where the factor of 4/5 accounts for the difference in aspect ratios, 52.66 uS is the NTSC active line time, and the factor of 2 is needed to convert from cycles to lines.

The video encoding employs a scheme whereby low frequency horizontal and vertical luminance components are transmitted at 59.94 frames per second (fps) while the higher frequency luminance components and color difference components are sent at one-fifth the rate, or 11.988 fps. To conserve bandwidth, the highest frequency diagonal resolution components are not sent.

3.2 VIDEO SOURCE SIGNAL

3.2.1 TIMING AND BANDWIDTH

The video source required for encoding is a 787.5 line, 59.94 fps progressive signal. This can be expressed as a 787.5/59.94/59.94 source. The horizontal scan rate of the source signal is exactly 3 times NTSC (47.2 kHz). Note that scan rates are intentionally related to NTSC to accommodate methods which minimize cochannel interference. Of the 787.5 lines per frame, 720 are active lines with 17.65 uS (5/6) active line time, this leaves 3.53 uS (1/6) for retrace.

In order to support the horizontal resolution of 1020 lpw, the minimum horizontal bandwidth of the source signal is:

$$1020/2/17.65 \text{ uS} = 28.9 \text{ MHz}$$

Although the 787.5/59.94/59.94 signal is the final input to this system, a very desirable method of deriving this input would be to start with a higher line rate source, such as a 1050 line 59.94 Hz per frame progressively scanned

source, which has been vertically filtered to eliminate vertical aliasing and then scan converted to 787.5/59.94/59.94.

It is also possible to derive the 787.5/59.94/59.94 signal by scan converting a 1050 line 59.94 Hz per frame 2:1 interlace source. Although this is presented as a possibility, motion artifacts result from the conversion process.

3.2.2 COLORIMETRY

The system will utilize the SMPTE Step II color processing (currently under development in committee as an extension of SMPTE 240M). This will allow direct use of component signals from sources meeting the proposed SMPTE Step II standard without transcoding. Signals using the existing Step I (SMPTE "C" phosphors) standard are transcoded before transmission.

Utilization of the Step II process reduces visibility of chroma noise, improves resolution in saturated colors, and makes the colorimetry independent of the receiver.

The processing as presently envisioned is as follows:

Taking primaries are:

color	x	y	comment
red (R)	.67	.33	NTSC red
green (G)	.21	.71	NTSC green
blue (B)	.15	.06	EBU (PAL/SECAM) blue

The camera is matrixed linearly to correspond to these primaries. Luminance, Y is derived from a linear matrix as well.

Y, R, and B are each compressed by a specified opto-electrical transfer function to form Y′, R′, B′; and then matrixed to form "color-difference" signals R′-Y′ and B′-Y′.

At the receiver, the inverse matrix is applied to obtain Y′, R′, and B′, which are de-compressed to Y, R, and B. These are linear signals which are rematrixed to form the RGB primary signals required by the display. The display primary RGB signals are subsequently gamma corrected before being applied to a CRT display.

R′-Y′ and B′-Y′ are amplitude scaled to form P_R and P_B for transmission according to SMPTE 240M. This allows undistorted transmission of the specified taking gamut up to 100% level, 100% saturated colors. Transmitter compression and receiver de-compression may be non-complementary beyond this point. This allows accurate transmission of surface colors, with minimum color noise.

3.3 VIDEO ENCODING

In order to take full advantage of spectrum compatibility by creating a transmitted signal which can exist in an NTSC Taboo environment, the video encoding must convert the HDTV source signal into a signal that:

- is transmittable as suppressed carrier modulation on two quadrature carriers at the center of the channel, without subcarriers, and

- has NTSC related timing so that spectral components may be interleaved with existing NTSC cochannels

To accomplish this, the video encoding converts the HDTV source signal into 480 analog components per 1/59.94 seconds. Each component occupies 63.56 uS in time and has a nominal bandwidth of 2.675 MHz. The 480 components are paired into 240 pairs which are transmitted sequentially as suppressed carrier amplitude modulation (AM) on two quadrature carriers in the center of the R.F. channel. The modulation process is discussed in further detail in Section 4.

In addition to the 240 pairs of analog components, which occupy the equivalent of 240 NTSC horizontal lines, a block of data and synchronizing signals are also sent each 1/59.94 seconds. This block of data and synchronizing signals make up the vertical blanking interval (VBI) and separate the analog components between successive fields. The data and synchronizing signals will occupy the equivalent of 22 and 23 NTSC horizontal lines on successive frames which is approximately the same as the vertical interval in an NTSC signal.

The basic principle used for video encoding is to divide the 2-dimensional (2D) spectrum of the source signal into components and to send only 480 components which are a good match to the human visual system. The encoding scheme transmits a high temporal rate (59.94 fps) low-frequency component for fast motion rendition along with a slower temporal rate (11.988 fps) high-frequency component.

In order to perform the video encoding described below, the RGB signals are converted into a luminance and two color

difference components. The encoding for the luminance and color difference signals will be described separately.

