Michael Robin, a 20-year veteran of the Canadian Broadcasting Corporation Engineering Headquarters, is an independent broadcast consultant located in Montréal, Canada. He is the co-author of Digital Television Fundamentals, published by McGraw-Hill and a contributing writer to Broadcast Engineering Magazine.
mrobin@miranda.com

Technical Notes - 2

Ancillary Data Multiplexing

SPACE FOR ANCILLARY DATA

Figure 1 shows a simplified view of the 4:2:2 525/60 scanning line and the location of video words during the active line interval and the horizontal blanking interval. The combined (multiplexed) number of luminance (Y) and chrominance (C_B and C_R) samples per total line (words per total line) is 1716, numbered 0 through 1715. The digital active line accommodates 720 Y samples, 360 C_B samples, and 360 C_R samples or a total number of 1440 words per active line, numbered 0 through 1439. The resulting number of words in the digital blanking interval is 276, numbered 1440 through 1715.

The component digital standards do not provide for the sampling of the analog sync pulses. Two timing reference signals (TRS) are multiplexed into the data stream on every line immediately preceding and following the start of the active line data. Of the 276 data words in the horizontal blanking interval eight are reserved for the transmission of the TRS. Words 1440, 1441, 1442 and 1443 used to transmit the end of active line (EAV) TRS message and words 1712, 1713, 1714 and 1715 used to transmit the start of active line (SAV) TRS message.

Each TRS consists of a four-word sequence. Using a 10-bit hexadecimal notation these words are represented as follows:

3FF 000 000 XYZ

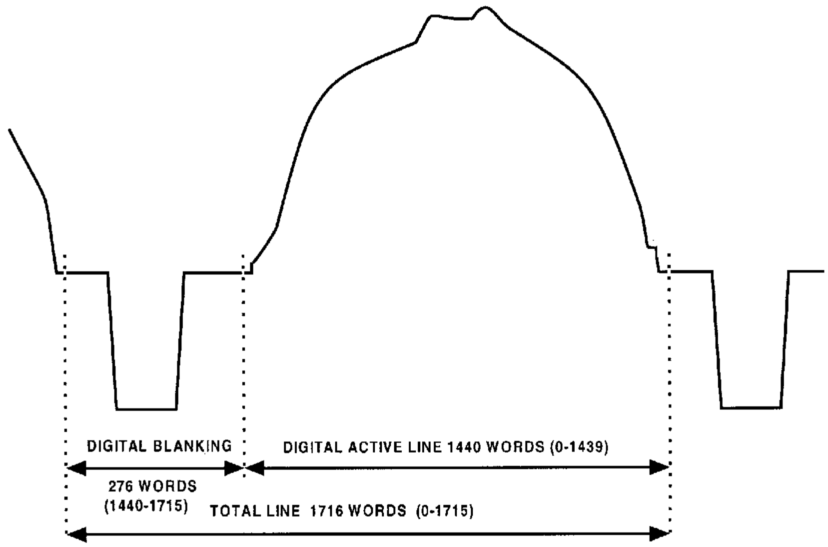

FIGURE 1 SIMPLIFIED VIEW OF 525/60 SCANNING LINE SHOWING LOCATION OF VIDEO WORDS DURING THE ACTIVE INTERVAL AND THE BLANKING INTERVAL

The first three words are a fixed preamble and unambiguously identify the SAV and EAV information. XYZ represents a variable word and defines the field identification, state of the vertical blanking and the state of horizontal blanking.

This leaves 268 words, 1444 through 1711, which can be used to transmit ancillary data. During the vertical blanking duration, large blocks of data, up to 1440 words, can be transmitted within the interval starting with the end of SAV and the start of EAV. Only eight-bit words can be used in the vertical blanking interval. Certain restrictions on the lines that can be used exist allowing only the use of lines 1 through 19 and 265 through 282. Lines 10 (vertical interval switching instant) and the following line 11 are not used to prevent switching clicks. Lines 9 (fields I & III) and 272 (fields II and IV) are reserved for error detection and handling (EDH) signals

