

Design Technique for Multiplexing Asynchronous Digital Video and Audio Signals

JONATHAN H. STOTT

Abstract—To apply digital transmission to the distribution of television program material it is necessary to encode both sound and picture signals, protect them from the effects of transmission errors and multiplex them into a form suitable for the transmission link that is to be used. Coding, error protection and multiplexing have all received extensive study but their combined application in a system has received little discussion in the literature. This paper shows how a system was designed for a particular experimental application with a specified total bit rate. Particular emphasis is given to the multiplexing arrangements, including the necessary use of justification techniques. The equipment designed has been used in successful field trials of digital transmission by both cable and satellite.

1. INTRODUCTION

This paper describes the design of experimental equipment which assembles two similar 60 Mbit/s data streams, each providing one video channel and six high-quality sound channels, and combines them to provide a single 120 Mbit/s data stream.

The equipment was developed for use in joint BBC-UK Post Office (UKPO) field trials of digital transmission. Its first application was in tests of an experimental 120 Mbit/s digital line system in 1975 [1] and in 1976 it was used for experimental digital transmissions at 60 Mbits/s via an Intelsat IV satellite [2].

The main aim of the paper is to describe the multiplexing requirements and the design approach used to satisfy them.

2. SYSTEM REQUIREMENTS

2.1 Primary Objectives

The original requirement was to provide a signal for experiments with the 120 Mbit/s digital line system. What form should the 120 Mbit/s signal take? It would clearly have been possible to convey one video signal using well-known techniques, namely 8-bit linear pulse-code modulation (p.c.m.) with around $3 \times f_{sc}$ * sampling frequency. This requires about 120 Mbits/s if simple error protection is included using a concealment technique [3]. However, it was felt that little new experience would have been gained by the use of this simple system. Two contemporary developments suggested a more efficient approach:

(i) the invention of a sub-Nyquist, $2f_{sc}$ sampling frequency technique for coding a composite PAL video signal [4]; and

(ii) the development of 2048 kbit/s digital 6-channel sound multiplex equipment [5].

Devereux [6] demonstrated that existing differential p.c.m. (d.p.c.m.) and hybrid d.p.c.m.** (h.d.p.c.m.) video coding techniques could be used in conjunction with sub-Nyquist sampling to offer a substantial total saving in bit rate. For example, the combination of $2f_{sc}$ sampling with 6-bit d.p.c.m. or h.d.p.c.m. coding requires only 53.2 Mbits/s, although some additional bit rate is desirable for error correction coding [7, 8]. Two such video signals may be accommodated within 120 Mbits/s with capacity to spare; sending extra signals thus becomes possible.

Any digital transmission system eventually adopted for television program distribution will carry both video and audio signals; experience of multiplexing such signals in digital form was therefore desirable. The 2048 kbit/s sound multiplex equipment appeared to be a convenient source of digital sound signals for the experiment since it was already under development within the BBC.

It appeared attractive to form a 120 Mbit/s signal from two similar 60 Mbit/s 'packages', each including one video channel and making some provision for high-quality sound. This idea was supported by the prospect of further trials with other transmission systems, some of which would be restricted to about 60 Mbits/s, for example digital transmission through a satellite transponder with 40 MHz nominal bandwidth.

A block diagram of suitable sending equipment is shown in Fig. 1.

2.2 Package Configuration

The inputs and outputs of the necessary multiplexing equipment will now be specified in greater detail.

The UKPO experimental digital line system uses the HDB3 ternary code [9] for interfacing at 120 Mbits/s; separate interfacing units were constructed to convert between this and binary NRZ data with clock. A tolerance of ± 15 p.p.m. (parts per million) is specified for the clock frequency, nominally 120 MHz.

The sound multiplex equipment produces a serial stream at 2048 kbits/s ± 50 p.p.m. The six sound signals which it may carry are each sampled at 32 kHz and quantized to 13-bit accuracy; the samples are then compressed to $10 \frac{1}{16}$ bits/sample using near-instantaneous companding [5]. The equipment incorporates its own provision for error protection using a concealment technique.

** In this context, 'hybrid d.p.c.m.' means that both p.c.m. and d.p.c.m. code-words are transmitted [6].

Manuscript received November 14, 1977.

The author is with the Research Department, British Broadcasting Corporation, Tadworth, Surrey, England.

* f_{sc} = color subcarrier frequency = 4.43361875 MHz for the PAL 625-line systems.

0090-6778/78/0500-0601\$00.75 © 1978 IEEE

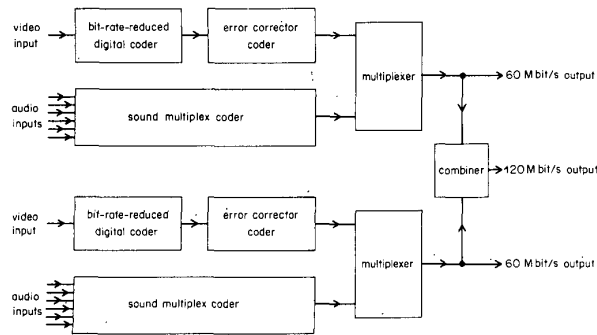


Fig. 1. Sending equipment: general block diagram.

The video coding equipment samples the video signal at twice the color subcarrier frequency and quantizes each sample to 8-bit accuracy. The 8-bit samples are then reduced to 6 bits using d.p.c.m. or h.d.p.c.m. coding. The digital input to the transmission equipment is thus in parallel form, comprising 6-bit words at $2f_{sc}$. The tolerance of the clock frequency is therefore that of the color subcarrier, namely ± 0.25 p.p.m. (national standard) or ± 1.1 p.p.m. (international standard).

An error correction coder increases the video bit rate before the multiplexer proper. The extent to which the bit rate is so increased forms part of the process of system design and is considered in Section 3.

The input and output bit rates of the multiplexer discussed here are unrelated and indeed may vary within their individual tolerances independently of each other. To multiplex such asynchronous signals requires the use of a technique known as justification. This will be described in Section 4.

