A COMPUTER BASED VIDEO SYNTHESIZER SYSTEM

Donald E. McArthur

June 1977

Experimental Television Center, Ltd.

164 Court Street

Binghamton, New York 13901

607-723-9509

Supported by NYSCA and NEA.

I. Introduction

As science advances, with the resulting advances in technology, we have new tools and new capabilities which influence our world in many ways. This new technology not only influences the traditional art forms but also produces new forms of art. The development of high speed electronic components and circuits, the cathode ray tube, the video camera, and inexpensive video tape recorders enabled the development of video art. Advances in integrated circuit design and fabrication techniques have led to the development of small but powerful computer systems which can be utilized by the video artist to achieve a new dimension of control over the video image. With a computer-based video synthesizer (CBVS), one can generate a sequence of images while controlling each individual image with detail and precision that is many orders of magnitude greater than is possible with manual control.

The ability to control the dynamics of the image is especially useful to the artist if the system is capable of generating the image in real time. With this requirement in mind, the natural choice of devices for converting electrical signals to visual images is the conventional video system. This choice also gives the capability of recording the video compositions with a conventional video tape recorder and of broadcasting to a large audience through existing network systems.

There are basically two modes of operation of the system: interactive-compositional mode and automatic-production mode. In the compositional mode, the artist can enter programs and parameters through the keyboard, observe the resulting sequence of images, and then modify parameters through either the keyboard or a real time input and thus build up a data set for a complete piece. The data set, representing all the aesthetic decisions made by the artist, is stored in the computer at each stage of the composition. When the composition is finished the system will operate in the automatic-production mode generating the final video signal in real time with no intervention by the artist. The artist may also choose to use a combination of these two modes in an interactive performance or to allow an audience to interact with the system operating automatically. The system is structured so that all of these variations can be accomodated by appropriate programming.

The system may be operated as a <u>generating synthesizer</u> which produces a video signal entirely from internal signals or as a <u>processing synthesizer</u> which utilizes video signals of external origin such as a camera. Either of these two types of operations is carried out by a configuration of elements modules, each of which performs a class of functions, with the specific function during one frame being determined by the control parameters recieved from the computer.

Since the computer functions only to generate the parameters which govern the behavior of the synthesizer modules, a video signal will be generated without operation of the computer. If the computer is stopped, the system will simply repeat the current frame until the parameters are changed. Thus the artist may choose to stop the computer and examine a single frame, or he may alter the program so that a given sequence is displayed very slowly or repeated very rapidly.

II. General Design Considerations

The NTSC video format is shown in figure 1. The time interval represented by one line is the time of one horizontal sweep or 1/15,734 sec. = 63.5 microseconds. The number by each line indicates the number of times that each format is repeated. There are 483 active scan lines and forty-two lines of vertical blanking making a total of 525 lines to compose one frame (two fields). The first line (a) consists of the horizontal sync pulse, which is five microseconds long occuring during a 11 microsecond blanking interval, followed by color burst and the picture information which last 52.6 microseconds. The next line (b) shows the first equalization pulse which is followed by five more and the beginning of vertical sync (lines c,d). Vertical sync starts in the middle of the line for this field and at the beginning of the line for the next field (line m). Lines e,f,g, and h show the completion of vertical sync with serrations and six more equalization pulses. Vertical blanking is shown in lines i and j. The second field starts in line j. Line k shows the 241 lines which interlace with the first field to complete the frame. Finally there are 6 more equalization pulses before (line l) and after (line n) vertical blanking (line m). Line o shows twelve more lines of vertical blanking. This sequence is repeated 30 times per second making a total of 108,000 frames per hour.

The portion of the video signal which carries intensity information is a continuously varying voltage and may be analyzed into components of different frequencies: low frequency components correspond to course structures in the image, and high frequency components correspond to fine structures. Although the video signal may, in general, have an arbitrary shape, a monitor will reproduce an image which corresponds to only a limited frequency range. This range is called the bandwidth. Thus components of the video signal which are outside the bandwidth of the monitor will not be visible. In a system any device which limits the bandwidth of the video signal will degrade the fine structure or spatial resolution of the image.

A video signal may be represented by a finite set of samples. These samples may be stored, eg. in a TBC or frame buffer, and then used to reproduce the video signal. In this process some of the information may be lost, but the lost information corresponds to high frequency components or fine structure of the image. If the sampling rate is sufficiently high, the lost information will be outside the bandwidth of the monitor, and no impairment of the image will occur. The minimum theoretical sampling rate which will retain the video information within a given bandwidth is called the Nyquist rate and is equal to twice the maximum frequency. Thus for broadcast quality, a minimum of 4.2 MHz. x 2 x 52 μSec. = 437 samples per line are required.