Table 1 summarizes the ancillary data space available with the ITU-R601 4:2:2 format. The horizontal ancillary (HANC) capability is listed in the upper row of the table and indicates the bit-rate available for insertion of ancillary data in the horizontal blanking interval. The vertical ancillary (VANC) capability is listed in row 2 of the table and indicates the bit-rate available for insertion of ancillary data in the vertical blanking interval. The total ancillary data space, listed in row 3 of the table, represents the sum of the HANC and VANC capability of the system. This value may be reduced by 10% to 20% by the data formatting used. Row 3 of the table lists the nominal (total) bit-rate.
The essential bit-rate required by the standard is shown in row 4 of the table. It results from the elimination of non-essential samples in the horizontal and vertical blanking intervals. Ancillary data may include digital audio, time code, EDH as well as user and

control data.

TABLE 1 4:2:2 525/50 ANCILLARY DATA SPACE

Horizontal ancillary data space (HANC)	268 Words/line x 525 lines/frame = 140,700 Words/frame 140,700 Words/frame x 29.97 frames/s = 4.216779 MWords/s 4.216779 MWords/s x 10 bits/Word = 42.16779 Mbps
Vertical ancillary data space (VANC)	1440 Words/line x 38 vertical interval lines = 54,720 Words/frame 54,720 Words/frame x 29.97 frames/s = 1.6399584 MWords/s 1.6399584 MWords/s x 8 bits/Word = 13.1196672 Mbps
Total ancillary data space	42.16779 Mbps (HANC) + 13.1196672 Mbps (VANC) = 55.3 Mbps Data formatting and exclusions may reduce this value by 10% to 20%
Total bit-rate	1716 Words/total line x 525 lines/frame x 29.97 frames/s x 10 bits/Word = 270 Mbps
Essential bit-rate	270 Mbps - 55.3 Mbps = 214.7 Mbps

AUDIO DATA MULTIPLEXING

The most important use of the ancillary data space is for the insertion of audio signals accompanying the video signal. The 4:2:2 component digital signals have a considerable amount of overhead. They can easily accommodate eight AES/EBU signals (eight stereo pairs or sixteen individual audio channels) still leaving a considerable amount of overhead for other uses.

The ANSI/SMPTE 272M document defines the manner in which AES/EBU digital audio data, AES/EBU auxiliary data and associated control information is embedded into the ancillary data space of the bit-serial digital video conforming to the ANSI/SMPTE 259M standard. The 4:2:2 525/60 component digital signal can accommodate 268 ancillary data words in the unused data space between the end of active video (EAV) timing reference and start of active video (SAV) timing reference. Figure 2 shows the ancillary data packet structure for the 4:2:2 component digital interface. Each packet can carry a maximum of 262 10-bit parallel words. A 6-word header precedes the ancillary data and contains:

• A three-word ancillary data flag (ADF) marking the beginning of the ancillary data packet. Their values are 000, 3FF, 3FF respectively.

• An optional data identification (DID) word identifying the user data.

• An optional data block number (DBN) word

• A data count (DC) word

A variable number of data words, not exceeding 255, follows. The packet is closed by a checksum (CS) word allowing the receiver to determine the validity of the packet. Multiple, contiguous, ancillary data packets may be inserted in any ancillary data space. They must follow immediately after the EAV, for the HANC, or the SAV, for the VANC, to indicate the presence of the auxiliary data and the start of a packet. If there is no ADF in the first three words of an ancillary data space it is assumed that no ancillary data packets are present.