3. CHOICE OF VIDEO CODING AND ERROR CORRECTION CODING

Clearly, the sum of the bit rates of the individual signals which are to be multiplexed to produce one 60 Mbit/s signal must not exceed 60 Mbits/s. If we allocate 2048 kbits/s for a sound-signal multiplex, 57.952 Mbits/s remain for the video signal (including error-correction coding) and any necessary control signals for framing and justification.

Step 1: What are the possible basic video bit rates?

Assuming $2f_{sc}$ (i.e. sub-Nyquist) sampling and some form of differential coding giving an integer m bits/sample we obtain the following table:

m	bit rate (Mbits/s)
4	35.47
5	44.34
6	53.2
7	62.07

As m is reduced, either the picture quality suffers or the cost of a codec (coder and decoder) increases, or both. Thus the highest usable value of m is to be preferred.

6 bits/sample, occupying 53.2 Mbits/s, seemed a suitable choice and was attractive as a suitable codec was already avail-

able. Operation at a reduced number of bits per sample is, of course, possible for comparison, in which case dummy information would be sent in the unused bit positions. In fact, an h.d.p.c.m. option using 5 bits/sample was included in the equipment.

Step 2: What form of error-correction coding should be employed?

Previous work [7, 8] demonstrated the effectiveness of the Wyner-Ash (8, 7) code in correcting errors in a digital video d.p.c.m. or h.d.p.c.m. signal. Using this form of error-correction coding on all six bits in each sample, the total video bit rate becomes:

$$(53.2 \times 8)/7 = 60.8 \text{ Mbits/s, which is too great.}$$

In the previous work the (8, 7) code was used to correct only the four most significant bits of the 6-bit word, and this arrangement would reduce the bit rate sufficiently for the present problem. However, the previous results suggested that it would be better on balance to correct all six bits although this meant using the slightly less effective*** Wyner-Ash (16, 15) code.

A (16, 15) codec is slightly more complex than one for the (8, 7) code, but this is balanced by the awkward arrangement of data and parity bits which would be entailed with correction of only 4 bits out of 6 with the (8, 7) code.

Using the (16, 15) code in our application makes the total video bit rate equal to:

$$(53.2 \times 16)/15 = 56.75 \text{ Mbits/s.}$$

This represents a comfortable fit within the 57.95 Mbits/s available for video and control signaling; the (16, 15) code was thus chosen.

4. JUSTIFICATION

Before proceeding further with the system design a brief description of 'justification', sometimes known as 'pulse stuffing', will be given.

4.1 The Need for Justification

Suppose that bit stream A ('the input') of bit rate f_A is to be incorporated into another stream B ('the output') of bit rate f_B ($f_B > f_A$). Clearly, if f_A bears a fixed rational relationship to f_B , i.e. if $f_A : f_B = p : q$ where p and q are integers, then a cyclic format for stream B could be contrived. For example, every q bits in stream B could contain p bits from stream A and $(q - p)$ other bits which could either be derived from other input channels or merely fill the space with dummy information. We will see later (Section 5.5) that the problem of multiplexing data and parity information in the error-corrector coder is of this form.

However, where no fixed rational relationship between f_A

*** Using a (16, 15) code in place of (8, 7) reduces the acceptable transmission-path error rate (for the same decoded output error rate), multiplying it by about 0.62 (see Reference 7).

and f_B holds, such a fixed cyclic format is not possible. In such a situation justification is required.

4.2 General Principles of Justification

The basic idea of justification techniques is to allocate a dual function to certain digit time slots[†] in the output stream. These may carry either input information or dummy information *on a time-varying basis* as appropriate to ensure that the average rate of transmission of input information equals the input bit rate. Obviously some form of signaling must also be provided to indicate whether these time slots carry real or dummy information at any given time.

Such dual function time slots are known as '*justifiable time slots (JTS)*' and the signaling as to their content is carried by a '*justification control signal (JCS)*'.

The output stream is usually divided into frames of fixed length within which certain digit time slots are designated as JTS or for the JCS. If there is more than one input, as in multiplexing, then each input has its own JTS and JCS positions allocated within the frame. With such a frame structure, two implementations are often used, as discussed in the literature [10]: 'positive justification' and 'positive/negative justification'. Other systems are possible, however.^{††}

4.3 Positive Justification

For this system each input is allocated one JTS per frame. Clearly, the maximum permissible input bit rate is transmitted when input data occupy the JTS in every frame; conversely, the minimum input bit rate will require sending dummy information in the JTS in every frame. The range of operation must always lie between these extremes if no data are to be lost, and normally the frame length is chosen so that neither extreme is approached very closely.

We may define a few terms, paraphrasing CCITT Recommendation 702:

The *justification rate* is the average rate at which dummy information is carried in the JTS.

The *justification ratio* is the ratio of the justification rate to the maximum possible justification rate (when every JTS is dummy). In other words, the justification ratio represents the fraction of the time that the JTS carries dummy information.

The *nominal justification rate* and the *nominal justification ratio* are the values of the above parameters when the input and output bit rates assume their nominal values.

4.4 Positive/Negative Justification

For this technique each input is given two JTS locations in every frame, JTS 1 and JTS 2.

When input and output rates have their nominal values, one JTS, say JTS 1, carries input information and the other, JTS 2, carries control information. When the input rate is higher than nominal (or, equally, the output rate is lower than nominal)

JTS 2 will sometimes also carry input information. Conversely, when the input rate is lower than nominal (or, equally, the output rate is higher than nominal) JTS 1 will sometimes carry dummy information.

The nominal justification rate is thus effectively zero; as input and output rates vary over their tolerances both positive and negative values of justification rate will occur.

Note, that if nominal input and output rates were always maintained, the system simplifies to the form first discussed in Section 4.1 above. The purpose of justification applied in this way would be merely to absorb the tolerances of the input and output rates whose nominal values must have a rational relationship. If no rational relationship exists, then positive justification would be required even if input and output rates remained exactly constant at their nominal values.^{†††}

Positive/negative justification would appear to have greatest application in the situation where a designer can choose related input and output nominal bit rates. A possible example is the design of a telephone/telecommunication system where a multiplex hierarchy is built up from the lower bit rates upwards. The number of input channels to be multiplexed and the nominal output rate at each level of the hierarchy are parameters over which the designer has some control.

