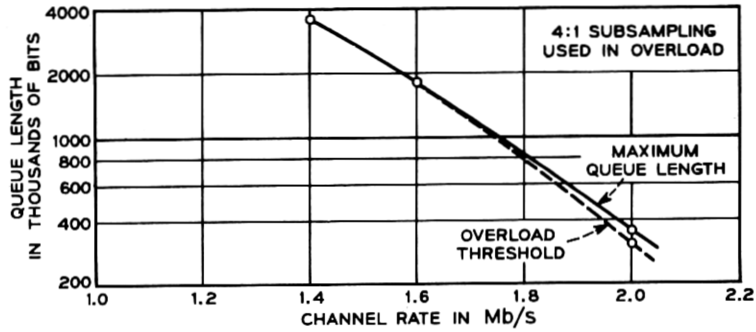

Fig. 15—Buffer requirements versus channel rate for a single source and 0.11-percent probability of overload. 4:1 subsampling used in overload. (Curve is higher than that of Fig. 13.)

tolerated since the degradation is expected to be less than with frame repeating. Exactly how much can be tolerated is not known, but if the frequency of usage is allowed to rise to 0.23 percent, then the curve of Fig. 16 is obtained (which matches very closely the buffer requirements of Figure 13 where frame repeating is used).

In the above schemes a drastic data-rate reduction is made if buffer overflow threatens. This has the effect of emptying the buffer so that

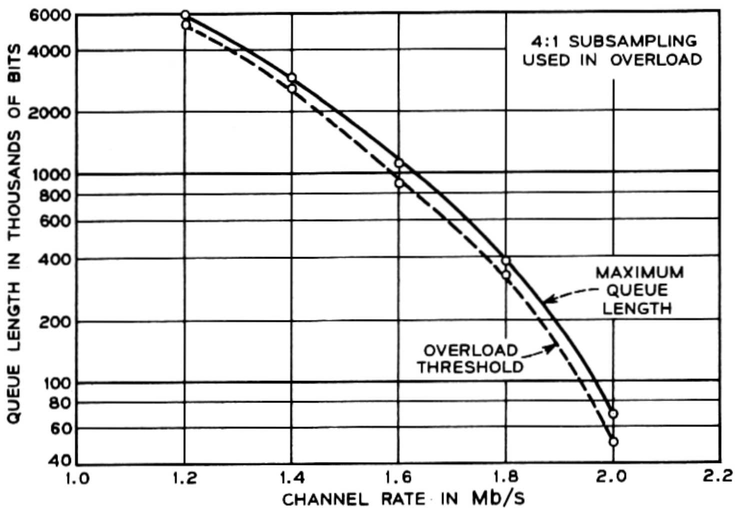

Fig. 16—Buffer requirements versus channel rate for a single source and 0.23-percent probability of overload. 4:1 subsampling used in overload. (Curve is close to that of Fig. 13.)

the next few fields can be sent with good resolution even though rapid motion is occurring. Thus, during violent movement only 1 in 4 or 1 in 8 frames is degraded. The visible degradation is annoying, however, during the whole period of violent motion. Thus, although only 0.11 percent of the frames are degraded, the visible impairment may be objectionable during all violent motion (0.88 percent of the time) since the intervening unimpaired frames do not reduce the visibility of the phenomenon.

When channel sharing is contemplated, system performance depends on the long-term average data rate. Thus, 2:1 subsampling should be employed not only when the buffer starts to fill but during all periods of active movement; similarly, 4:1 subsampling should be used during all periods of very violent movement. To simulate techniques of this type the number of significant changes in the previous field was used to control the coding in the present field. If the number of changes