Each of the samples may be represented by a binary number with B digits. In general, a binary number with B digits has 2^B discrete values, and a B-digit representation corresponds to 2^B discrete gray levels. Since the video signal may have any one of a continuous set of values, an error is introduced when it is represented by a discrete number. Figure 2 shows the correspondence between binary and decimal numbers, the correspondence between a continuous video signal and its discrete representation, and the resulting error. This error is called the quantization error, and its root-mean-square value is given by $\sqrt{\int_0^1 \left(\frac{1}{2} \frac{1}{2^B} t \right)^2 dt} = \frac{1}{\sqrt{12}} \frac{1}{2^B}$.

If B is sufficiently large, the spacing between the discrete gray levels will be small compared to the intrinsic noise of the system and will not be visible on a monitor. The minimum acceptable value of B may be estimated by considering the quantization signal-to-noise ratio which is given by

$$S/N = 20 \log_{10} \left\{ \sqrt{12} \; 2^B \right\} = (6.02\,B + 10.79)\ \text{db}.$$

The signal-to-noise ratio for a one-half inch VTR is about 40 db, so about five binary digits are required for a comparable quality (monochrome) image.

For the analysis of equipment requirements for generating a video signal, we may utilize the measure of information given by the mathematical theory of communication. If a message occurs with probability P, then the amount of information (measured in bits) is

$$I = \log_2 (1/P).$$

Thus, the answer to a question which can be answered by yes or no with equal probability carries the quantity of information of

$$\log_2 1/(1/2) = \log_2 2 = 1 \text{ bit}.$$

Since a binary digit has two possible values, zero and one, each may carry at most one bit of information. Thus there are 5 x 437 x 483 = 1,055,355 bits of information in one frame of monochrome video signal (assuming thirty-two gray levels and 4.2 MHz. bandwidth). For generating a color signal, one may represent each of the red, green, and blue primary signals with a five bit word specified at time intervals of 100 nSec; this corresponds to an information rate of 150,000,000 bits per second. Other examples of amounts of information are typewriter keystroke, about six bits, and one typewritten page (double spaced), 9,000 bits. The rate at which the human brain can process information has been estimated to be about forty bits per second.

The mathematical theory of communication introduces another concept, redundancy, which is useful in analyzing a video synthesizer system. A message which contains N bits of information may be coded in a way that uses more than N binary digits. For example, the message pair (yes/no) may be coded as (111/000) using three binary digits instead of one per message. In this case the coding is redundant. The pattern shown in figure 3 consists of 16 x 16 = 256 squares, each of which is either all black or all white. Thus by representing white by one and black by zero, any pattern of this format can be represented by a code consisting of 256 binary digits. If all possible patterns are allowed, then one pattern carries 256 bits of information. A circuit which generates a video signal corresponding to this pattern codes the message into a form with 1,055,355 binary digits (assuming the spatial and intensity resolution as above). The redundancy of the video signal is further increased if the set of possible patterns is restricted by requiring that the total pattern is built of sixteen 4 x 4 blocks each being one of the following: all black, all white, black and white as in the upper left hand corner, or the latter with black and white interchanged. Then two binary digits code the choice for each block, and since there are sixteen blocks, thirty-two bits of information are contained in the entire pattern.

The disparity between the rate at which the human brain can process information, which limits the rate that an artist can manipulate controls of a video synthesizer, and the rate that information must be generated in order to produce a video signal clearly shows that the synthesizer must be structured in such a way to exploit redundancy in the video signal. A microprocessor typically requires 2 to 10 μSec. to execute a single instruction; thus it cannot possibly be used to generate a point every 100 nS. as required for a video signal in real time. On the other hand, one video field last 1/60 = 16.67 mS., a duration in which several thousand instructions can be executed. Thus a microprocessor is capable of generating signals according to a complex algorithm utilizing information supplied by the artist to produce control signals for high-speed special-purpose devices at a field-by-field rate.

Thus we are led to the hierarchical structure shown in Figure 4. Typical channel capacities are shown for each interconnection. The synthesizer consists of high speed special purpose circuits which generate a video signal with a character determined by the control signals supplied by the computer. The control signals fix the behavior of these circuits for an entire field. The computer takes information supplied by the artist, information defining the composition as a whole, and from it determines the control parameters required by the synthesizer for each field. Thus a gradual change in some picture parameters can be specified by the artist by a small set of numbers. In the simplest case only two numbers are required, the frame count of the first and last frame of the sequence. The computer utilizes this information to calculate the corresponding picture parameters for each field; thus it may produce several thousand control values.