Positive/negative justification appears rather more complicated to instrument than positive justification, seemingly without offering useful advantages.[§]

4.5 Apparatus

Having outlined the general principles of justification, we now consider how to apply them in practice, where input and output nominal bit rates are unrelated, and where positive justification is therefore used.

The most important point to note is that we are considering a situation where input information is provided at a steady rate but is transmitted in an irregular fashion; for any given input the average rate of information transmission equals the input rate but the instantaneous rate varies. This implies a need for storage. An approximate analogy is a gas-holder. This is steadily provided with gas from the gas-producing plant but is subjected to a fluctuating demand. The average consumption must equal the rate of production or the gas-holder will become completely full or empty. Furthermore the capacity of the gas-holder must be large enough to cope with the prevailing short-term changes of demand (e.g. the cooking of Sunday lunch). We may think of this as corresponding to the gaps in frame structure where a frame-alignment sequence, justification-control signal or information from another input is transmitted. The gas-holder cannot cope with a long-term change of demand (e.g. from summer to winter) and in this case the rate of production would be changed. The analogy is slightly imperfect at this point since a change in the ratio of

[†] A digit time slot is a cyclic time interval, which can be recognized and defined uniquely, allocated to a single digit (CCITT Rec. 702).

^{††} For example, a system devised by J. P. Chambers was used for pilot field trials [1] in which 119.7 Mbits/s of information were to be carried in 120 Mbits/s. In that system it was the frame length which was changed as necessary on a time-varying basis.

^{†††} Positive/negative justification could be used with non-integer relationships; however, there seems little reason to do so. The system would amount to positive justification for nominal rates, and the region around nominal, and would only have a true positive/negative function when at one extreme of the range.

[§] It is interesting to note that the UKPO has specified positive justification for their digital hierarchy [10].

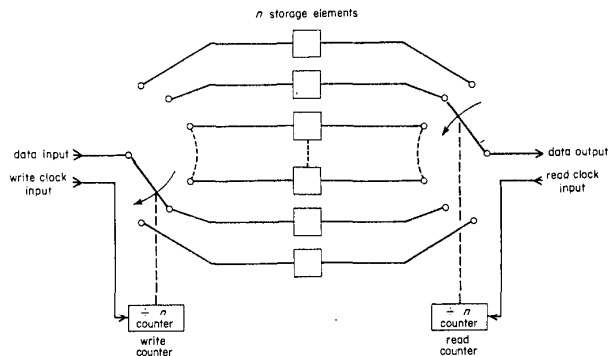
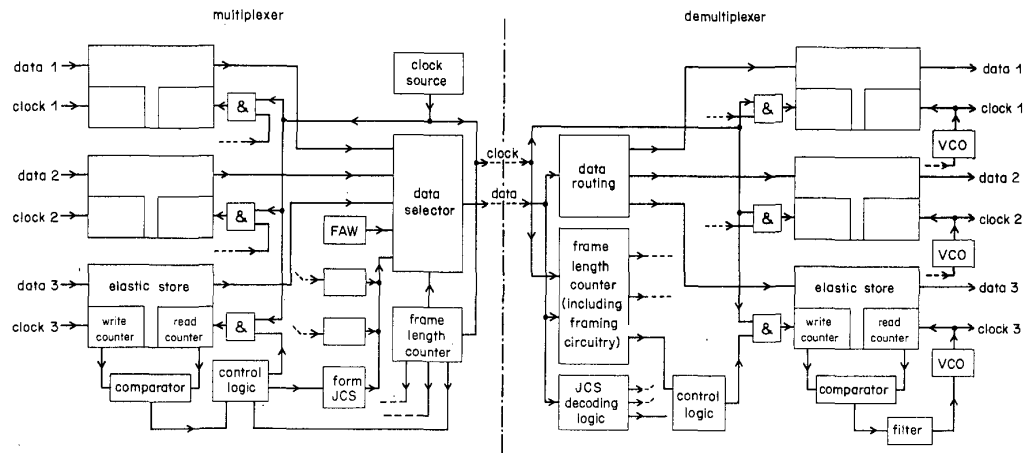
Fig. 2. n -stage elastic store arrangement.

Fig. 3. Illustration of multiplexing using justification.

input and output rates is accommodated by a change of justification ratio—in other words an increase of input rate is matched by an increase in average transmitted rate.

Thus we require storage of variable capacity—sometimes called 'elastic' storage. Fig. 2 shows a typical arrangement for providing this. Data are written into n storage elements in turn, directed by a $\div n$ counter—the 'write counter'. Another counter, the 'read counter', causes data to be read out of the stores in turn at the rate demanded by the read clock. If no data are to be repeated or lost the switches must not overtake each other. The amount of data in the store at any time is given by the number of switch positions by which the write counter leads the read counter. As with the gas-holder, the store must lie between empty and full without quite reaching the extremes.

Fig. 3 shows how elastic stores are incorporated into a multiplexer/demultiplexer combination which uses justification techniques. The figure shows three inputs being multiplexed together, but any number of inputs could be used. Each input is written steadily into its own elastic store in the form of Fig. 2.

The transmission-path bit rate is determined by the clock source shown at top right of the multiplexer. It feeds the frame-length counter which divides the transmitted output into frames and controls a data selector to determine what information will be fitted into each part of the frame. This is

laid down by the frame structure (see Section 5). Each frame will contain a frame-alignment word (FAW) which therefore constitutes one input to the data selector. Each input is allocated time slots in the frame; when an input is selected, a gate is opened supplying transmission-rate clock pulses to the read clock of that input's elastic store.

If no dummy information is ever sent, the elastic store will gradually empty. A comparator compares the states of the write and read counters. At some point shortly after the beginning of a frame the output of the comparator is examined to determine whether justification—sending a dummy bit—is necessary in that frame. The result of this decision is conveyed in the justification control signal (JCS). When the justifiable time slot (JTS) for the input is reached, the output of the corresponding elastic store is selected. If a data bit is sent, the read counter is advanced normally; but if justification is required the read counter is not advanced and a dummy bit is sent. In this latter case, the contents of the store are not decreased in this time slot. To continue with the gas-holder analogy we may suppose that the contents of the gas-holder, measured daily at a particular time, would gradually fall as winter approached. Eventually a point would be reached where gas production would have to be increased to restore the former level.