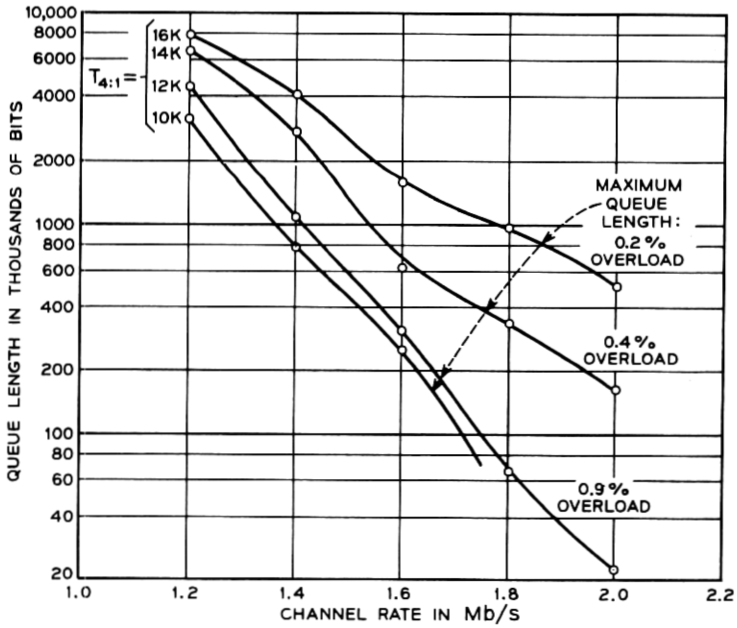

Fig. 17—Buffer requirements versus channel rate for a single source where subsampling is controlled according to the number of changes in the previous field. If the number of changes exceeds $T_{2:1}$ then 2:1 subsampling is used; if the number of changes exceeds $T_{4:1}$ then 4:1 subsampling is used. Variations in $T_{2:1}$ do not affect the results if $T_{2:1} \ll T_{4:1}$. Also shown for each value of $T_{4:1}$ is the percentage of the time in overload, i.e., when 4:1 subsampling is used.

exceeds the value $T_{2:1}$ then 2:1 subsampling is used; if the number of changes exceeds $T_{4:1}$, i.e., during violent motion, then 4:1 subsampling is used. In practice, only frame differences corresponding to sampled picture elements should be examined since the interpolation itself introduces frame differences independent of movement. Buffer size versus channel rate curves for various values of $T_{4:1}$ are given in Fig. 17 along with the percentage of time that overload, i.e., 4:1 subsampling, is used. Variations in $T_{2:1}$ do not seem to affect these results as long as $T_{2:1}$ is much smaller than $T_{4:1}$.

The value of $T_{4:1}$ for which the buffer requirements most closely approximated those of Figures 13 and 16 where queue length control is used is between 13,000 and 14,000 changes per field. This corresponds to very violent motion (about half the picture changing) and occurs only about 0.5 percent of the time (see Fig. 2). Statistics of the bit stream with $T_{4:1} = 13{,}400$ changes per field and $T_{2:1} = 3000$ changes per field are given in Figures 18 and 19. From these figures it is again apparent that large peaks in the bit rate occur very rarely, but when they do occur they can last for an appreciable period of time.

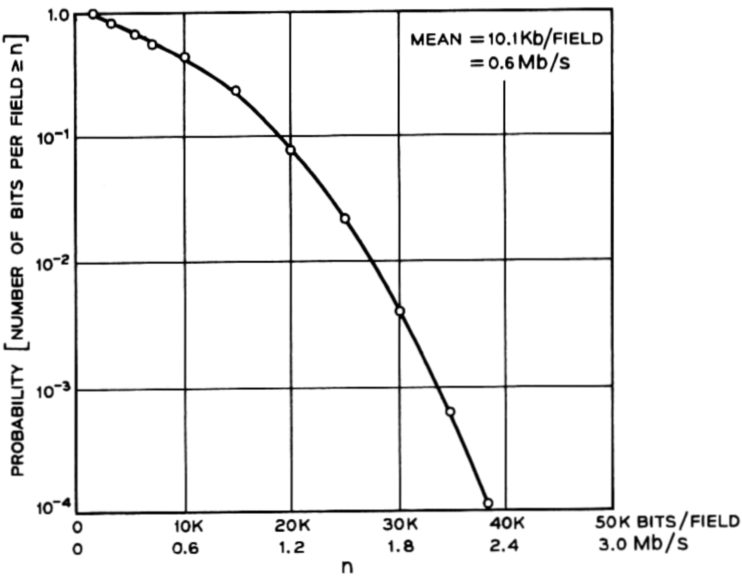

Fig. 18—Statistical characteristics of the single-source bit stream with subsampling controlled by the number of changes in the previous field. $T_{2:1} = 3000$; $T_{4:1} = 13{,}400$. Probability that the number of bits per field exceeds n is plotted versus n. Also shown is n in Mb/s.