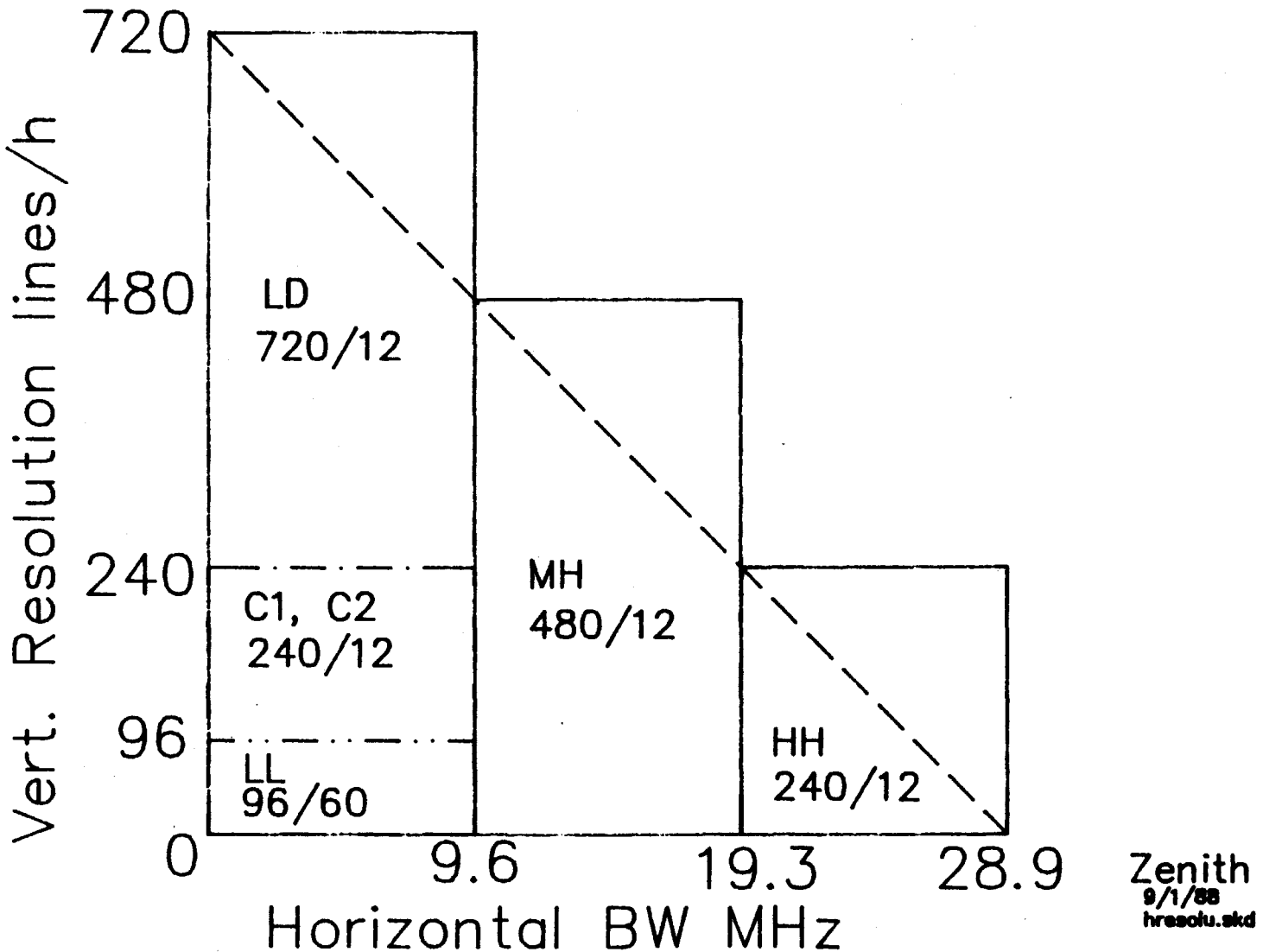
3.3.1 LUMINANCE ENCODING

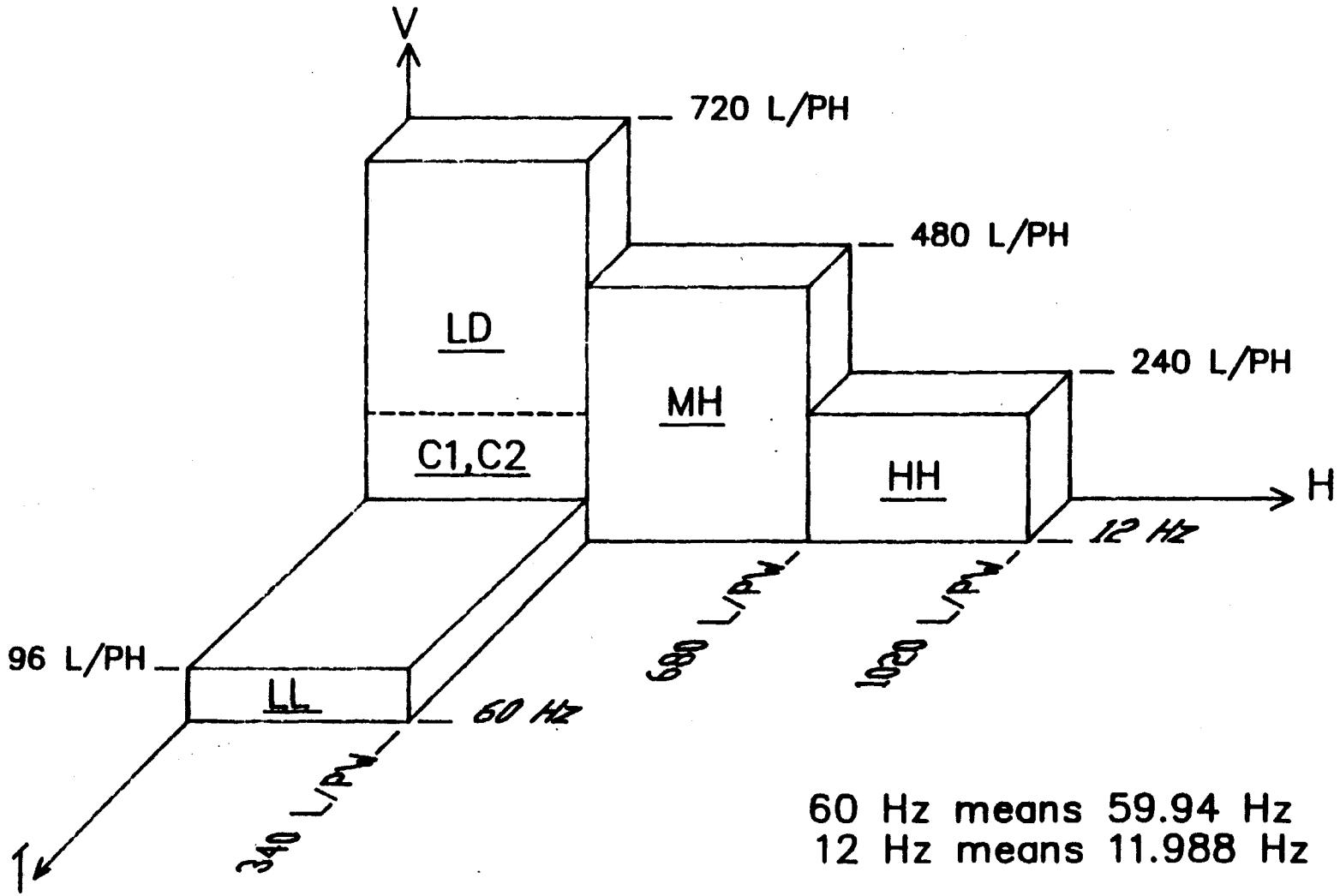
The luminance encoding process divides the HDTV luminance signal into four components which represent different portions of the 2D frequency spectrum. These four components are shown in Fig. 3-2 as LL, LD, MH, and HH. The LL and LD components together represent the lower third (0-9.63 MHz) of the horizontal frequencies of the source image with full 720 line vertical resolution. The MH component represents the middle third (9.63-19.27 MHz) of the horizontal frequencies with only 480 line vertical resolution and the HH component represents the upper third (19.27-28.9 MHz) of the horizontal frequencies with 240 line vertical resolution. The encoding process sends the LL components at 59.94 fps while all the other components are sent time multiplexed for a final rate of 11.988 fps. Fig. 3-3 is a diagram showing the spatial-temporal resolution of each component. Fig. 3-3 also shows the spatial-temporal resolution of the color difference components C1 and C2 which will be discussed later.

A block diagram of the encoding process is shown in Fig. 3-4. The HDTV source is first separated into three 9.63 MHz bands using three filters, each covering one third the horizontal frequency band. Each frequency band is vertically filtered and vertically resampled to the different vertical resolutions for transmission.