The processes we have described are reversed in the demultiplexer. There is a frame length counter which must become synchronized with the received frame structure by making use

of the frame-alignment words. This being done, the function of each piece of received information is known, and data originally coming from, say, input 1 are directed to the elastic store for output 1. The justification control signal in every frame is identified and decoded so that it is known whether the JTS pertaining to each input contains real or dummy information. All real information is written into the corresponding elastic store, the dummy information being discarded.

The elastic stores in the demultiplexer have an exactly opposite function from those in the multiplexer; they are provided with information in irregular bursts and must feed it out again smoothly. To do this it is necessary to provide the store with a stable, regular, read clock which has the same rate as the original input clock which fed the multiplexer. This smooth clock is derived using the fact that information is fed into the demultiplexer elastic store at a rate whose *average* equals the original multiplexer input clock rate. (This must be so, or the multiplexer elastic store would overflow or underflow.) A phase-locked-loop arrangement is used which includes a voltage-controlled oscillator (v.c.o.) and the write and read counters. A comparator detects whether the elastic store is more or less than half full. Its output controls the frequency of the voltage-controlled oscillator via a filter so that the store remains, on the average, half full. The loop bandwidth is made small so that the irregularities at the input are not followed by the v.c.o. from which the read clock is derived.

The fundamental limitation of the system is that it is not possible to smooth out the variations completely; in other words, there will be residual jitter on the output clock. We may think of two types of jitter introduced by the multiplexer/demultiplexer: (i) that arising from the regular gaps in the frame structure (e.g. for the FAW or data from other inputs) and (ii) that arising from justification.

The jitter of type (i) has a fundamental component at frame rate and may therefore be removed by making the loop bandwidth in the demultiplexer somewhat smaller. This is normally quite feasible as the frame rate is usually fairly high to avoid excessively long frames.

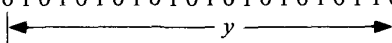
Type (ii) jitter is more of a problem since it may have components at very low frequencies which will not be attenuated by the demultiplexer phase-locked loop. The following example illustrates how these low frequencies arise. Suppose that input and output rates are such as to give a justification ratio of exactly $\frac{1}{2}$. If we write 0 and 1 to denote frames containing real and dummy information in the JTS we obtain the following sequence:

... 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 ...

and so on. The resulting jitter will have its lowest frequency component at $\frac{1}{2}$ frame rate—again not very troublesome.

If the ratio of input and output rates is changed very slightly the sequence might, for example, become:

... 0 1 0 1 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 1 0 1 0 1 0 1 0 1 ...



The sequence now repeats after y frames, in other words the fundamental jitter component is at frame rate/ y . As the justification ratio approaches $\frac{1}{2}$ more closely the fundamental frequency becomes lower and lower without limit.

A similar 'beating' effect will occur as the justification ratio approaches any exact fraction. This effect has received considerable attention in the literature [10, 11]. It appears to be wise to avoid justification ratios which are near unity, zero, or a ratio of integers (a/b) where b is small. Ref. 10 suggests 0.42 (close to $3/7$) as a good compromise. In fact we will see in Section 5.2 that the present design problem has insufficient variables to allow a free choice of justification ratio.

5. MULTIPLEX DESIGN

5.1 Structure—Parallel or Serial?

The apparatus has to generate a serial output and is presented with parallel digital video data; it must therefore incorporate elements of both parallel and serial operation. We must choose where to draw the dividing line between them.

The Wyner-Ash error corrector operates in parallel; there are six separate error correctors, one for each level of significance of the video sample word. This configuration has been shown [7] to be desirable because:

(a) It reduces the speed of logic operations, making circuitry both easier and cheaper to build.

(b) When this parallel structure is eventually serialized the six single-error-correcting codes are effectively interlaced. The serial stream thereby acquires a measure of burst-error correction.

The multiplexer proper thus receives one serial input (2048 kbit/s sound) and one parallel (56.75 Mbit/s error-correction-coded video). Two approaches are suggested:

System 1

Serialize the video signal and then, using justification techniques, multiplex the resulting serial video signal with the existing serial sound signal. The multiplex thus has a serial frame structure.

System 2

Convert the serial sound signal to a 6-bit parallel form. This parallel signal is multiplexed with the 6-bit video signal, using justification, to produce a 6-bit, 10 M word/s signal which is finally serialized to produce a 60 Mbit/s serial output. The multiplex thus has a parallel frame structure.

System 1 might be termed the classical and obvious approach, whereas System 2 at first appears needlessly overcomplicated. We will now discuss the points in favor of System 2.

Parallel operation here has the usual advantage of needing a slower clock rate. This greatly eases any design difficulties, and although extra storage will be required compared with serial operation the total cost is reduced by the much lower cost of slow-speed devices.