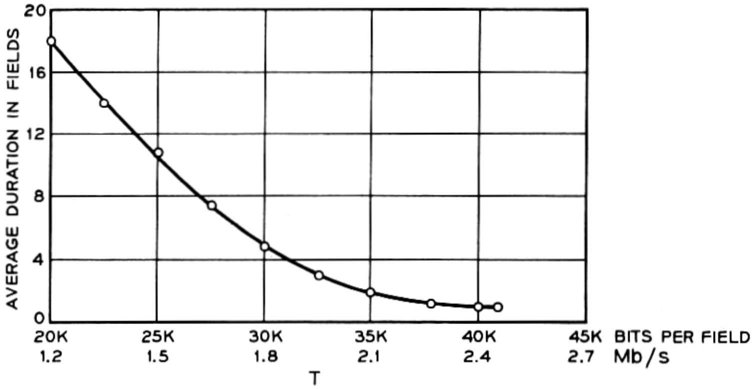

Fig. 19—Average duration of peaks in the bit rate above T for a single source with subsampling controlled by the number of changes in the previous field. $T_{2:1} = 3000$; $T_{4:1} = 13{,}400$.

Smaller values of $T_{4:1}$ require smaller buffers; however, the picture degradation will be more objectionable. Just how small $T_{4:1}$ can be made depends on the actual 4:1 subsampling scheme used and on the subjective requirements of the viewer. However, as we shall see in the next section, when combining sources prior to transmission, variations in $T_{4:1}$ do not seem to affect the results significantly. $T_{2:1}$ then becomes much more important in the performance of the system.

V. SIMULATION OF SHARED BUFFER-SHARED CHANNEL SYSTEMS

The system that will be considered in this section is shown in Fig. 1. Unbuffered data from each interframe coder or buffered data from which the stuffing information has been removed enter the prebuffers at the left to await transfer to the principal buffer via the time division multiplexer. The prebuffers should be capable of storing one or two lines of coded information if excessive transmission of source identification bits is to be avoided. The multiplexing switch can either rotate at a constant rate or stop at the output of each prebuffer only long enough to transfer data. A variable rotation rate will, in general, require less buffering and alleviate certain synchronizing problems. Similarly, the switch can access each prebuffer either sequentially or nonsequentially according to some control strategy which gives higher priority to more active sources. Nonsequential access will, in general, require smaller buffers; however, some means must then be provided for preventing one source from hogging the channel. Other configurations are also possible; e.g., with a nonuniform switch rotation rate the principal buffer might be removed by making the prebuffers larger.

On rare occasions the combined data from all sources will be too much for the system to transmit, and some data will have to be deleted. If the coder itself can be told that data are being deleted, then it can be done without introducing too much picture degradation in the same way as in the single-source coder.[1] However, if rapid feedback to the coder is not feasible, then the responsibility for minimizing picture degradation in the event of buffer overflow falls entirely on the various receiver systems. This could happen in the situation depicted in Fig. 20, where the signal has successfully passed through N shared channel transmission systems and upon reaching the Nth switching station it finds that there is a momentary shortage of transmission capacity requiring the deletion of data according to some predetermined strategy. For example, half the transmitted samples along a line might be deleted, thus requiring interpolation at the receiver, or an entire line of data