The 0-9.63 MHz horizontal frequencies are vertically filtered and resampled to 96 lines which are sent every

Figure 3–2 HDTV Resolution



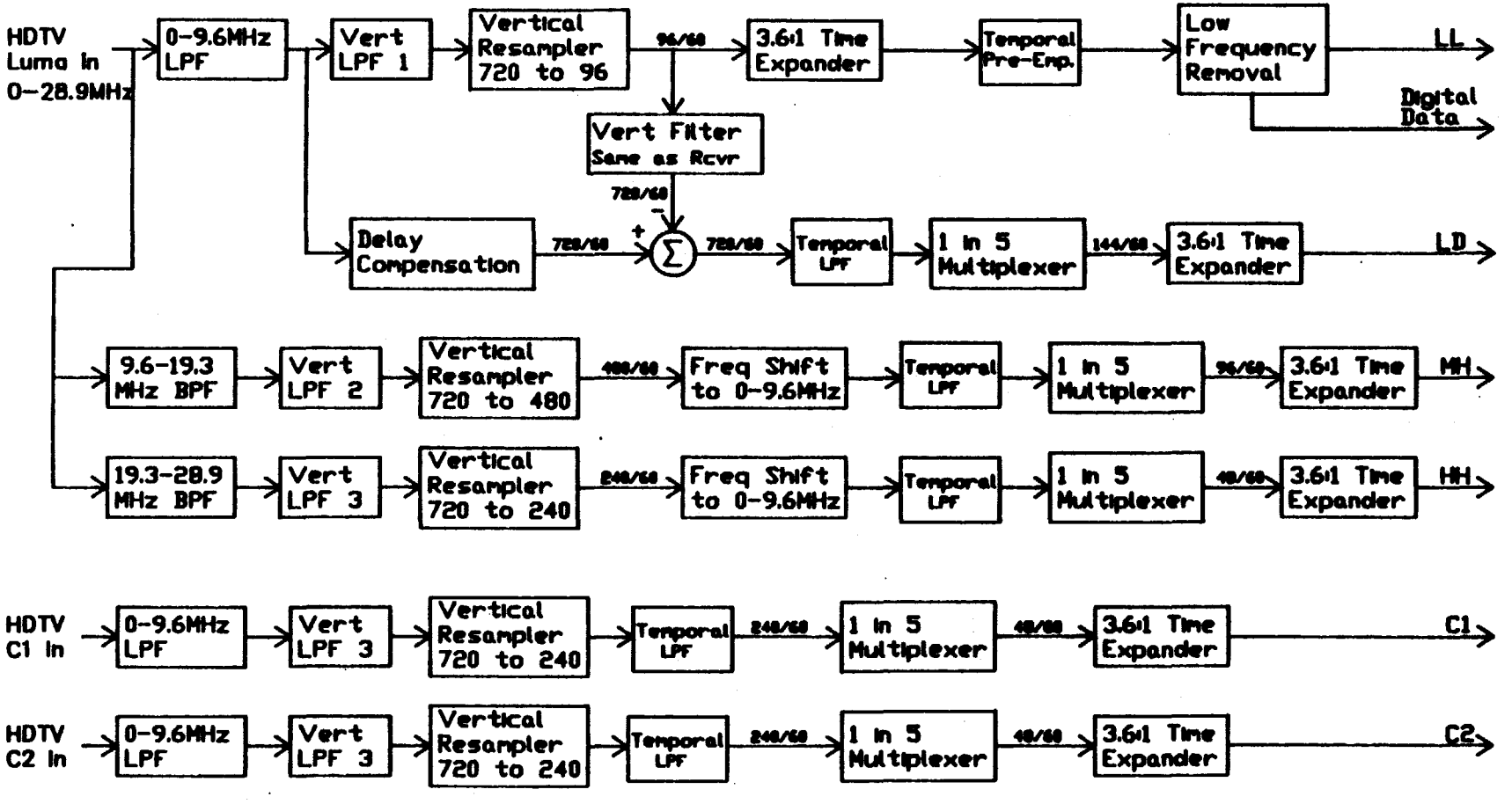


Spatial–Temporal Resolution

Figure 3–3

Zenith
9/1/88
3d_res2

HDTV Video Encoding (Part 1)



Zenith
9/1/88
enc_blk

Figure 3–4 (Part 1)

HDTV Video Encoding (Part 2)

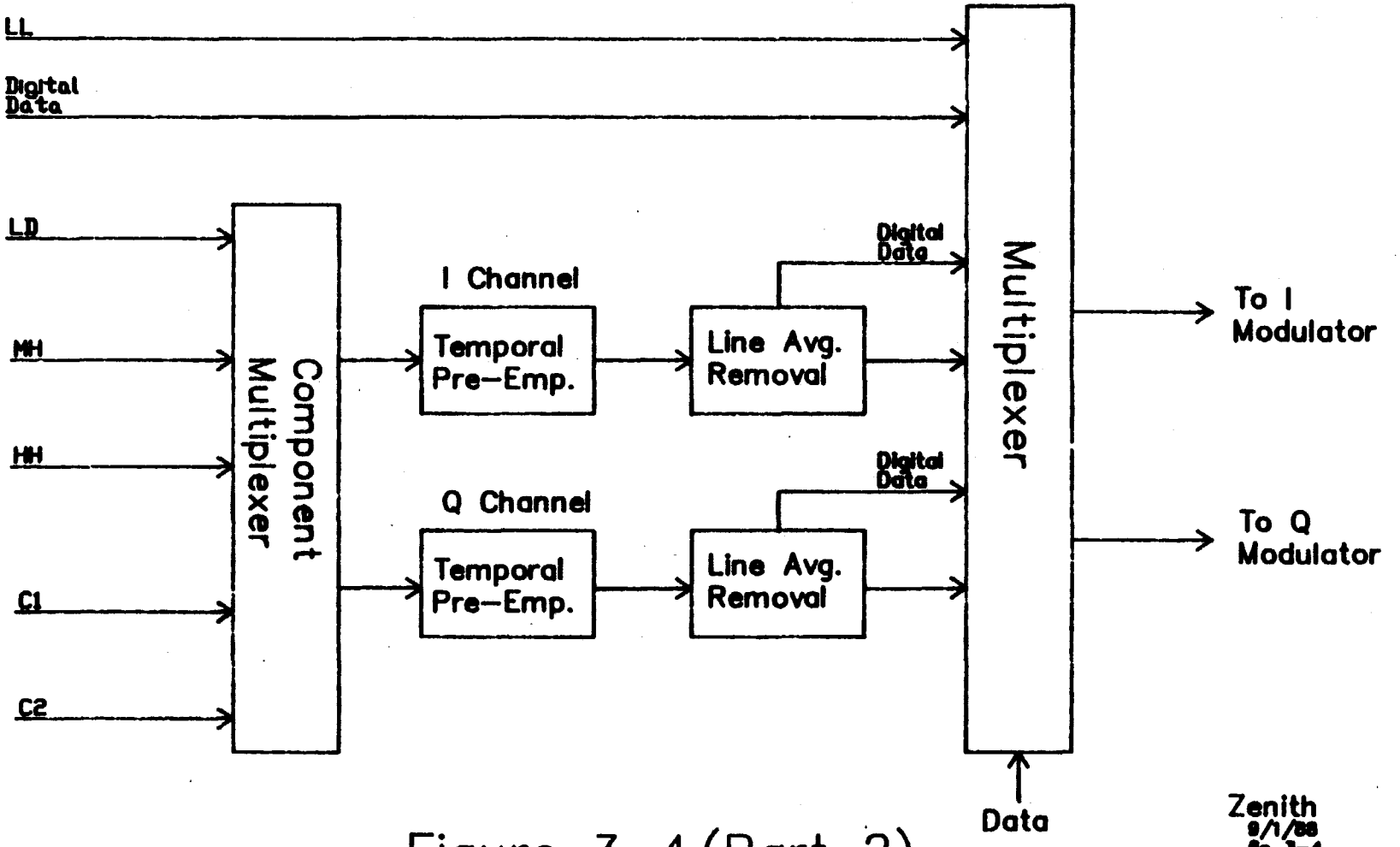


Figure 3-4 (Part 2)