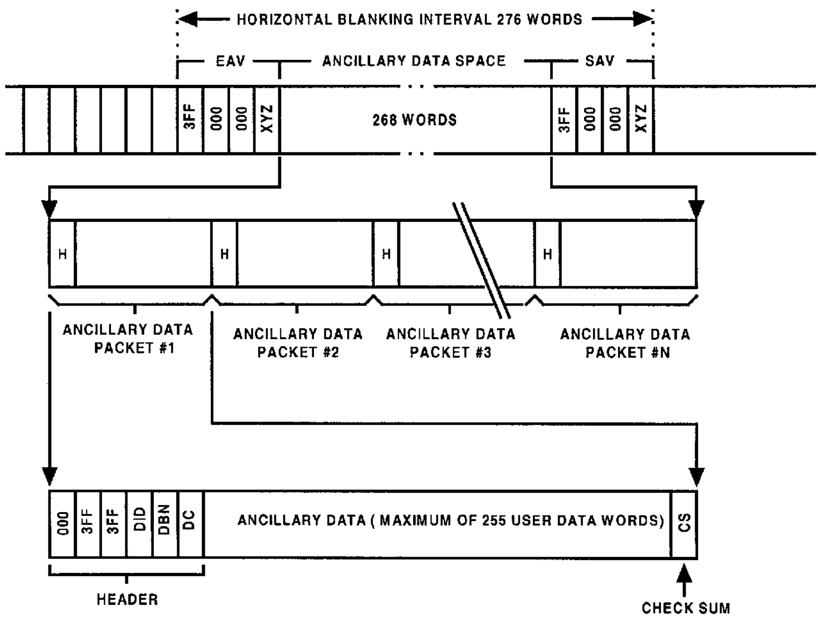

FIGURE 2 ANCILLARY DATA PACKET STRUCTURE FOR 4:2:2 525/60 FORMAT

The SMPTE Standard 272M proposes two modes of operation for embedding digital audio into a video digital data stream. The minimum implementation is characterized by 20-bit resolution, 48 kHz sampling, audio synchronous with video, only one group of four audio channels and a receiver buffer size of 48 audio samples. The full implementation is characterized by 24-bit resolution, sampling frequencies of 32 kHz, 44.1 kHz or 48 kHz, audio synchronous or asynchronous with video, up to four groups of four audio channels, a receiver buffer size of 64 audio samples and indication of relative time delay between any audio channel and the video data signal.

Figure 3 shows an example of the minimum implementation in which two data streams (AES/EBU data stream 1 and AES/EBU data stream 2) are formatted for embedding into a 4:2:2 525/60 component digital signal.