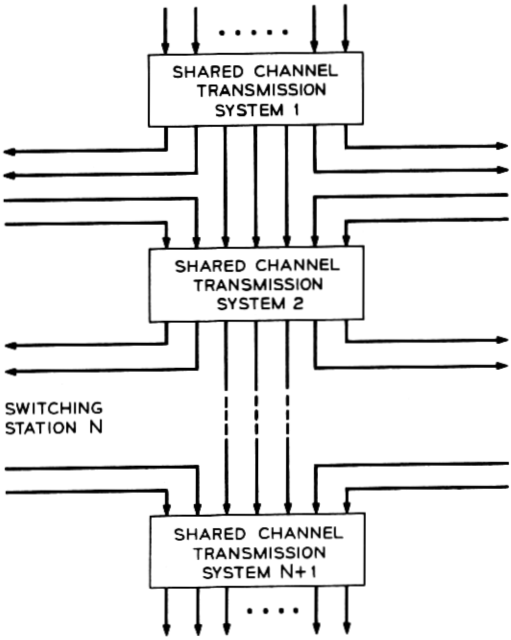

Fig. 20—Cascaded shared channel transmission systems where feedback to the original coder is impractical.

might be dropped at the Nth switching station and replaced by the average of the line above and the line below at the receiver.[11]

In the simulations which were carried out to study the channel sharing system of Fig. 1, 2:1 or 4:1 subsampling was assumed to have already taken place for each incoming field in which the number of significant changes exceeded $T_{2:1} = 3000$ or $T_{4:1} = 13{,}400$, respectively. A value of three thousand changes per field represents moderate to active motion on the part of the subject and is probably near the minimum value which could be used to control 2:1 subsampling in practice, since subsampling during slow movement results in noticeable aliasing defects in the picture. A value of 13,400 changes per field corresponds to violent motion, where nearly 50 percent of the picture area is changing. The choice of $T_{4:1}$ should not be crucial to the picture quality as long as it is in the range of violent motion; thus, for comparison purposes it was chosen so that the resulting single-source buffer requirements approximated those of Fig. 16, where subsampling was controlled by the queue length in the buffer. The effects of varying $T_{2:1}$ and $T_{4:1}$ will be discussed later.

Since the data were recorded only at the field rate, the queue length within the transmitter buffer could be examined in the simulations only at certain times. Thus, as in the single-source case, small queue lengths are not represented accurately. Unlike the single-source case, however, a certain amount of in-field data smoothing takes place because of the data interleaving[5] mechanism at the multiplexer switch. If more than just a few sources are combined, it is highly unlikely that television lines from the various sources will come from the same area of the picture. Thus, while one line from one source may come from an active area near the bottom of the picture, other lines from other sources will come from quiet areas near the top of the picture. Furthermore, as the transmission rate per source is reduced, the amount of buffering per source required for in-field smoothing decreases. Thus, if the data from each source occurred in the bottom half of the field, and all the sources were more or less in vertical synchronism, then 16K of buffering per source would be required for in-field data smoothing at 2 Mb/s per source. However, only 8K of buffering would be required at 1 Mb/s per source, and much less is required if, as seems likely, the sources are not in vertical synchronism.

The maximum queue length per source is plotted versus the channel rate per source in Figures 21 to 25 for systems in which 1, 2, 5, 12 and 20 sources respectively share the same transmission channel. As before,

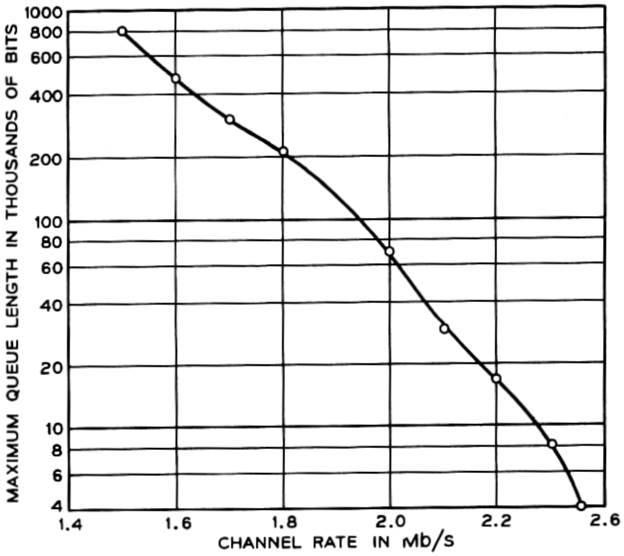

Fig. 21—Buffer requirements versus channel rate for a single source. (Same as Fig. 17 with $T_{4:1} = 13{,}400$, $T_{2:1} = 3000$.) 4:1 subsampling used when the number of changes in previous field exceeds $T_{4:1}$. Variations in $T_{2:1}$ have little effect.