1/59.94 seconds. This component, which represents the low horizontal and low vertical (LL) luminance information, carries the bulk of the motion information which can be seen by the human eye, and therefore is sent at a fast update rate. At the receiver, 720 active lines of luminance information are reconstructed from the 96 LL components by means of a vertical interpolation filter. The rest of the vertical detail in the 0-9.63 MHz low horizontal band is sent as difference lines (LD). The encoder generates these LD lines by subtracting the 720 lines that the receiver would reconstruct from the 96 LL components from the original low pass filtered source lines. Although there are 720 difference lines every 1/59.94 seconds, only 144 (1/5) of these are sent in 1/59.94 seconds. It takes 1/11.988 seconds to send all 720 difference lines. A temporal low pass filter (LPF) is used to reduce aliasing components of the LD components during motion.

To conserve bandwidth, only the active portion of the video is transmitted with no "blanking interval" between lines. The active portion of LL and LD are time expanded by 3.6:1 which reduces their bandwidth to 9.63/3.6 = 2.675 MHz and increases their duration to 63.56 uS. After time expansion, Fig. 3-4 shows that temporal pre-emphasis is applied to LL. Temporal pre-emphasis is not needed for encoding the video but instead is used for reduction of interference and will be discussed in more detail in Section 4, "Transmission and Modulation Techniques". Before transmission, the low-frequency component (<200 kHz) of each LL line is removed and sent as digital data during the vertical interval. The process of removing the low-frequencies will also be discussed in more detail later.

The middle third of the horizontal frequencies, MH, are resampled to 480 lines of vertical resolution and are also

multiplexed so that only 1/5 of them are sent; 96 of these components are sent every 1/59.94 seconds. A temporal LPF is used to reduce aliasing of the MH components during motion. The result of the vertical filter and vertical resampling is that the static vertical resolution for the mid horizontal frequencies will be 2/3 that of the low horizontal frequencies. For transmission, these components are frequency shifted to 0-9.63 MHz and the active portion of the line is time expanded by 3.6:1.

The highest horizontal frequencies, HH, which are resampled to 240 lines of vertical resolution, are also temporally LPF'ed and sent 1/5 as often. There are 48 of these components per 1/59.94 seconds. These horizontal frequencies only have 1/3 the static vertical resolution compared to the low horizontal frequencies. These components are also frequency shifted to 0-9.63 MHz and the active portion is time expanded by 3.6:1 before transmission.

As shown in Fig. 3-2, the net effect of the different vertical resolutions transmitted for each horizontal frequency band is a reduction in diagonal resolution. The dashed diagonal line in Fig. 3-2 has been added for reference and would represent the resolution obtained from diagonal sampling.

The four luminance components of the encoded HDTV signal described above each have a bandwidth of 2.675 MHz after time expansion and are transmitted in pairs as suppressed carrier amplitude modulation on two quadrature carriers in the R.F. channel. In the luminance processing described above, 384 luminance components are sent every 1/59.94 seconds. With each component occupying 63.56 uS, each transmitted pair of luminance components will take up

the same time as one NTSC line. Therefore, a total of 192 NTSC lines are used to send luminance information every 1/59.94 seconds. The remaining lines are used to send data and color difference information.

To generate the two signals which modulate the inphase (I) and quadrature (Q) RF carriers, a component multiplexer is used. The component multiplexer time-multiplexes the LD, MH, HH, and color difference components into two signals which after further processing are multiplexed with the LL components for modulation of the I and Q RF carriers. The two signals from the multiplexer are first each temporally pre-emphasized and then have the line average of each component (LD, MH, HH, C1, C2) removed and sent digitally as data during the VBI. The line average removal process will be explained later.

3.3.2 COLOR DIFFERENCE ENCODING

Figs. 3-2, 3-3, and 3-4 also show the color difference components which are processed with techniques similar to the luminance components. The two color difference components, C1 and C2, each bandlimited to 9.63 MHz, are vertically filtered and vertically resampled to provide 240 lines of resolution for each. As with the luminance, these components are then temporally low pass filtered, multiplexed down to 1 in 5, and the active portion is time expanded by 3.6:1.

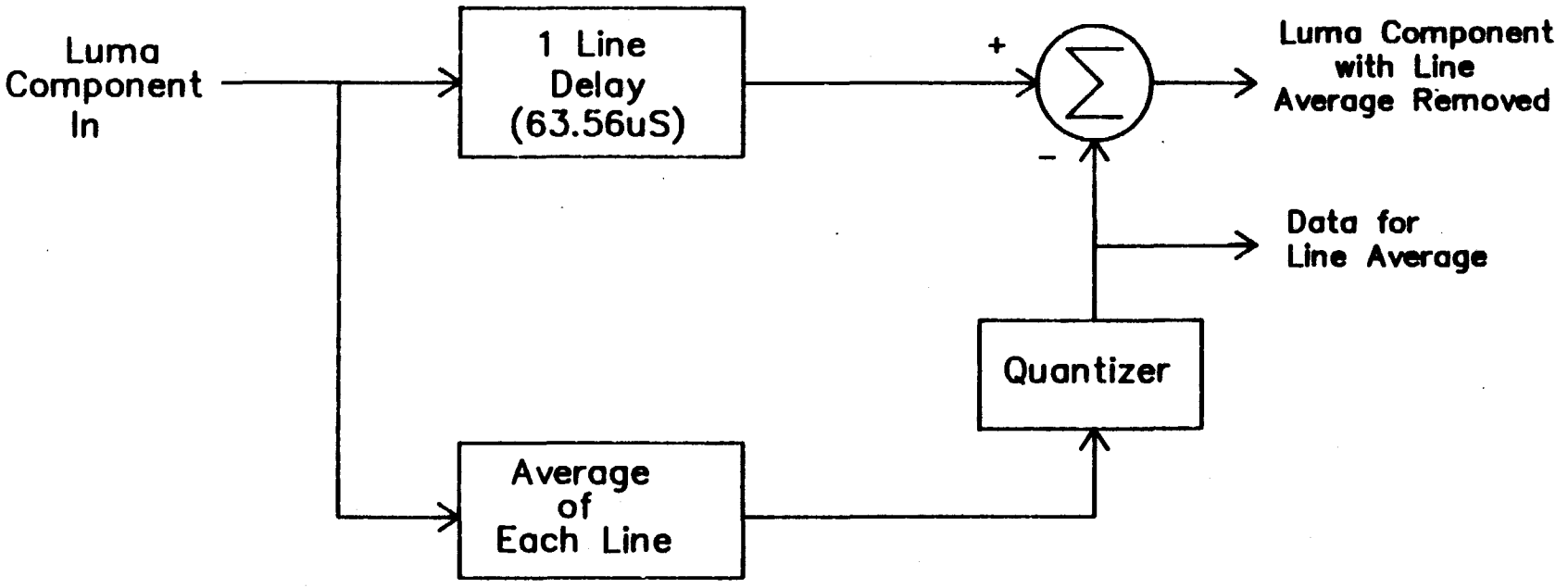
A total of 96 color difference components are sent as pairs every 1/59.94 seconds. The 96 color difference components occupy the same time as 48 lines of NTSC information. For display at the receiver, the 240 color difference components will be interpolated for all 720 active lines to give a smooth color presentation.