III. System Structure

The CBVS consists of two parts: the computer section shown in the lower section of figure 5 and the video section shown in the upper section of the figure. Both sections operate simultaneously and independently, communicating through the buffer memory which has a capacity of 1,024 sixteen-bit words. Each of these words is either a picture element, a number which controls some function of the video section and determines some aspect of one field of the video image, or it is a picture feature, a number determined by the video section and may depend on an external signal such as a video camera signal. The buffer memory is connected to the computer bus through a sixteen-bit parallel interface which is structured in such a way that each word in the buffer memory is addressable and may be read or written in exactly the same way as words in the main computer memory. This memory-mapped I/O system simplifies the software which controls the buffer memory. In order to update an element such as a control D/A, the computer must execute an instruction which stores the new value in the location corresponding to that element.

During the active scan time, the control computer reads features from the buffer memory and generates elements for the following field and stores them in the buffer memory. During the vertical blanking interval, information is transferred through the element bus from the buffer memory to the element modules or from the feature modules to the buffer memory. The designation of a particular area of the buffer memory as an element or feature is under program control. During the transfer between the buffer memory and the element bus, the computer is locked out of the buffer memory. On completion of the transfer, the interface generates a vectored interrupt which requests the computer to generate parameters for the next field.

The computer system consists of a DEC LSI-11 microprocessor which has a sixteen-bit word length and an instruction execution time of about 7 microseconds, a Teletype Keyboard and printer connected through a serial interface, 20 K of dynamic memory, and a dual floppy disk system with a capacity of 256,256 bytes per diskette. An additional serial interface is also available for connecting through a modem to other computer systems. The entire system is dedicated to the synthesizer system.

The overall timing is determined by a 9.7552434 MHz clock which is phase locked to the subcarrier (3.579545 MHz). This frequency is chosen to insure a coherent subcarrier and to divide the active portion of the scan line into 512 pixels. The red, green, and blue signals are generated independently, and the chroma encoding is done with analog circuits; thus there is no advantage to following the common practice of making the pixel rate an integer multiple of the subcarrier frequency. With this clock frequency, a full nine bit word is used to define the horizontal position on the active portion of the raster. Figure 6 shows the X and Y wave forms. The X-Y module generates twenty bits of timing information (ten bits for horizontal, including the blanking period, and ten for the line count). This module also generates sync, drive, burst flag, and the transfer request $\overline{TR}$ signal which controls the timing of the buffer memory.

Timing details of the interface and buffer memory are shown in figure

7. The transfer request $\overline{TR}$ goes low at the beginning of vertical blanking initiating an arbitration for access to the buffer memory. If the computer is accessing the buffer memory, the current bus cycle is completed, then $\overline{READY}$ goes high, and the buffer memory controller cycles through memory making the required element and feature transfers through the element bus. When this is completed, $\overline{READY}$ goes low, control of the memory is returned to the computer, and an interrupt is generated requesting data for the following field. As indicated in the diagram, during the Nth field the computer is generating data for the N+1th field.

The timings of the signals on the element bus are indicated in figure 8. During the transfer, the memory controller generates the addresses, A_ϕ-A_9, the clock signals, $\overline{\phi_1}$, $\overline{\phi_2}$, and $\overline{\phi_3}$, and the status signals $\overline{CME}$ indicating a transfer from the memory to an element and $\overline{CFM}$ indicating a transfer from a feature module to the memory. The signals $\overline{ETF}$ and $\overline{FTE}$ are generated by the synthesizer modules and initiate a controller Element/Feature mode change. The three phase clock system is used to control modules which have the structure shown in figure 9. Functions which use data from the computer during the vertical blanking interval are disabled when the buffer memory accesses that particular element by a signal generated using $\overline{\phi_1}$. This allows access to the buffer memory during $\overline{\phi_2}$. The third clock $\overline{\phi_3}$ generates a memory write signal.

Time delays in the digital processing modules could produce errors and shifts of the image to the right. This is prevented by deskewing the output of each with a latch clocked by the master clock (9.755 MHz). Compensation for the resulting 102.5 nSec. delay in each module is provided by starting the X count at the beginning of the horizontal blanking interval rather than at the end. An additional shift to the right or left is then achieved by adding (mod 512) a constant supplied by the computer. The default value of this constant is 404 + number of elements.

IV. Element and Feature Modules

The structure described above supports a variety of element and feature modules which may be chosen and configured according to the taste of the artist. Our experience indicates that a large amount of work can be produced with a relatively small number of elements in a standard configuration. Whenever possible a new element added to the system is configured in such a way that if the control word is set equal to zero it has no effect on the system. Thus a minimum amount of reprogramming is required following system expansion.

Two general classes of modules have been developed, digital and hybrid. The hybrid elements are high-speed digital-to-analog converters used for generating the red, green, and blue video signals which are converted to NTSC format in the standard way and low-speed D/A converters used for generating control voltages, field-by-field controllable, used to operate existing voltage-controlled analog image processing systems such as keyers, raster manipulaters, etc. Another hybrid element is the analog video switching matrix. Four bits of one control word are used to select one of sixteen inputs for one output.