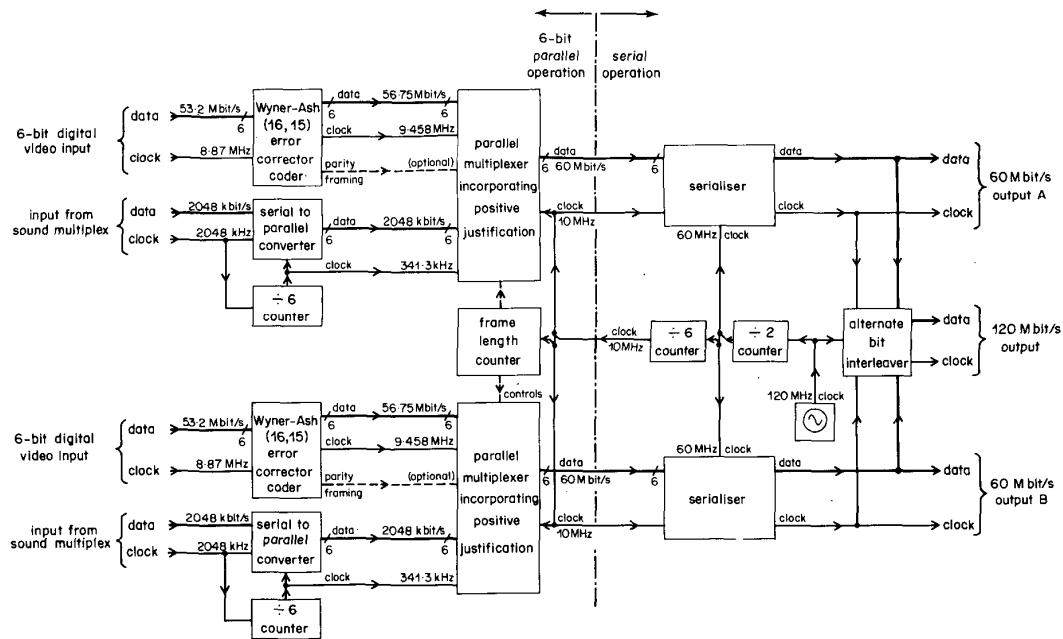


Fig. 4. Detailed block diagram of sending equipment for 120 Mbit/s operation.

Framing is simplified with the parallel structure. There are three aspects to be considered:

- (i) multiplex framing,
- (ii) word framing of the video signal,
- (iii) parity framing of the video error-correction system.

Multiplex framing is the problem of recognizing the frames into which the total multiplexed signal is divided. It is necessary so that demultiplexing and removal of dummy information may take place.

Video word framing relates to the problem of regrouping video bits into sample words, with the bits in each word in the right order from most to least significant.

Parity framing relates to distinguishing data and parity words in the video bit stream.

If a *serial* structure is used, all three of the above operations will have to be performed separately, and will require some form of framing code to be added to the (in this case) serial video stream in order that video words may be retrieved from it. This framing must be independent from the multiplex framing because of the independence of the video and output bit rates.

With a *parallel* structure these two functions (i.e. (i) and (ii) above) may be combined. The whole frame is made of words of equal length to the video words (6 bits), and sound, frame-alignment words, JCS and JTS are all in 6-bit form. Thus if we specify that video words are inserted with, say, the most significant bit first, then we know that the first video bit following a frame-alignment word is a most significant bit (m.s.b.) and the last video bit before the next FAW is a least significant bit (l.s.b.).

The method of parity framing does not significantly affect the choice of structure since the multiplex frame cannot con-

§ § The sound multiplex includes its own framing which need not be considered here.

tain a fixed multiple of 16 video words. Error-corrector framing must therefore be achieved independently and we will discuss the methods of doing this in Section 5.5.

Using parallel operation the justifiable time slots for sound and video become 6-bit words. The amplitude of the jitter resulting from justification will thus be increased, and its frequency lowered. This would normally be an undesirable state of affairs since lower frequency jitter is very difficult to remove with the phase-locked loop in the demultiplexer. However, the specifications of jitter tolerance for both sound [12] and video [13, 14, 15] in fact allow for an increase of jitter amplitude with decreasing frequency for very low frequencies. Provided careful attention is paid to the bandwidth of the demultiplexer phase-locked loops, parallel operation need not therefore cause insuperable difficulties with timing jitter.

Parallel structure thus has many advantages without any serious disadvantages; it has therefore been adopted. The block diagram of the sending equipment with this structure is shown in Fig. 4.

The equipment actually built includes a degree of flexibility so that different arrangements may be used for other experiments than the original 120 Mbit/s tests. For example, the satellite tests [2] required a single 60 Mbit/s package in the form of two 30 Mbit/s streams which were obtained by a rearrangement of the system of Fig. 4.

5.2 Frame Length and Related Parameters

Having opted for a parallel structure, with multiplexing on a word-by-word basis (where 1 word = 6 bits), we pass to the choice of the multiplex frame length, x words, say. In this situation where all the bit rates are already specified we wish to choose a fairly short frame length in order to minimize both circuit complexity and framing lock-up time.

For the reasons outlined in Section 4.4, positive/negative justification has little to offer for the situation of unrelated bit

rates. Positive justification has therefore been specified, requiring 1 JTS word per input per frame.

Although one word per frame could provide JCS signaling for both sound and video justification it seems advisable to protect against error bursts by distributing this information through the frame. It was decided therefore to provide three JCS words, each containing both sound and video JCS information, and to use a majority vote system in decoding. Within these words, which must each contain 2 bits of justification control information (i.e. one for video and one for sound), either further redundancy could be provided or the spare bits could be used for auxiliary signaling.

Each frame will therefore consist of the following components:

- y frame-alignment words (FAW) where $y \geq 1$
- 3 justification control signaling words (JCS)
- video words, including one justifiable time slot for video (JTSV)
- sound words, including one justifiable time slot for sound (JTSS).

It was thought that $y = 2$ would be preferable to $y = 1$ in order to give ample ruggedness of the 60 Mbit/s multiplex.

The number of words in each frame devoted to an input of rate f (including its JTS) is given by the expression:

$$(xf)/(60 \times 10^6) \text{ rounded up to the next integer.}$$

The difference between the expression and its rounded up value gives the justification ratio. Substituting the values of f for sound and video we can write an expression giving a lower bound for x :

$$Sx + Vx + 3 + y \leq x$$

$$\begin{array}{cccc} \text{sound} & \text{video} & \text{JCS} & \text{FAW} \\ \text{words} & \text{words} & & \end{array}$$

$$\text{where } S = \frac{2.048}{60} = 0.034133333$$

$$\text{and } V = \frac{56.75032}{60} = 0.945838667$$

$$\text{i.e. } x \geq \frac{3 + y}{0.020028}$$

$$\text{if } y = 1 \quad x \geq 119.72$$

$$\text{if } y = 2 \quad x \geq 249.65.$$

These are lower bounds, all higher value integers are not necessarily usable. For example trying $x = 250$:

$$S \times 250 = 8.533 \text{ rounding up to 9 with justification ratio } 0.4667$$

$$V \times 250 = 236.459 \text{ rounding up to 237 with justification ratio } 0.5403$$

Total video plus sound words = 246 per frame.