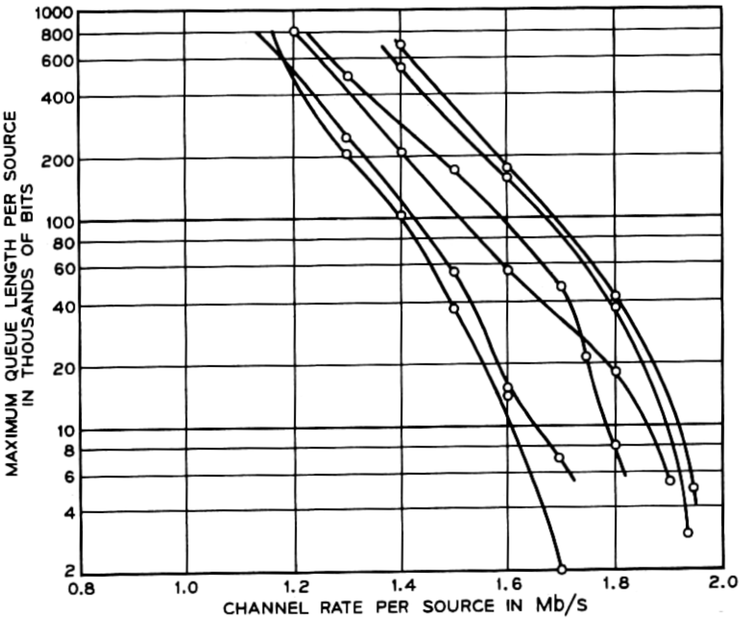

Fig. 22—Buffer requirements versus channel rate when data from *two sources* is combined prior to transmission. Multiplicity of curves results from different starting positions in the data. $T_{4:1} = 13{,}400$; $T_{2:1} = 3000$.

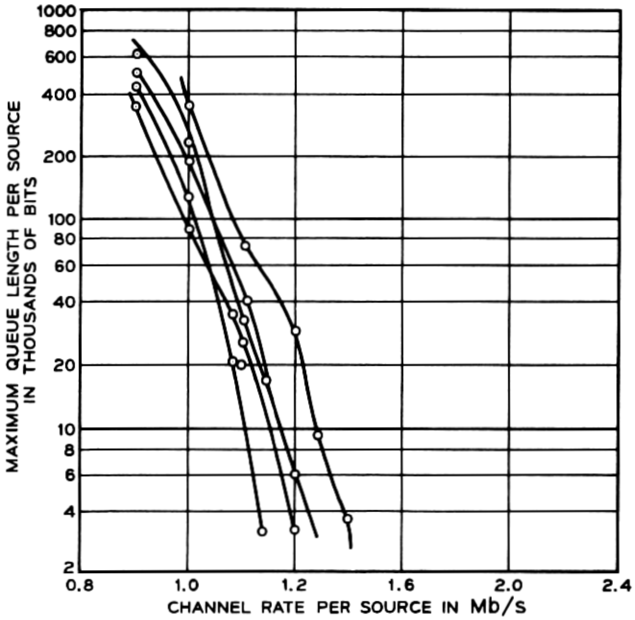

Fig. 23—Buffer requirements versus channel rate when data from *five sources* is combined prior to transmission.