Digital processing elements include: constant, X + constant, Y + constant, twelve-channel sixteen-line demultiplexer with output complement, and four-channel four-bit by sixty-four word memory.

One of the possible ways of interconnecting digital element modules is shown schematically in figure 10. With this arrangement the X and Y signals are processed by the three sequences $(A_1 A_2 A_3 A_4)$, $(B_1 B_2 B_3 B_4)$, and $(C_1 C_2 C_3 C_4)$ to produce three different patterns or textures. These three signals are then combined by element D to produce a composite image. Finally, the digital-to-analog converter produces an analog video signal. The form of the video signal generated by this system depends on both the choice of configuration of modules and on the control parameters supplied by the computer. Since the control parameters are constant during one frame, the video signal may be represented by an equation of the form

$$V(t) = f\left[X(t),\ Y(t),\ E_1(t), \ldots\ldots\ldots E_n(t)\right]$$

where V(t) is the video signal which varies with time t. The structure of V(t) and of the resulting image is determined by the function f which is defined by the configuration of element modules. The time dependence of the video signal is shown explicitly; during one field, only X and Y change while the control values $E_1(t)$, $E_2(t)$, $E_3(t), \ldots E_n(t)$ are held constant. The time dependence of the video signal has been divided into two classes: firstly, the variation from one field and the next is determined by the computer through the element values, and secondly, the variation during one field is determined by the element modules through the X and Y signals for fixed element values.

V. System Operation

The power and versatility of this system may be seen by considering an application to an extremely simplified version of the system consisting of a pair of camera signals which are mixed by a voltage controlled mixer. With this system, only one element, a low-speed digital-to-analog converter, utilizes the element bus, and it generates a control voltage which determines the mixer operation for each frame. A flow chart of a program and an example of a data set are shown in figure 11. To use this system, the artist only needs to specify the numerical values in the data set.

The data set in this example corresponds to the mixer selecting the first camera until frame number 600, then fading to the second camera until frame 700, holding the second camera until frame 2000, then fading back to the first camera by frame number 2200. The values 8 and 4 in the parameter column determine the rate at which the fade takes place, and the numbers in the service routine column label indicate the selected service routine: a 1 for no change, a 2 for increasing the control voltage by P units, and a 3 for decreasing the control voltage by P units.

After the artist has stored these twelve numbers in the computer memory, the program may be started, and the computer will initialize the system and go to a background program where it waits for an interrupt. After every other interrupt the frame count is increased by one and compared with the frame count entry in the data set. Then the appropriate service routine (SR) is selected, and the element value is increased or decreased as required. Finally, the computer returns to the background program and waits for the next interrupt. This process is repeated for every field thus generating the sequence of control voltages and fading from one camera to the other.

The artist can observe the resulting sequence of images and then make changes in the data set to achieve the desired result. This technique may be extended to more complex systems involving several elements and feature modules with corresponding programs and data sets. Thus the artist can produce complex sequences with precise control of each frame. When the composition is finished, the system will automatically generate the video signal which may be displayed on a monitor or recorded on a VTR.