The HDTV color difference resolution will be one third the horizontal and vertical resolution of the HDTV luminance. In comparison, in the horizontal direction, this is 4.48 times the NTSC chroma resolution and in the vertical direction, this is .5 times. The vertical resolution of the HDTV color difference signal is the same as current NTSC receivers which use a one-line comb filter. The color resolution of the final HDTV display will be balanced in the horizontal and vertical directions and is satisfactory for HDTV.

3.3.3 REMOVAL AND DIGITIZATION OF LOW-FREQUENCIES

As mentioned above, the low-frequencies are removed from the LL components and the line average is removed from all the other analog components before transmission. The information removed is digitized and transmitted as data during the VBI. This process produces two important results. First, there will be no d.c. information transmitted in any of the analog components, and second, a large power reduction is realized. Both of these results are important for minimizing interference into existing services and other HDTV transmissions.

Fig. 3-5 shows a block diagram of how the line average is removed from each component. First, the line average is calculated by integrating over the duration of each component. Then the result is quantized and subtracted from the analog component which has been delayed to match the integration time. The result of the integration is quantized to the same number of bits that is transmitted in the VBI so that the receiver decoder is exactly complementary to the encoder.



Block Diagram of Line Average Removal

Figure 3–5

Zenith
9/1/88

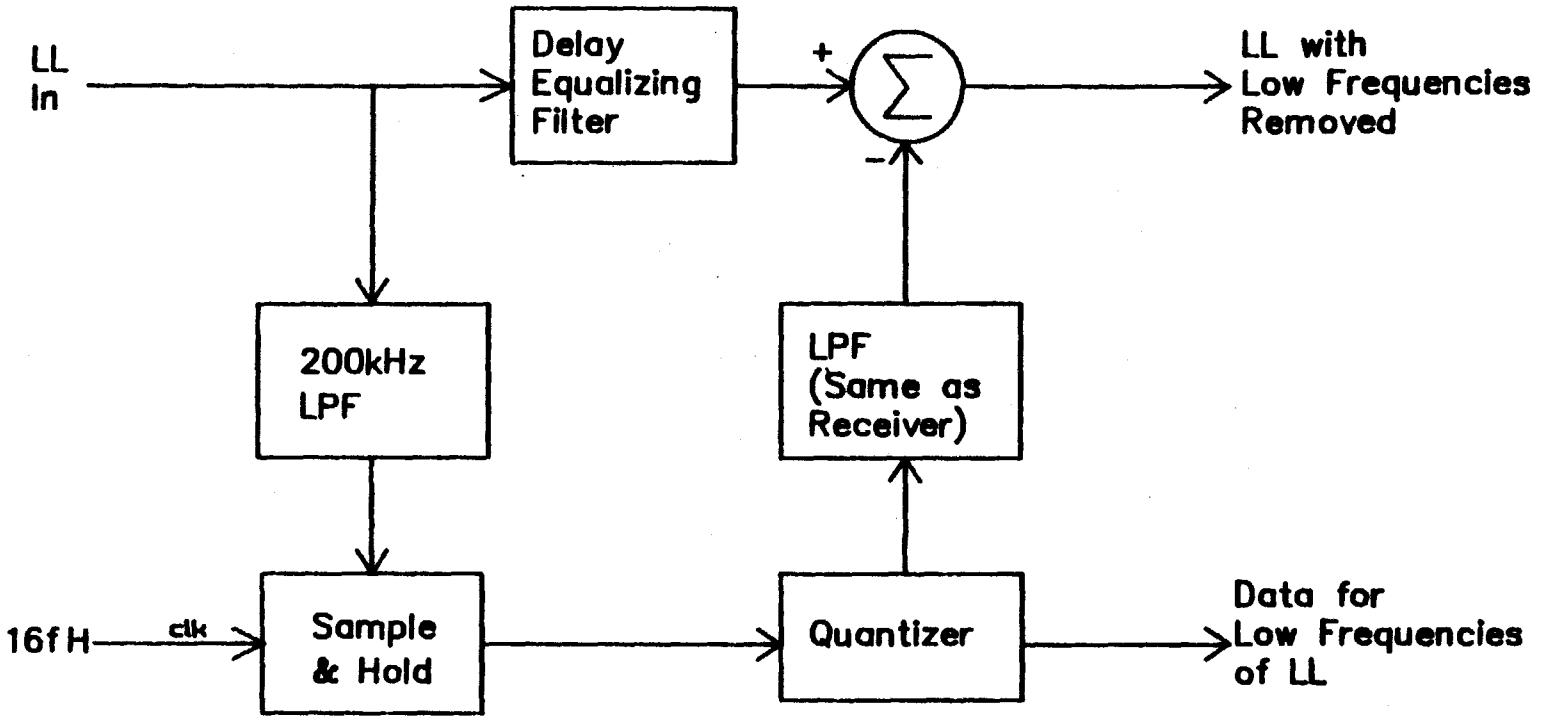
fig_3–5

In order to further assure correct restoration of the line average, a reference level is sent during the VBI which the receiver uses for automatic gain equalization (AGE). The receiver uses the reference level to establish gain constants between the analog signal and the digital data.

For the LL components, the low-frequencies are removed in two steps. First, the line average is removed as just described, and then the low-frequencies, below 200 kHz are removed by the method shown in Fig. 3-6. As shown in Fig. 3-6, each LL component, after time expansion, is low-pass filtered to 200 kHz and then resampled into 16 horizontal samples. These samples are quantized and interpolated in a LPF to produce a smooth representation of the low-frequencies which are then subtracted from the LL components. Since the quantization corresponds to the number of bits which are transmitted in the VBI and the LPF is an exact duplicate of what is used in the receiver, the receiver will be able to restore low-frequencies of the LL components with no error. Note that since the processing uses subtraction for the low-frequency removal, aliasing introduced by using only 16 horizontal samples will be canceled exactly at the receiver. In order to maintain matching between the D/A converted digital signal and analog signal levels, the receiver will use the same reference level in the VBI mentioned above for gain equalization.

3.3.4 MOTION RENDITION

The end result of the encoding is that low horizontal and vertical frequency luminance components, LL, are updated every 1/59.94 seconds and provide very good motion reproduction where the human eye is most sensitive to movement. For higher frequency components and color, where



Block Diagram of Low Frequency Removal

Figure 3–6

Zenith
9/1/88
fig_3–6

the eye is less sensitive to movement, intermediate updates occur every 1/59.94 seconds but it takes 1/11.988 second for a complete update.