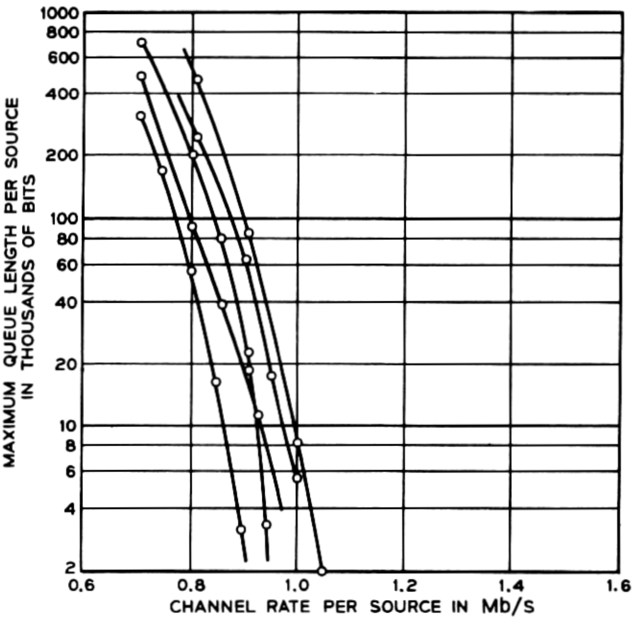

Fig. 24—Buffer requirements versus channel rate when data from *twelve sources* is combined prior to transmission.

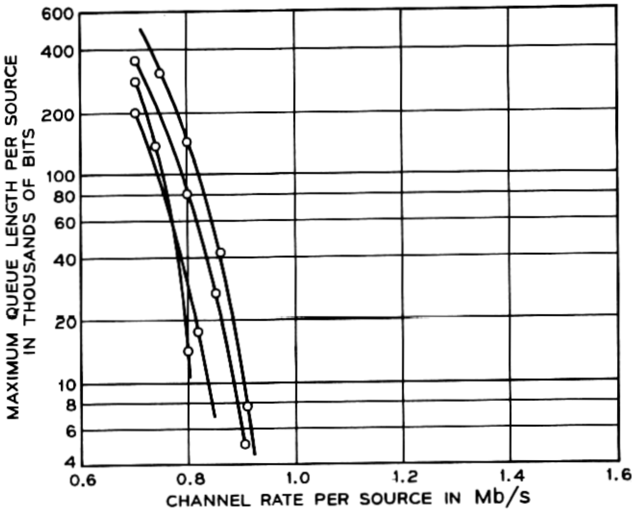

Fig. 25—Buffer requirements versus channel rate when data from *twenty sources* is combined prior to transmission.

the simultaneous occurrence of several conversations is simulated by starting each one at a different place in the data. Different sets of starting positions yield different curves. 2:1 subsampling was assumed when the number of changes per field exceeded $T_{2:1} = 3000$, and 4:1 subsampling was assumed when the number of changes per field exceeded $T_{4.1} = 13{,}400$. If buffer sizes of the order of 70,000 bits per source are not considered unreasonable, then the approximate channel rate per source required to transmit N_S sources, without buffer overflow, is given in Table I.

In every case where more than one source was simulated, changing the starting positions in the recorded data caused significant variation in the maximum queue length measurement. Results differing by a factor of ten were obtained in some cases. This occurs because the maximum queue length depends crucially on the lining up of data peaks from separate sources, and this is likely to be quite erratic as the starting position for each source is varied.

When more than four sources shared a channel, variations in $T_{4:1}$, the number of changes per field beyond which 4:1 subsampling was assumed, had little effect compared with the effect shown in Fig. 17 on a single source. Variations in $T_{2:1}$, however, had more and more of an effect as the number of sources per shared channel increased, whereas

TABLE I—CHANNEL RATE VERSUS NUMBER OF SOURCES (BUFFER SIZE = 70,000 BITS/SOURCE; NO OVERFLOW)

Number of Sources, N_S	Channel Rate per Source (Mb/s)
1	2.0
2	1.7
5	1.2
12	0.90
20	0.85

with a single source no effect was observed at all. This occurs because of the higher sensitivity to the average data rate as opposed to the peak data rate when the number of sources per shared channel is large.

The curves also tend to become much steeper as the number of sources is increased. This is to be expected since as the rate of the combined data becomes more and more uniform, buffering will become less and less effective in reducing the required channel rate. If the principal transmitter buffer is reduced in size to the point where it is barely large enough to smooth data peaks within a field (the simulations are invalid for buffer sizes smaller than this), then the approximate channel rate per source required to transmit N_S sources, without buffer overflow, is given in Table II. Comparison with Table I shows that for more than five sources, large buffers result in relatively little improvement in bit rate.