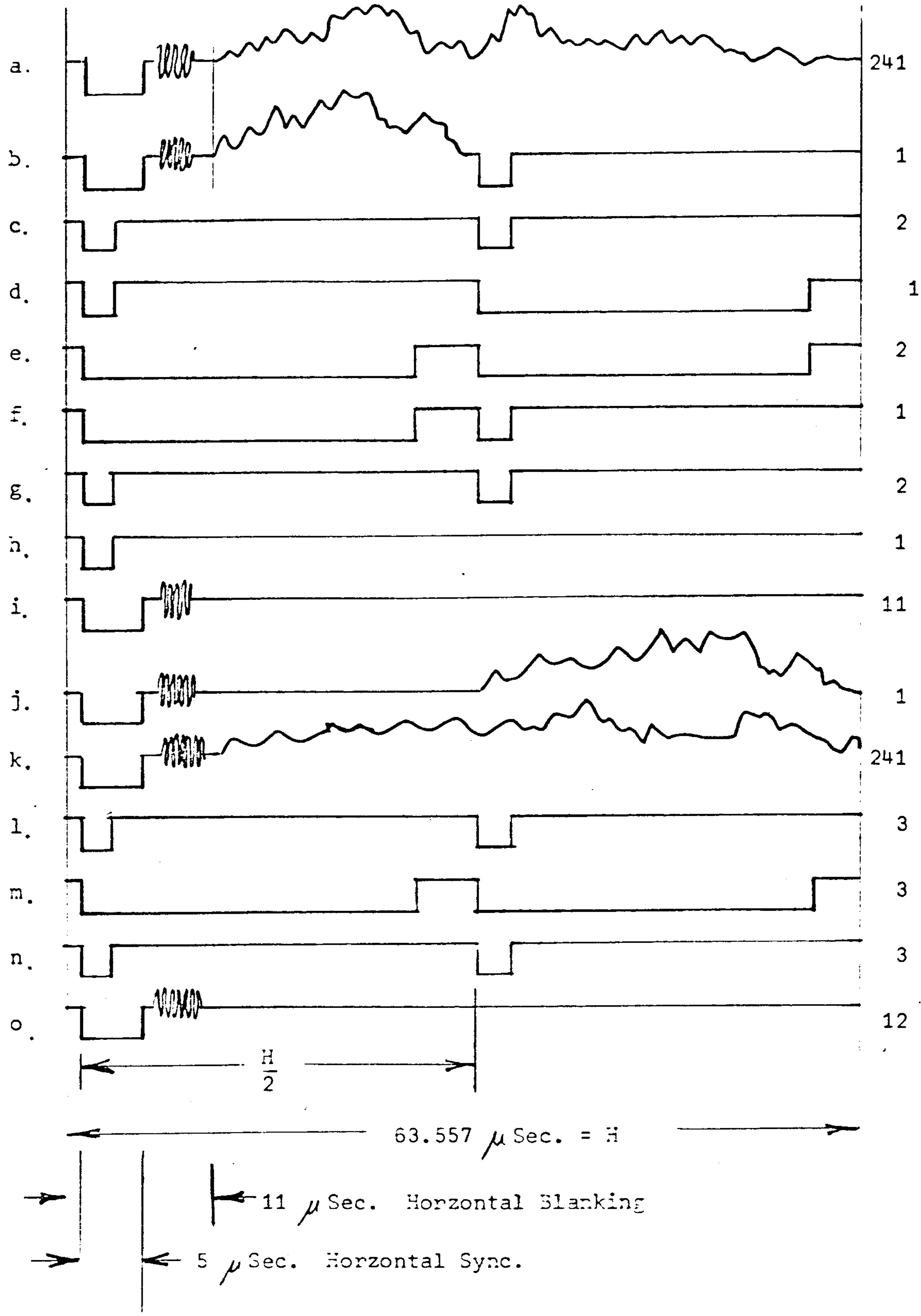
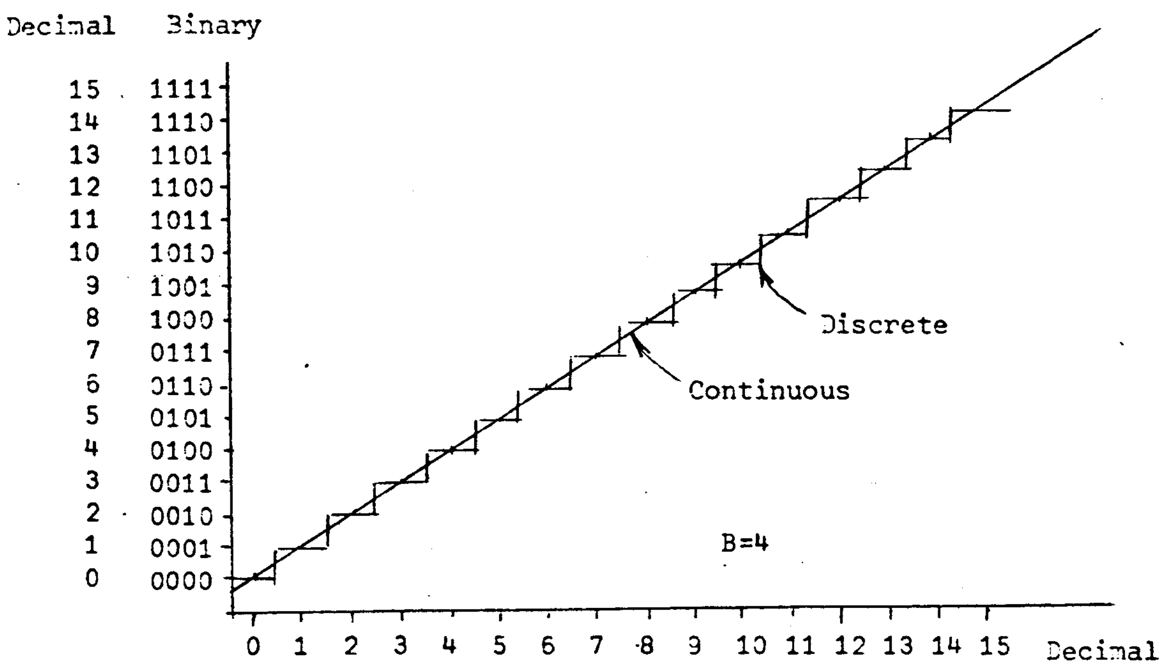

a.
241
b.
1
c.
2
d.
1
e.
2
f.
1
g.
2
h.
1
i.
11
j.
1
k.
241
l.
3
m.
3
n.
3
o.
12
H/2
63.557 μ Sec. = H
11 μ Sec. Horzontal Blanking
5 μ Sec. Horzontal Sync.