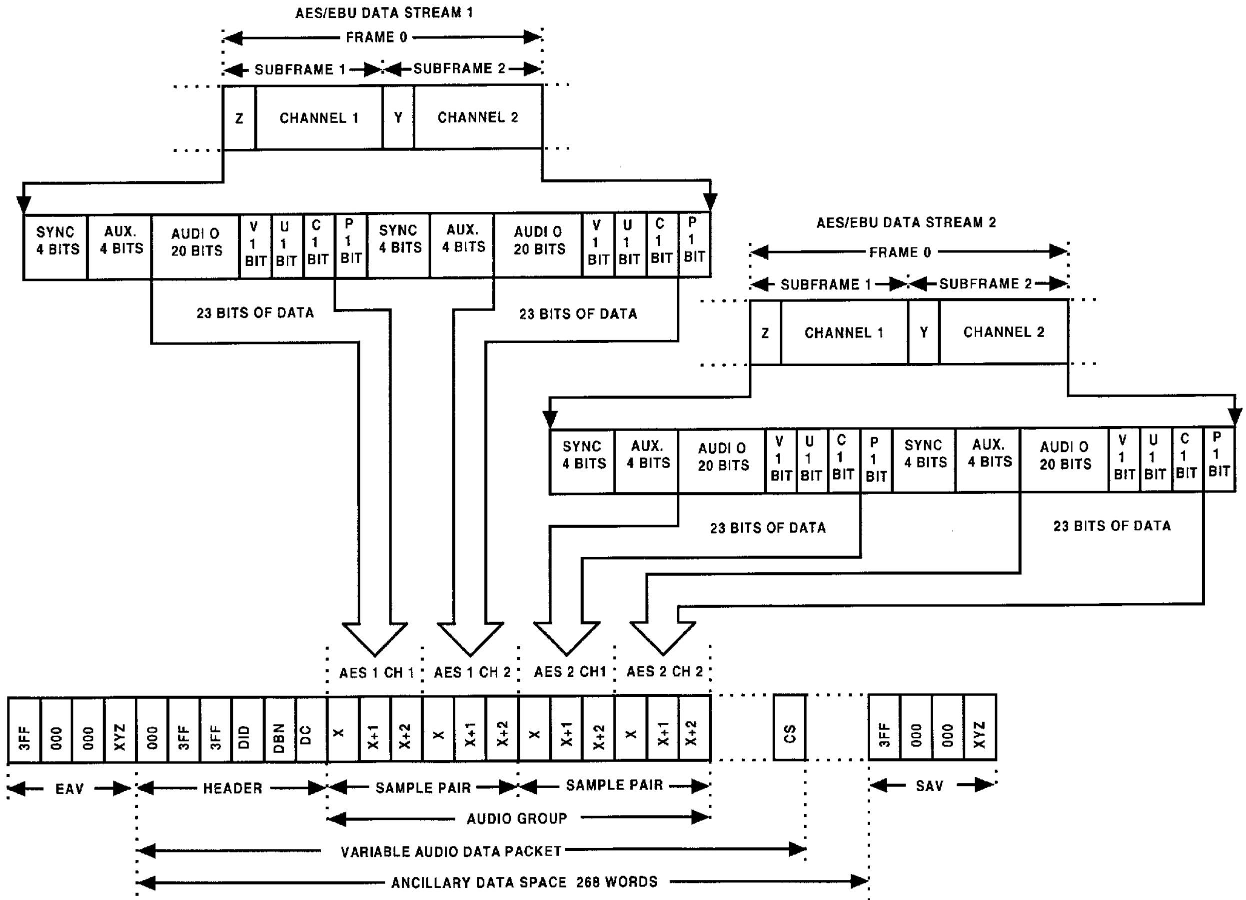

FIGURE 3 AUDIO DATA PACKET FORMATTING FROM TWO AES/EBU DATA STREAMS

•A six-word header starts the audio data packet.

•To begin the embedding sequence, Frame 0 of AES/EBU data stream 1 provides data from its Subframe 1 and Subframe 2. Each of these subframes is stripped of the four Sync bits, the four Auxiliary bits and the P bit. The remaining 20 bits of audio and the V,U and C bits, a total of 23 bits of Subframe 1 are mapped into three consecutive 10-bit words identified as X, X+1 and X+2 of AES1/CH1.

•The 23 bits of Subframe 2 are similarly mapped into three consecutive 10-bit words identified as X,X+1 and X+2 of AES1/CH2.

•AES1/CH1 and AES1/CH2 form a sample pair.

• To continue the embedding sequence, Frame 0 of AES.EBU data stream 2 provides data from its Subframe 1 and Subframe 2. These data are similarly reduced to 23 bits and result in Sample Pair AES2/CH1 and Sample Pair AES2/CH2.

•A number of consecutive Sample Pairs form an Audio Group. The number of sample pairs per Audio Group varies from line to line.

•The audio data packet closes with a CS word.

•Subsequent sample pairs will accommodate Frame1 of AES/EBU data stream 1 and data stream 2, Frame 2 of AES/EBU data stream 1 and data stream 2 and so on until the 192 frames (each constituting one AES/EBU block) of each of the two AES/EBU data streams are embedded.

•From then on a new block of 192 frames coming from the two AES/EBU data streams will be embedded and the process will continue.

•At the receiving end the packets are extracted and fill a 64-sample buffer from which the original data are extracted at a constant bit-rate and reformatted.

Table 2 shows the audio data structure represented by the three 10-bit data words. The channel number is indicated by two bits and a parity is calculated on the 26 bits, excluding all b9 address bits. Figure 4 shows a simplified block diagram of an audio multiplexer and Figure 5 shows a simpliified block diagram of an audio demultiplexer.