Although the fast update rate of the lowest frequency luminance components, LL, contribute to good motion rendition, motion compensation is applied to all the other terms. Motion compensation consists of both encoder and receiver techniques. Encoder motion compensation is achieved by predicting the motion related errors which the receiver makes and minimizing the errors by modifying the transmitted components. Receiver motion compensation is aided by data sent during the vertical interval which indicates on a line-by-line basis, which lines require compensation. Within each line indicated by the encoder, the receiver determines which pixels actually need correction. The receiver compares the incoming line information with previous lines which are stored in a frame memory in the receiver and only allows motion compensation at places in a line where significant differences are found in the surrounding areas. Static areas are not affected by the correction.

3.3.5 COMPONENT MULTIPLEXING

There are a total of 480, 63.56 uS, 2.675 MHz luminance and color difference components which must be transmitted every 1/59.94 seconds. When transmitted in pairs (on quadrature carriers), they take up the equivalent of 240 NTSC lines which leave 22 or 23 NTSC lines every 1/59.94 seconds for data. As indicated above, the data is transmitted during the vertical interval which may be optionally frame-locked to an interfering channel to minimize interference into that channel. The luminance and color difference components are transmitted during the

normal NTSC active time and are paired and sequenced for transmission. Pairing is chosen to minimize sensitivity to any crosstalk which might occur between the quadrature channels under non-ideal transmission conditions. The following 5 line pairing sequence is used and is repeated 48 times per 1/59.94 seconds to achieve a total of 240 lines:

Line #	I Channel	Q Channel
1	LL	LL
2	LD	LD
3	LD	MH
4	MH	HH
5	C1	C2

3.4 DIGITAL AUDIO AND DATA

The VBI data occupies 22 or 23 NTSC lines. This data consists of two channels of compressed digital audio, low-frequency video data, synchronization signals, captioning, and error protection. The data is transmitted using 16 state quadrature amplitude modulation (16 QUAM) of the same carriers on which the video is modulated. Fig. 3-7 shows a computer simulated eye diagram of the detected signal on the I and Q channels. The eye openings separating the four levels of each channel, which produce the 16 states, can clearly be seen.

The data clock rate on the I or Q channel is related to the bandwidth of the R.F. channel and is 340 x 15.73427 kHz = 5.34965 MHz. The recovered baseband spectrum of a random data sequence is shown in Fig. 3-8. Fig. 3-8 shows that the data spectrum is flat up to the beginning of the Nyquist slope.

The equivalent of two NTSC lines are dedicated to synchronization signals including the reference level for

EYE DIAGRAM
of an in-phase or quadrature detected
random 16-QUAM pulse sequence in a
6 MHz channel

arbitrary scale units

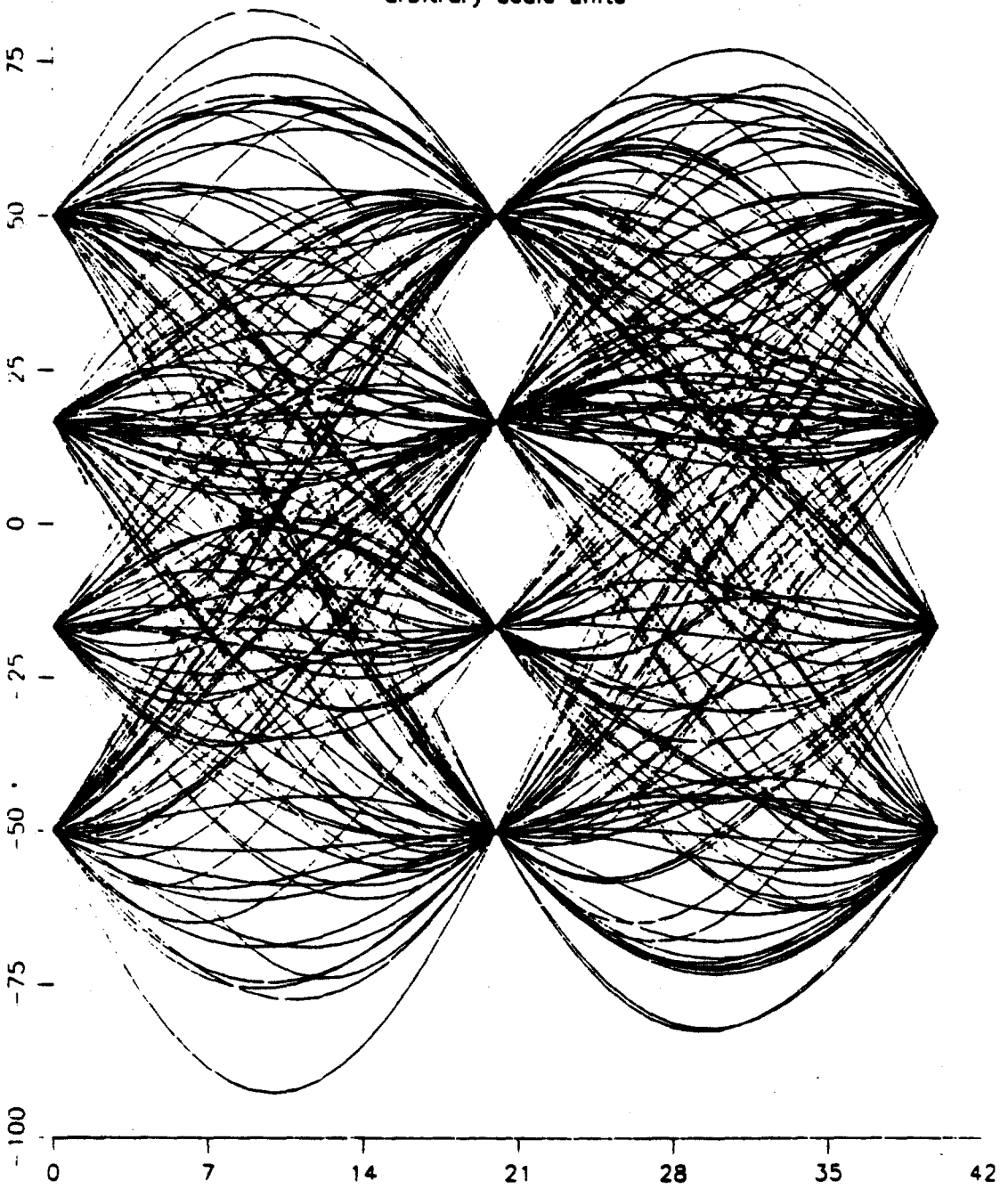


Figure 3-7

Zenith
9/1/88

BASEBAND SPECTRUM
of a random 16 QUAM pulse sequence
Pulse rate 5.35 MHz
Cosine Spectrum roll-off rate 12%