Figure 1


Decimal
Binary
15
1111
14
1110
13
1101
12
1100
11
1011
10
1010
9
1001
8
1000
7
0111
6
0110
5
0101
4
0100
3
0011
2
0010
1
0001
0
0000
Discrete
Continuous
B=4
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Decimal

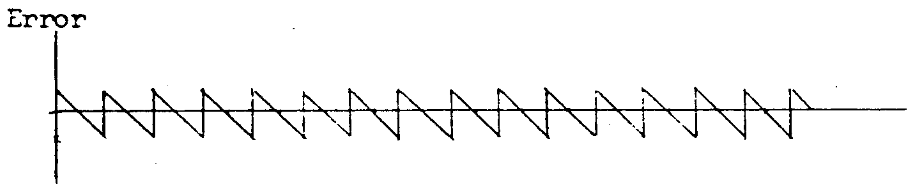

Error

Figure 2

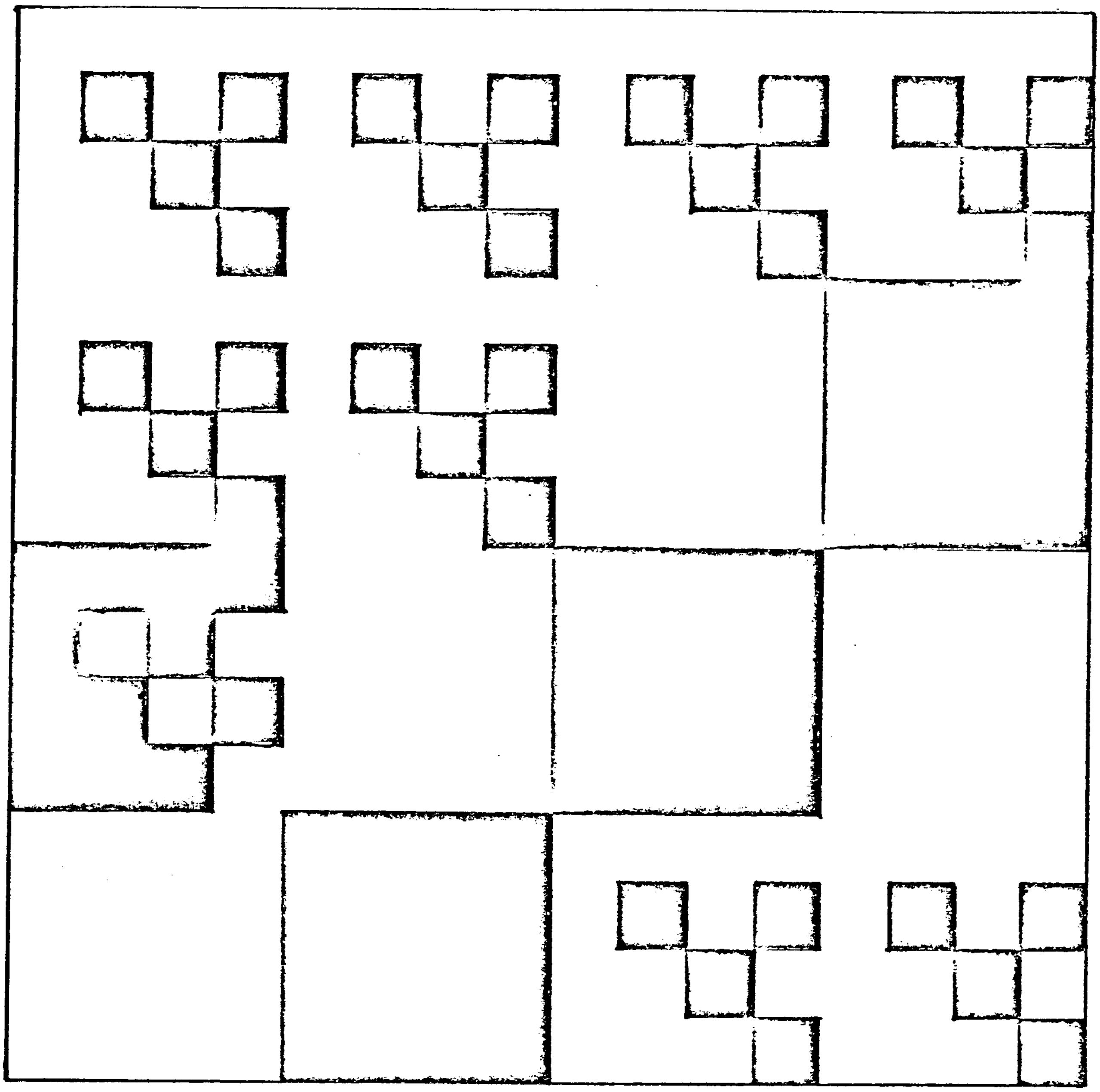


Figure 3

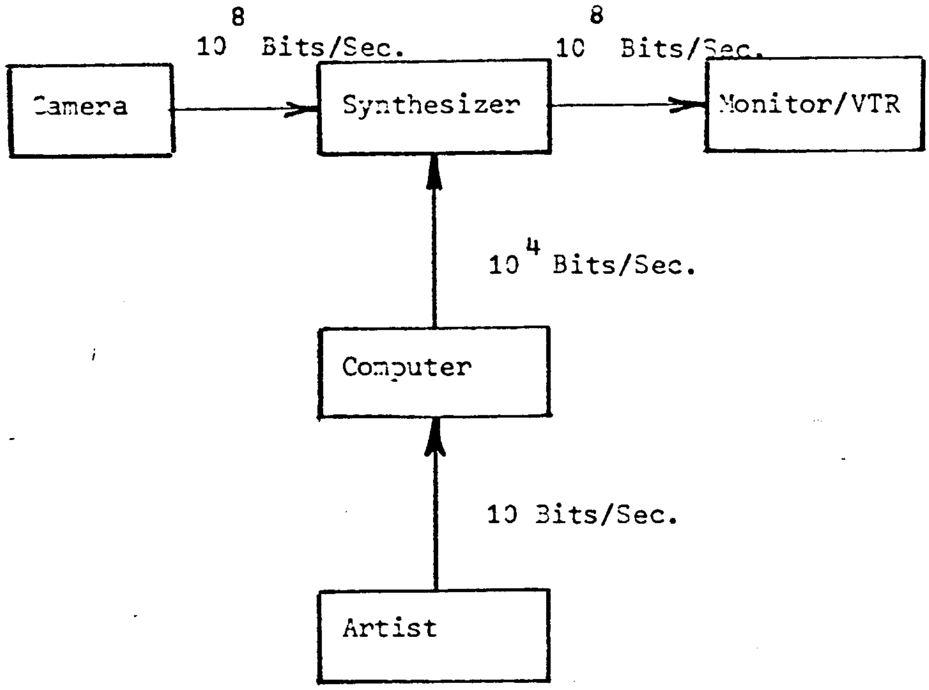

10^8 Bits/Sec.
10^8 Bits/Sec.
Camera
Synthesizer
Monitor/VTR
10^4 Bits/Sec.
Computer
10 Bits/Sec.
Artist