Thus there is only room for 3 JCS + 1 FAW.

Working on this way, restricting the choice to those values

of x which conveniently factorize, §§§ we reach $x = 300$ for which: $S \times 300 = 10.24$ rounding up to 11 with justification ratio 0.76, $V \times 300 = 283.7516$ rounding up to 284 with justification ratio 0.2484. Total video plus sound words = 295 per frame leaving room in this case for the desired 3 JCS + 2 FAW. We may note that the nominal justification ratios are close to $\frac{3}{4}$ and $\frac{1}{4}$ respectively. These are not ideal, but at least we avoid the worst figure of $\frac{1}{2}$. We do not have control of enough variables (e.g. choosing our own output rate) to optimize the justification ratio as is possible when designing a hierarchical system *ab initio*.

A further check shows that the frequency tolerances of input and output may be safely accommodated.

The next section describes the detailed structure chosen; we will see that this frame length allows a tidy solution.

5.3 Detailed Frame Structure

Summarizing the preceding section, we have the following words to locate in the 300-word frame:

10 sound + 1 JTSS

283 video + 1 JTSV

3 JCS

2 FAW.

The 2 frame-alignment words may conveniently be located together, effectively forming a bunched 12-bit framing pattern. These are designated words 0 and 1 in the frame (which then runs from word 0 to 299).

The 10 'permanent' sound words may be equally spaced throughout the frame in every 30th word location. This has the advantage that less elastic storage is required than if the sound words were grouped together or spaced irregularly. The remaining 'special' words, i.e. 3 JCS, JTSS, JTSV and 2 FAW may now be placed in the mid-points between sound words. In fact we need only place 'special' words in every other space. Note that the JCS words should precede the JTSS and JTSV. All the remaining word positions are of course allocated to video words. Fig. 5 shows the structure chosen in detail.

This structure is convenient in that the word locations are easily decoded when the frame length is defined by a $\div 300$ counter comprising $\div 15$ and $\div 20$ counters in cascade.

5.4 Demultiplexing Framing Strategy

There are two aspects of demultiplexer framing strategy in this particular system.

The basic problem is that the demultiplexer must recognize the frame structure using the frame-alignment words so that it can demultiplex the signal into its component parts. Furthermore, it must do this reliably and quickly even in the presence of transmission errors.

The secondary problem arises when the two 60 Mbit/s streams have been interleaved to produce the 120 Mbit/s signal. It is then necessary to separate the combined signal back into the two streams, say A and B . We must therefore provide means to identify these two streams.

§§§ It is convenient to choose a value of x which factorizes in such a way that it is relatively easy to generate the logic waveforms which define the components of the multiplex frame. Moreover, a multiple of 30 is attractive since the ratio S of sound words to the total output is close to $1/30$.

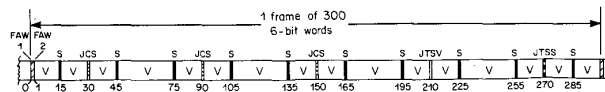


Fig. 5. Frame structure of the 60 Mbit/s multiplex.

The solution to this problem of channel identification and separation is simple and will be discussed next, as it has a bearing on the detailed strategy of frame-alignment.

The two 60 Mbit/s streams must be synchronous in order that they can be interleaved to produce a 120 Mbit/s output. Considerable simplifications occur if the two channels of the multiplexer share not only a common clock, but also common housekeeping waveforms related to the frame structure. Thus, for example, the frame-alignment words occur in the two streams at the same time, and so on for the other time-slots through the frame. If we specify that corresponding bits from the *A* and *B* streams are interleaved always placing the bit from the *A* stream first, the structure is unambiguous. Indeed, it is not even necessary to have different frame-alignment words for the *A* and *B* channels.

The general method of framing is simple. Initially, the received stream is searched until the frame code is found, after which it is to be expected that the frame code will recur at regular intervals of one frame length. The signal is examined at these points to confirm the presence of the frame code where it is expected; if the code is consistently absent, then a fresh search must be made.

Points concerning framing which need consideration are:

(i) It is possible for a sequence identical to the frame code to occur in the data—the problem of data ‘imitating’ the frame code.

(ii) In the presence of transmission errors the correct frame code will not always be present in the expected position although frame synchronization remains correct. The circuitry must therefore distinguish between such occurrences and genuine loss of synchronization.

A conventional way of dealing with (ii) above is to count the number of successive framing codes which are incorrect. When a certain value is exceeded synchronization is deemed to be lost. As this threshold value is increased so the ruggedness of the system is increased. However, this can lead to a conflict with (i) above, since the time to recognize a false synchronization produced by imitation is also increased.

A slight modification has been made to this technique whereby effects (i) and (ii) may be catered for independently. Before describing this solution, some further important points particular to our design problem will be listed:

(a) allowance is to be made for 1×60 Mbit/s, 2×60 Mbit/s and 120 Mbit/s interleaved operation;

(b) circuit design at 120 MHz is limited by the speed of available devices. Furthermore, it is desirable to limit the extent of circuitry operating at 120 MHz and even 60 MHz on the grounds of cost.

Point (a) is dealt with by basing the framing operation at the 60 Mbit/s level. If operation is desired with a 120 Mbit/s interleaved signal, then a simple circuit to split this into two 60 Mbit/s streams is used. This ‘de-interleaving’ operation also helps with point (b) since it requires the minimum of high-speed logic.

Noting that the frame code is effectively 12 bits long (1×60 Mbits/s) or 24 bits (2×60 Mbits/s), we can afford to make a further simplification in the interests of point (b). It is preferable to search in the serial stream for the frame code, the alternative—operating in the parallel stream—being extremely complicated. However, serial-stream searching would require a long register and much associated fast circuitry. The compromise adopted is to search for only the first frame-alignment word (FAW 1), i.e. one half of the frame code, in the serial stream, and thereby gain synchronization. Subsequently the presence of the whole frame code where expected is checked, in the serial stream for FAW 1 and in the parallel 10 Mword/s stream for FAW 2.