arbitrary amplitude scale

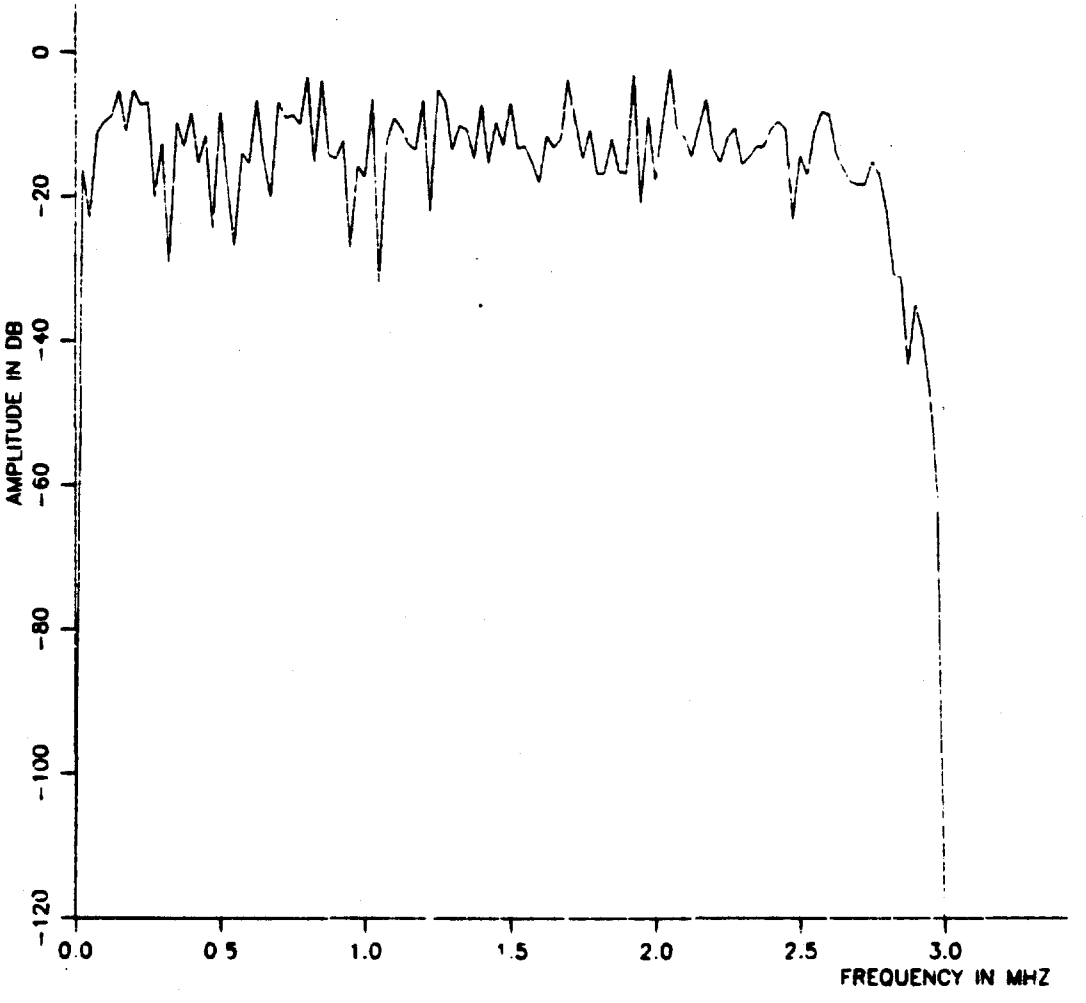


Figure 3-8

Zenith
9/1/88

the AGE mentioned above. At the data clock rate of 5.34965 MHz, the remaining portion of each VBI line allows the following number of bits to be transmitted each 1/59.94 seconds:

340 clks/NTSC line x 4 bits/clk x 20.5 lines = 27,880 bits

Error protection, using interleaved concatenated codes, capable of correcting at least one error in every 10 bits is used. The error protection carries an overhead of 33%. The resultant data rate is then:

27,880 x 3/4 = 20,910 bits per 1/59.94 seconds

The video data required per VBI is summarized in the table below:

Number of Bits	Allocation
768	8 bits per line average for LL
864	6 bits per line average for LD
576	6 bits per line average for MH
288	6 bits per line average for HH
576	6 bits per line average for C1 and C2
288	2 bits per LD for motion compensation
9,216	6 bits per sample for low-frequency LL
12,576	Total video data bits per VBI

This leaves 8,334 bits per VBI which is enough to support two compressed digital audio signals and captioning. The number of bits available per VBI for audio is sufficient to accommodate two channels of either Digideck or Dolby encoded audio.

3.5 VIDEO DECODING

In the HDTV receiver, the 480 analog video components and the data are detected and used to reconstruct a 787.5 line progressively scanned HDTV picture. Fig. 3-9 shows a block diagram of the system used to detect the I and Q modulation and data. The tuner converts the 6 MHz HDTV signal into an I.F. signal where it is gain controlled by variable gain amplifier, A. Carrier regeneration, is accomplished using a Frequency and Phase Locked Loop (FPLL). The FPLL simultaneously provides wide acquisition and narrow bandwidth which are both needed in this system. Carrier synchronization will be discussed further in a Section 4. The FPLL provides two regenerated quadrature carriers which are used to demodulate the modulation on the I and Q carriers.

Since all the processing steps for both luminance and color difference signals are performed digitally, the demodulated I and Q signals are digitized by two A/D converters. In order to establish the correct ratio between the analog and digital D/A converted signals, the amplitudes of the signals being digitized are controlled by the AGE system via variable gain amplifier, A. The block labeled "controller" supplies the AGE signal as well as recovered data.

The analog components modulated on the I and Q signals are detected, digitized, and time-demultiplexed into the four luminance components, LL, LD, MH, HH, and the two color difference components, C1 and C2. The processing of the recovered luminance and color difference components produce 720 active lines of progressively scanned luminance and color difference signals which are rematrixed into RGB signals for display.