Figure 4

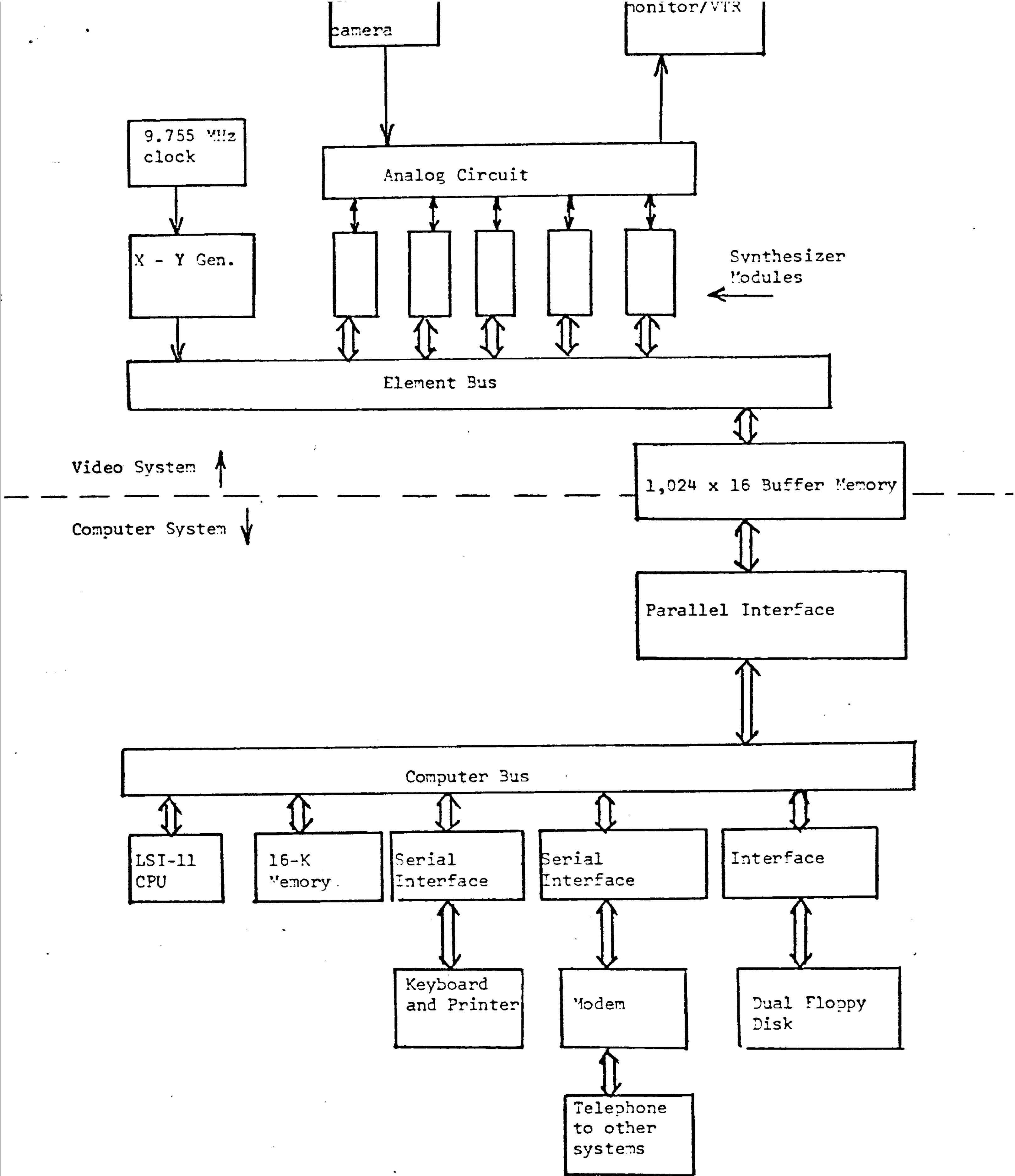

camera
monitor/VTR
9.755 MHz clock
X - Y Gen.
Analog Circuit
Synthesizer Modules
Element Bus
Video System
1,024 x 16 Buffer Memory
Computer System
Parallel Interface
Computer Bus
LSI-11 CPU
16-K Memory
Serial Interface
Serial Interface
Interface
Keyboard and Printer
Modem
Dual Floppy Disk
Telephone to other systems

Figure 5

X and Y Half Cycle Durations and Wave Forms