The detail of the framing strategy will now be described, assuming in the first instance that one 60 Mbit/s channel is in use. Once FAW 1 has been located the presence or absence of FAW 1 and FAW 2 where they are subsequently expected controls an up/down counter. Its state, say *N*, is increased by 1 for every incorrect FAW. It is decreased by 1, provided *N* is not already zero, for every correct FAW. If no errors are present and correct synchronization has been obtained the counter will quickly clock down to zero and retain that value. If transmission errors are present but synchronization is maintained then the counter will sometimes clock up, but will quickly clock down to zero. If synchronization is lost, the counter will steadily clock up and when a threshold value, say *T*, is reached, synchronization is deemed to be lost. The process of searching for FAW 1 now takes place. When it is found the up/down counter is set to a value, say *R*, the residue, whence it will eventually clock down to zero if correct synchronization has been obtained.

Fig. 6 illustrates the process in the form of a flow chart.

The important point to note is the use of the value *R*; it is the key to separating ruggedness and lock-up time since ruggedness depends on *T* and lock-up time on $(T - R)$. As *T* is increased, the transmission error rate required to break synchronization increases. On the other hand the lock-up time depends on how quickly a false lock caused by imitation of FAW 1 is detected. This is quickened by reducing $(T - R)$, the limit being set by the need to avoid loss of a newly acquired correct lock in the presence of errors. If we increase *T* but keep $(T - R)$ unaltered, the resistance to transmission errors when in the locked condition increases without any penalty in lock-up time.

If interleaved operation is in use an extra operation is introduced which is shown in the flow chart as a dotted line. As the de-interleaver arbitrarily splits the incoming stream in two there is no way of knowing which stream is which initially. The search for FAW 1 is therefore conducted looking for either correct operation or the ‘channels reversed’ condition. If the latter is found the de-interleaving operation is interrupted for one clock pulse so that the streams are now sent to their correct destinations. Clearly, this operation must be inhibited once lock is obtained. Note that it is possible to detect the channels reversed condition even if the *A* and *B* streams have identical frame-alignment words. This is a result of the time sequence of the bits wherein *A* always leads *B*.

The equipment constructed has the facility to vary the threshold *T* from 1 to 8 and the residue *R* from 0 to 7. Frame

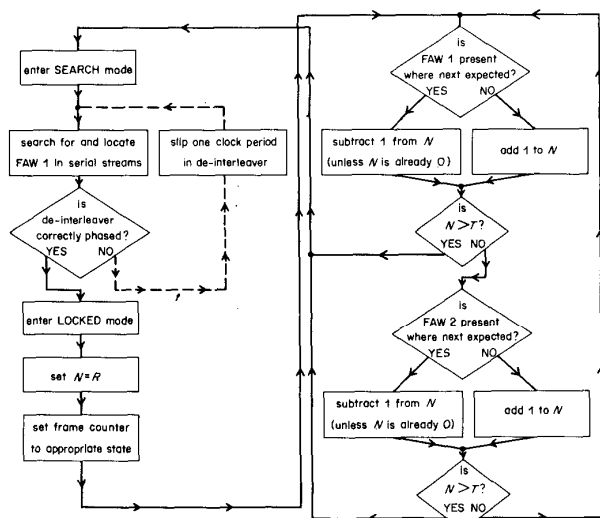


Fig. 6. Framing strategy flow chart. N , state of up-down counter. T , threshold. R , residue. †, only used for interleaved operation.

synchronization behavior in practice bears out the argument above. If T is varied from 2 to 8 the permissible random error rate varies from 4 in 10^5 to > 2 in 10^2 , and with $T = 8$ the highest observed time to achieve frame synchronization varies from 400 to 200 μ s as R is varied from 0 to 7; in the presence of errors at a rate of 2 in 10^2 this range is only increased by about 50 μ s. The total lock-up time (including the error corrector and settling of all phase-locked loops in the system) is about 0.3 s.

5.5 Parity Framing of the Video Error Corrector

The error corrector operates in parallel, each of the six bits in each video sample having its own (16, 15) Wyner-Ash coder and decoder. Each of the six Wyner-Ash coders regularly produces one parity bit for every 15 data bits fed into it. This is accommodated by having an output clock rate which is (16/15) times the input clock rate, with the parity bits incorporated in the output stream at every 16th bit position. This process takes place in all six error-correction coders, and a simplification results if all these coders work in step, sharing common control waveforms. Each will then produce a parity bit at the same time so that the parallel output consists of 6-bit words, at (16/15) times the input rate, of which every 16th 6-bit word is now a parity word.

The decoder must identify the parity words before correction may proceed. Two general techniques are possible:

- (i) to slave the decoder to the coder by an extra signaling channel;
- (ii) to deduce the location of the parity words from the information contained in the data and parity bits themselves. The option of using either technique is provided in the apparatus which has been constructed.

The slaving technique, method (i) above, is provided by making use of some of the spare capacity within the JCS words. When this technique is used the JCS redundancy is reduced so that only two of the six bits in each JCS word are

used for justification signaling and the remaining four bits are available for other purposes. One of these bits is used to signal parity framing, leaving three bits per JCS word for other auxiliary signaling.

In order to explain the alternative, 'self-seeking' technique (method (ii) above) we will first consider the situation when the input to the error correction decoder equals the coder output, in other words no errors have been introduced in transmission. If the decoder correctly identifies the parity word it will decode the output data correctly, with all the parity checks it performs being satisfied. If it does not correctly identify the parity word, that is to say it performs parity checks over the wrong groups of digits, then these parity checks will frequently not be satisfied. The decoder will then conclude that it has located some errors and will attempt to correct them by inverting the appropriate bits. Excepting the case of certain rare data signals which are related to the period of parity formation (15 data bits), the apparent rate of error correction will be high, approaching the maximum of which the decoder is capable (once per successive 'constraint length' [7] per decoder).

We may now see how to apply this to method (ii) for parity framing. The decoder incorporates circuitry which monitors the apparent rate at which error correction is taking place. When this rate exceeds a pre-set threshold, close to the maximum possible rate, the decoder concludes that it is looking in the wrong place for the parity, slips one word period and tries again. The correct framing phase will be found after a maximum of 15 such slips.