Exactly how large the principal buffer must be to smooth the data within a field is difficult to specify because of the many factors that were discussed earlier in this section. In the simulations, however, the data are already smoothed over a field period regardless of the buffer size. Thus, the results in Table II can be obtained by simply setting the size of the principal buffer equal to zero in the simulations.

TABLE II—CHANNEL RATE VERSUS NUMBER OF SOURCES (BUFFER LARGE ENOUGH ONLY FOR IN-FIELD SMOOTHING; NO OVERFLOW)

Number of Sources, N_S	Channel Rate per Source (Mb/s)
1	2.4
2	2.0
5	1.3
12	1.05
20	0.92

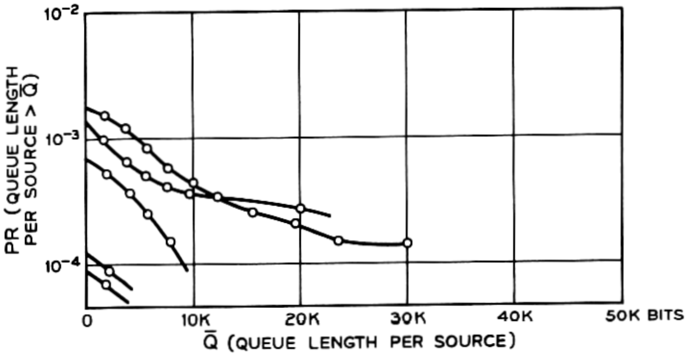

Fig. 26—Probability that the buffer queue length exceeds $\bar{Q}$ when data from *five sources* is combined. Channel rate = 1.25 Mb/s.

Up to now only the maximum queue length attained during the simulation has been considered. This maximum occurs only once, however. For nearly all the remaining time the queue length is much less than this. Thus, if a small probability of buffer overflow is allowed, the buffer size which would be required in practice could be somewhat less than is indicated in Figs. 21 to 25. Or, equivalently, for a given buffer size the required channel rate could be somewhat less.

Queue length probability distributions for 5, 12, and 20 sources and various channel rates are given in Figs. 26 to 28 for various starting

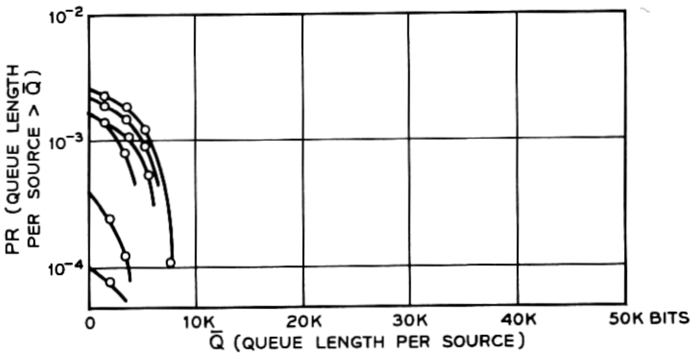

Fig. 27—Probability that the buffer queue length exceeds $\bar{Q}$ when data from *twelve sources* is combined. Channel rate = 1.0 Mb/s.

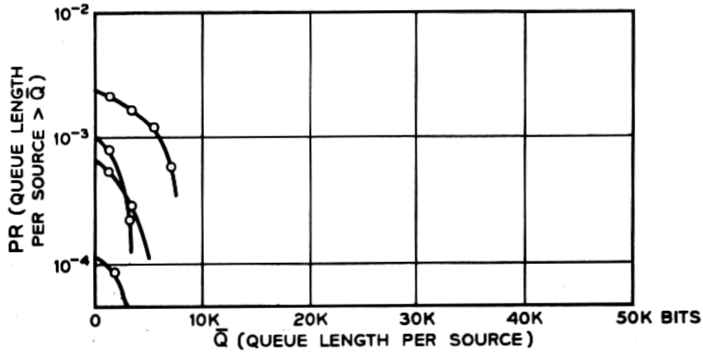

Fig. 28—Probability that the buffer queue length exceeds $\bar{Q}$ when data from *twenty sources* is combined. Channel rate = 0.9 Mb/s.

positions on the digital disk. These channel rates, which are also shown in Table III, were chosen so that with buffering of only a few thousand bits/source, i.e., enough for smoothing the data over a field period, overflow occurred less than about 0.1 percent of the time. Compared with the 0.5 percent probability of 4:1 subsampling, buffer overflow occurred one fifth as often.