Receiver Detector

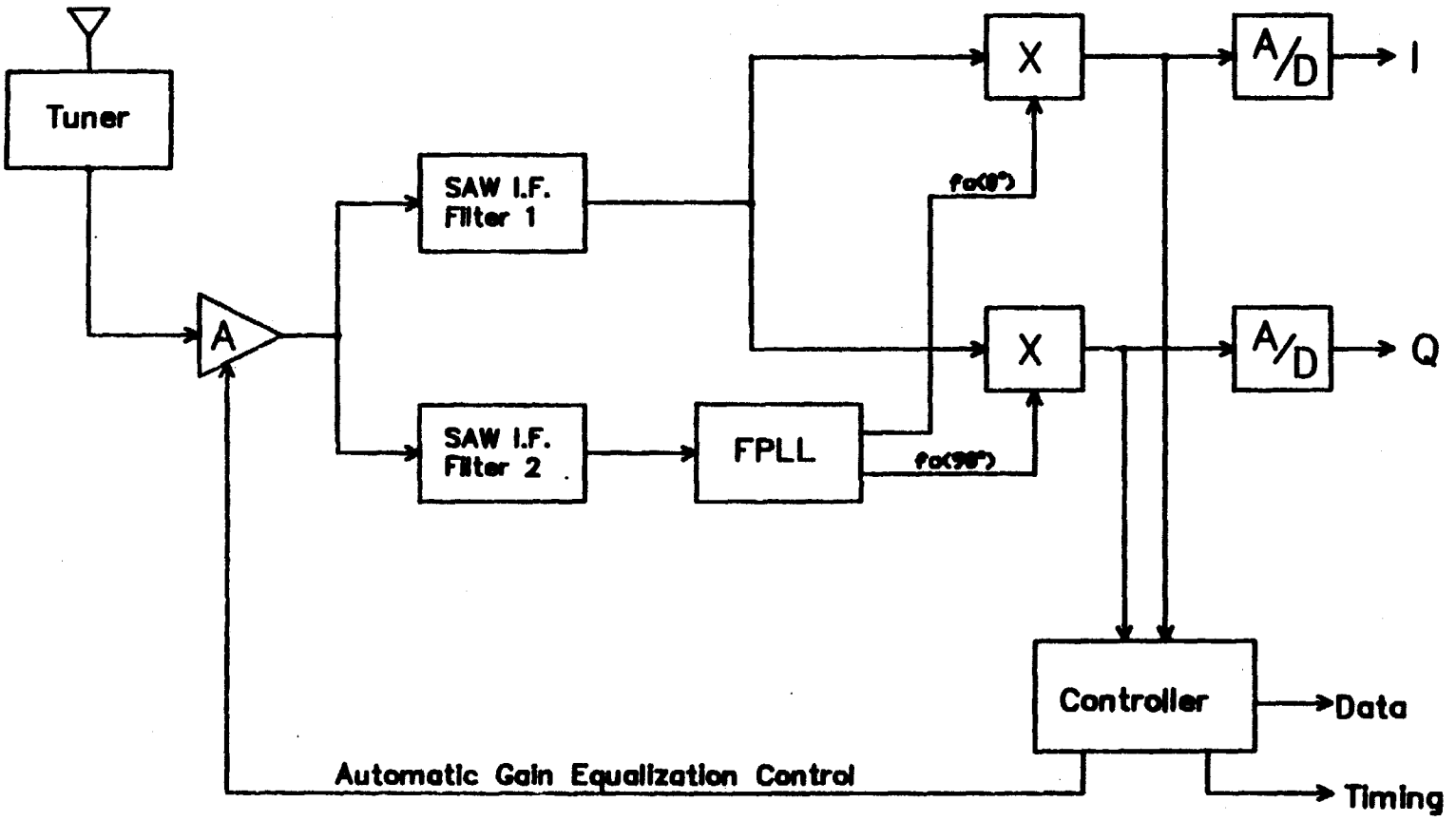


Figure 3–9

Zenith
9/1/88
fig_3-9

3.5.1 LUMINANCE DECODING

A block diagram of the luminance decoding is shown in Fig. 3-10. As shown in Fig. 3-10, each luminance component is decoded by first restoring the digitized low-frequencies. For the LL component, this involves reinserting the low-frequency information as well as the line average. For the other components, this only requires the reinsertion of the line average. By restoring the low-frequency LL with a digital LPF identical to the one used in the transmitter, exact restoration is achieved with no artifacts.

After low-frequency restoration, each component is subjected to temporal de-emphasis which will be explained in detail later. The storage memory required by the temporal de-emphasis, is also used for two other purposes. The first purpose is to store information from previous fields so that it can be repeated for display. The other purpose is to provide storage by which to time compress the signals back to 17.65 uS (active time) for display.

After the temporal de-emphasis, the LL, MH, and HH components still have the wrong number of lines vertically for display. These three components are then interpolated back to 720 active lines per frame using vertical interpolation filters for the sake of display. The components which were sent at 11.988 fps, namely LD, MH, and HH, are all combined together and motion compensated before being recombined with the 720 line LL component. Receiver based motion compensation is used to augment the motion compensation provided at the encoder. The receiver motion compensation is controlled by frame differences in LL combined with motion compensation data transmitted in the VBI.

Luminance Decoding

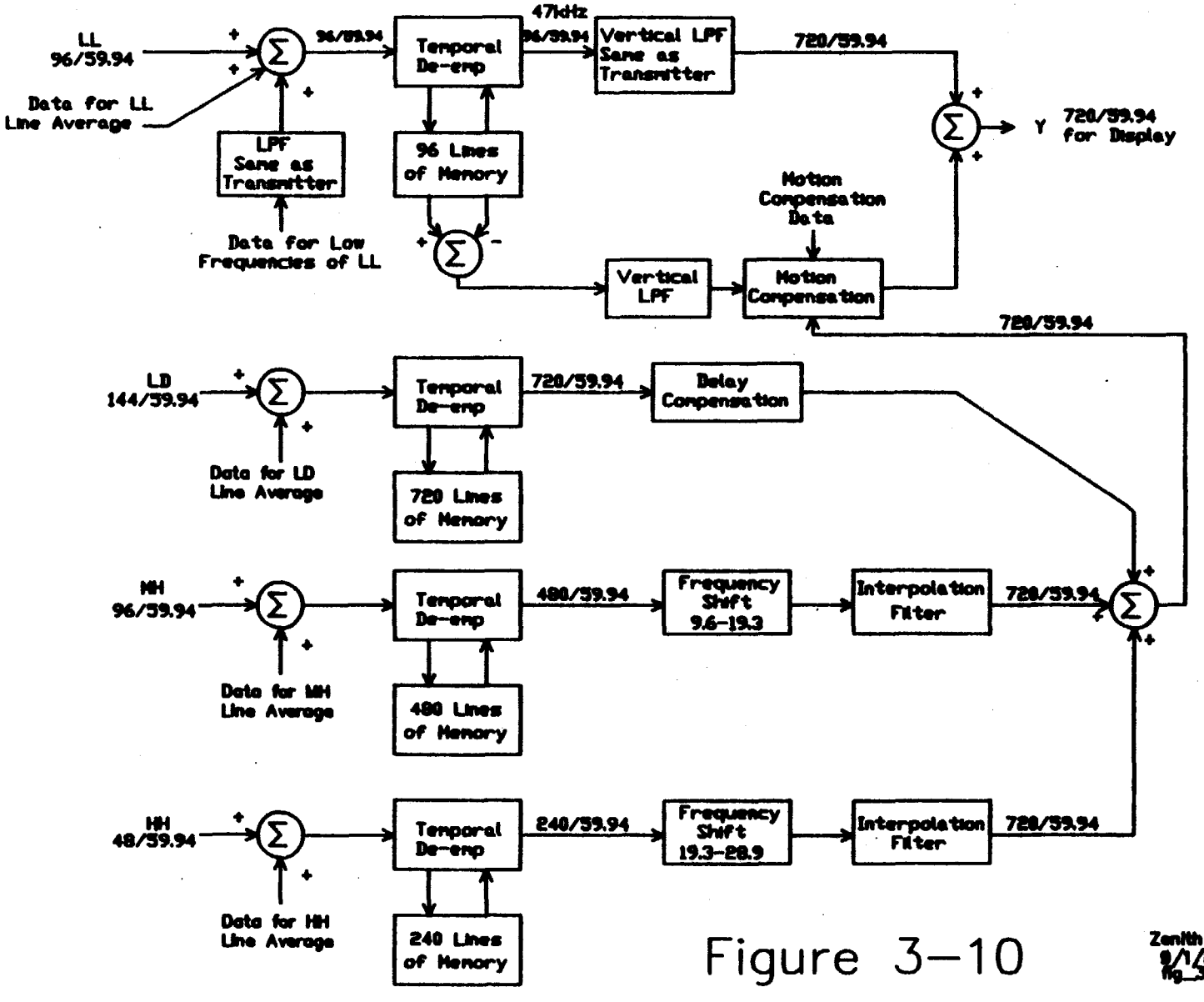


Figure 3-10

Zenith
9/1/88
fig_3-10