TABLE 2 FORMATTED AUDIO DATA STRUCTURE

BIT ADDRESS	WORD X	WORD X+1	WORD X+2
b9	not b8	not b8	not b8
b8	audio 5	audio 14	P
b7	audio 4	audio 13	C
b6	audio 3	audio 12	U
b5	audio 2	audio 11	V
b4	audio 1	audio 10	audio 19 (MSB)
b3	audio 0 (LSB)	audio 9	audio 18
b2	channel 1	audio 8	audio 17
b1	channel 0	audio 7	audio 16
b0	Z	audio 6	audio 15

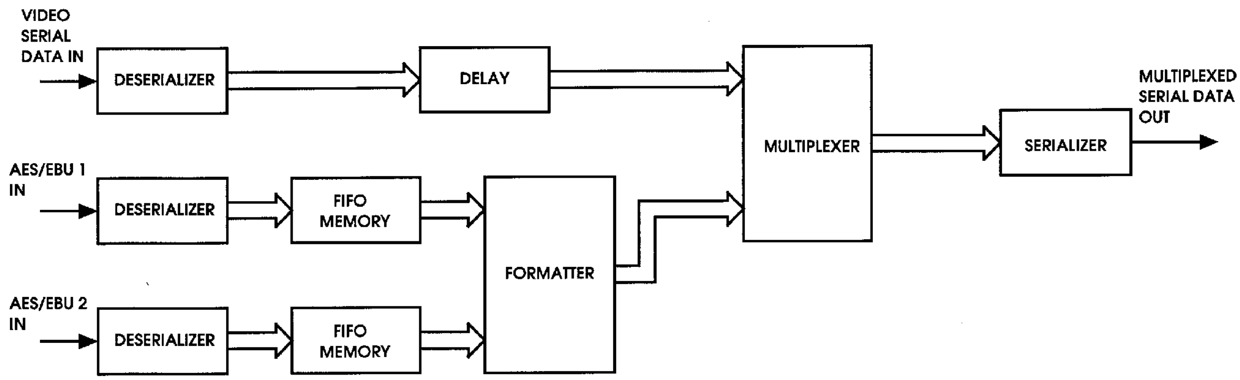

FIGURE 4 SIMPLIFIED BLOCK DIAGRAM OF AUDIO MULTIPLEXER

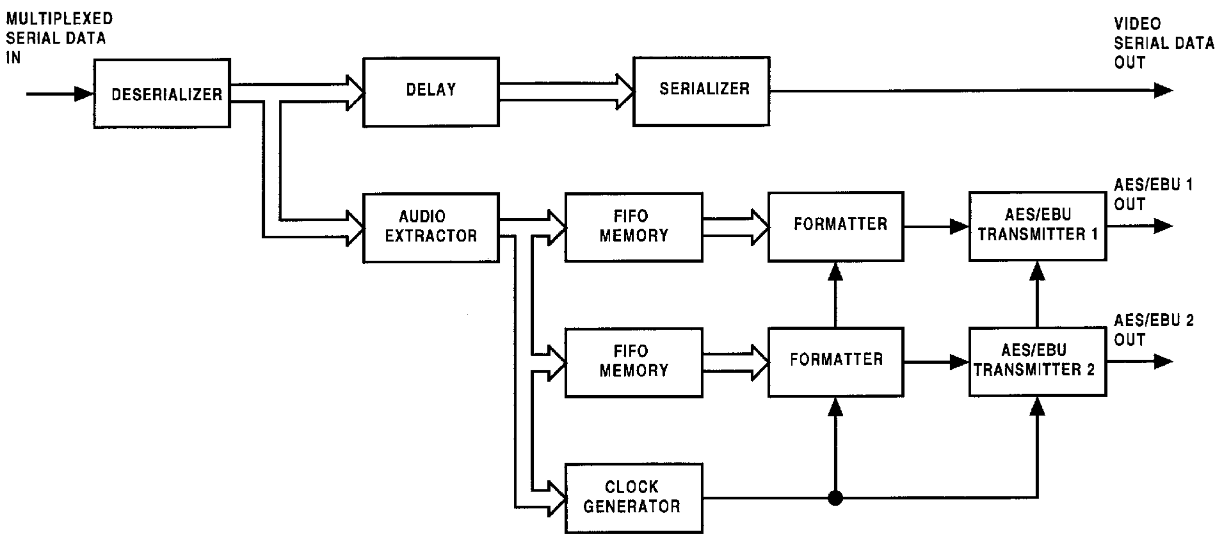

FIGURE 5 SIMPLIFIED BLOCK DIAGRAM OF AUDIO DEMULTIPLEXER