Comparison of Tables II and III shows that as the number of sources per shared channel increases, the bit rate saving obtained by allowing a small probability of buffer overflow becomes less and less. This result further substantiates the fact that as the number of sources increases, the combined data become uniform enough so that buffering is of very little importance.

TABLE III—CHANNEL RATE VERSUS NUMBER OF SOURCES. (BUFFER LARGE ENOUGH ONLY FOR IN-FIELD SMOOTHING. BUFFER OVERFLOW OCCURRED LESS THAN 0.1 PERCENT OF THE TIME.)

Number of Sources, N_S	Channel Rate per Source (Mb/s)
1	2.0
2	1.7
5	1.25
12	1.0
20	0.90

VI. CONCLUSION

The basic concept of sharing a buffer and a transmission channel between several interframe coders has been discussed with particular reference to the system in Ref. 1, which requires 2.0 Mb/s to transmit a single *Picturephone* signal. Several simulations were described, and it was concluded that five sources could be combined and transmitted on a 6-Mb/s facility with about the same amount of buffering as in Ref. 1; i.e., 1.2 megabits/second and 70,000 bits of buffering per source. With a minimal amount of buffering (a few thousand bits per source) a 2:1 reduction of the bit rate in Ref. 1 can be obtained if twelve sources are combined prior to transmission; i.e., 1.0 megabits/second per source. Combining more than twelve sources results in relatively little reduction in the required channel rate.

As in Ref. 5 the results of this paper can be used with appropriate scaling for other coding schemes as long as the ratio of addressing bits per cluster to amplitude bits per significant change (12 bits/4 bits, here) is maintained and the 2:1 and 4:1 subsampling is brought in at about the same time. Thus, if 9 bits were to be used for addressing a cluster and 3 bits per significant change, then the channel rate and queue length scales should be all multiplied by 2/3.

The simulations could have been carried out more easily had there been available a reliable statistical model of human movement in *Picturephone* scenes. Section III goes a long way toward providing such a model. For a single source the number of significant changes per field has an exponential probability density function in the region of interest; however, there is a strong field-to-field correlation resulting in large data peaks which can last for many fields. It is this characteristic that makes channel sharing so effective in reducing the required bit rate.

VII. ACKNOWLEDGMENTS

Mrs. M. A. Franke and F. W. Mounts constructed much of the hardware used to collect the data. K. A. Walsh acted as a subject. Helpful discussions were also held with J. C. Candy, R. F. W. Pease, J. O. Limb 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. Mounts, F. W., "A Video Encoding System with Conditional Picture-Element Replenishment," B.S.T.J., *48,* No. 7 (September 1969), pp. 2545–2554.
3. Seyler, A. J., "The Coding of Visual Signals to Reduce Channel-Capacity Requirements," Proc. IEE, *109,* (September 1962), pp. 676–684.
4. Limb, J. O., and Pease, R. F. W., "Exchange of Spatial and Temporal Resolution in Television Coding," B.S.T.J., *50,* No. 1 (January, 1971), pp. 191–200.
5. Limb, J. O., "Buffering of Data Generated by the Coding of Moving Images," B.S.T.J., this issue, pp. 239–259.
6. Seitzer, D., "Instantaneous Priority Multiplexing of Gray Pictures," IEEE Trans. on Commun. Technology, (October 1970), p. 690.
7. Urquhard-Pullen, D. I., "Improvements Relating to Multiplexed Transmission Systems," British Patent No. 1057024.
8. Cutler, C. C., and Mounts, F. W., private communication.
9. Seyler, A. J., "Probability Distribution of TV Frame Differences," Proc. IREE Australia, (1965), p. 355.
10. Pease, R. F. W., "Frame Difference Signals for a Personal Communication Television System of 271 Lines," internal memorandum.
11. Pease, R. F. W., and Limb, J. O., "A Simple Interframe Coder for Video Telephony," B.S.T.J., *50,* No. 6 (July–August 1971), pp. 1877–1888.