The drawback of this method is that it limits the maximum error rate which may be corrected. The error-monitoring circuitry mentioned above will of course register when transmission errors are being corrected. If the error rate is high the threshold may be exceeded although the decoder is correctly framed. The setting of the threshold is thus a compromise: too low a setting will restrict the performance unduly, whereas too high a setting may result in the decoder failing to frame up when the data being conveyed is insufficiently random. Furthermore, for a given data signal, the parity framing lock-up time will increase as the threshold is increased.

In practice method (ii) has proved entirely satisfactory—unwanted framing slips do not occur until the error rate is so high that the effects on picture quality, despite the use of error correction, are subjectively unacceptable. The threshold normally used is 5 in 10^3 which gives ample ruggedness with a parity framing time of typically 12 ms; this time is negligible compared with the phase-locked loop settling times.

When the 'self-seeking' technique is in use, the JCS words may carry justification signaling at maximum redundancy and hence maximum ruggedness. This combination of options has normally been used.

5.6 Demultiplexer Phase-Locked Loops

The phase-locked loops are designed to have very low bandwidths (within practical limits) so that maximum jitter attenuation is provided. The video loop has a -3 dB bandwidth of 7 Hz, while that of the sound loop is 10 Hz, and the form of the transfer response is such that jitter attenuation increases with frequency asymptotic to 12 dB/octave. No visible or audible

impairments due to jitter have been detected in practice, either in the laboratory or in the field trials [1, 2]. This is to be expected since the jitter attenuation is only low at low frequencies where the otherwise stringent specifications for jitter tolerance for both sound and video signals are considerably relaxed [12, 13, 14, 15].

6. CONCLUSIONS

It has been found to be practicable to build up a 60 Mbit/s digital transmission package which carries one System I PAL color video signal and six high-quality sound signals, with transmission-error protection for the video and the audio signals. Two of these 60 Mbit/s packages have been combined for transmission at 120 Mbits/s.

The bit rates produced in coding these video and sound signals were independent of each other and of the transmission rate. In consequence a technique called justification was required when the signals were multiplexed.

Equipment constructed according to the design principles set out in this paper has already been used in two highly successful field trials (one via an experimental UK Post Office 120 Mbit/s digital line circuit in 1975, and the other at 60 Mbits/s via an Intelsat IV satellite in joint BBC/UKPO experiments conducted in 1976), and is sufficiently flexible to be adaptable to a number of different transmission systems.

ACKNOWLEDGMENT

The author wishes to thank the Director of Engineering of the BBC for his kind permission to publish this paper.

REFERENCES

- Osborne, D. W. 1976 Experimental digital transmission of multiplexed video and audio signals at 120 Mbit/s. BBC Research Department Report No. 1976/16. An edited version of this Report is a paper presented at the International Broadcasting Convention 1976; IEE Conference Publication No. 145, pp. 84-88.
- Reid, D. F., Stott, J. H. and Osborne, D. W. 1977 Experimental digital transmission of multiplexed video and audio signals at 60 Mbit/s through a satellite. BBC Research Department Report No. 1977/32.
- Devereux, V. G. and Meares, D. J. 1972 Pulse code modulation of video signals: subjective effect of random digit errors. BBC Research Department Report No. 1972/14.
- Devereux, V. G. and Phillips, G. J. 1974 Bit-rate reduction of digital video signals using differential p.c.m. techniques. International Broadcasting Convention 1974. IEE Conference Publication No. 119, pp. 83-89.
- Croll, M. G., Osborne, D. W. and Spicer, C. R. 1974 Digital sound signals: the present BBC distribution system and a proposal for bit-rate reduction by digital companding. Ibid., pp. 90-95.
- Devereux, V. G. 1975 Digital video: sub-Nyquist sampling of PAL colour signals. BBC Research Department Report No. 1975/4.
- Stott, J. H., Oliphant, A. and Osborne, D. W. 1974 Digital video: error-correcting codes and a practical study of a Wyner-Ash error corrector. BBC Research Department Report No. 1974/40.
- Ratliff, P. A. 1974 Digital video: subjective assessment of an experimental Wyner-Ash error corrector. BBC Research Department Report No. 1974/41.
- Moore, T. A. 1974 Digital transmission codes: properties of HDB3 and related ternary codes with reference to broadcast signal distribution. *The Radio and Electronic Engineer*, 1974, 44, 8, pp. 421-427.
- Levett, F. A. W., and Sturch, P. M. E. 1975 Principles of asynchronous digital multiplexing. IEE Conference on Telecommunications Transmission 1975, IEE Conference Publication No. 131, pp. 79-81.
- Duttweiler, D. L. 1972 Waiting time jitter. *Bell System Technical Journal*, 51, 1, pp. 165-207.
- Manson, W. I. 1974 Digital sound signals: subjective effect of timing jitter. BBC Research Department Report No. 1974/11.
- Devereux, V. G. 1971 Pulse code modulation of video signals: subjective tests on acceptable limits for timing jitter in the decoded analogue samples. BBC Research Department Report No. 1971/42.
- Devereux, V. G. and Wilkinson, G. C. 1973 Digital video: effect of PAL decoder alignment on the acceptable limits for timing jitter. BBC Research Department Report No. 1973/1.
- Chew, J. R. and Stebbings, D. W. 1975 Specification of the timing stabilities of broadcast monochrome and PAL colour television signals. BBC Research Department Report No. 1975/14.



Jonathan H. Stott was born in London, England on October 31, 1950. He studied Engineering and Electrical Sciences at Churchill College, Cambridge University, England, receiving the B.A. and M.A. degrees in 1972 and 1976 respectively.

Before University he was with the Information Equipment Group of Ferranti (Scotland) for a year during 1968-9, working on the development of computer interfaces. Since 1972 he has been at the British Broadcasting Corporation Research Department, Tadworth, England, where he is currently a Research Engineer, working mainly on digital television projects.

Mr. Stott is an Associate Member of the Institution of Electrical Engineers.