x_1 102.5 nSec.

x_2 205 nSec.

x_3 410 nSec.

x_4 820 nSec.

x_5 1.64 μSec.

x_6 3.28 μSec.

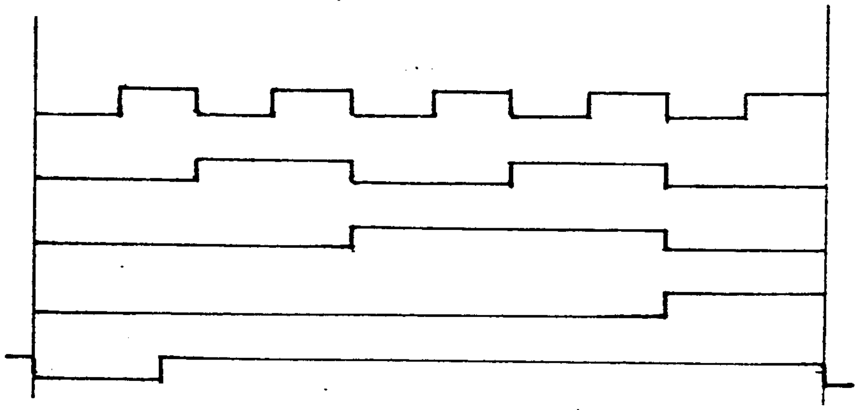
x_7 6.56 μSec.

x_8 13.12 μSec.

x_9 26.24 μSec.

x_{10} 52.48 μSec.

Horizontal Blanking



y_1 16.66 mSec.

Field Index

y_2 63.5 μSec.

y_3 127 μSec.

y_4 254 μSec.

y_5 508 μSec.

y_6 1.01 mSec.

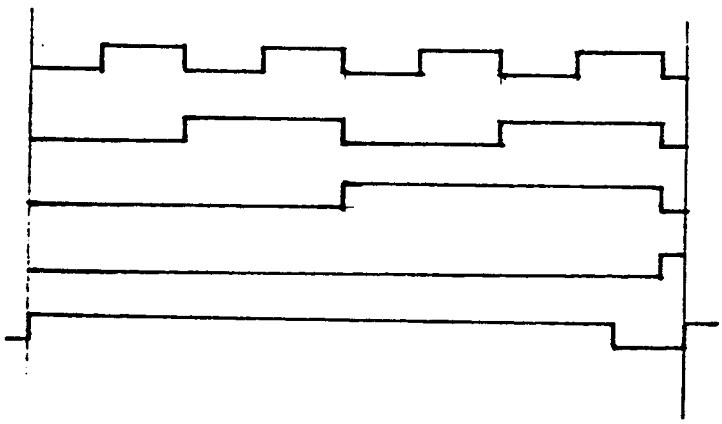
y_7 2.03 mSec.

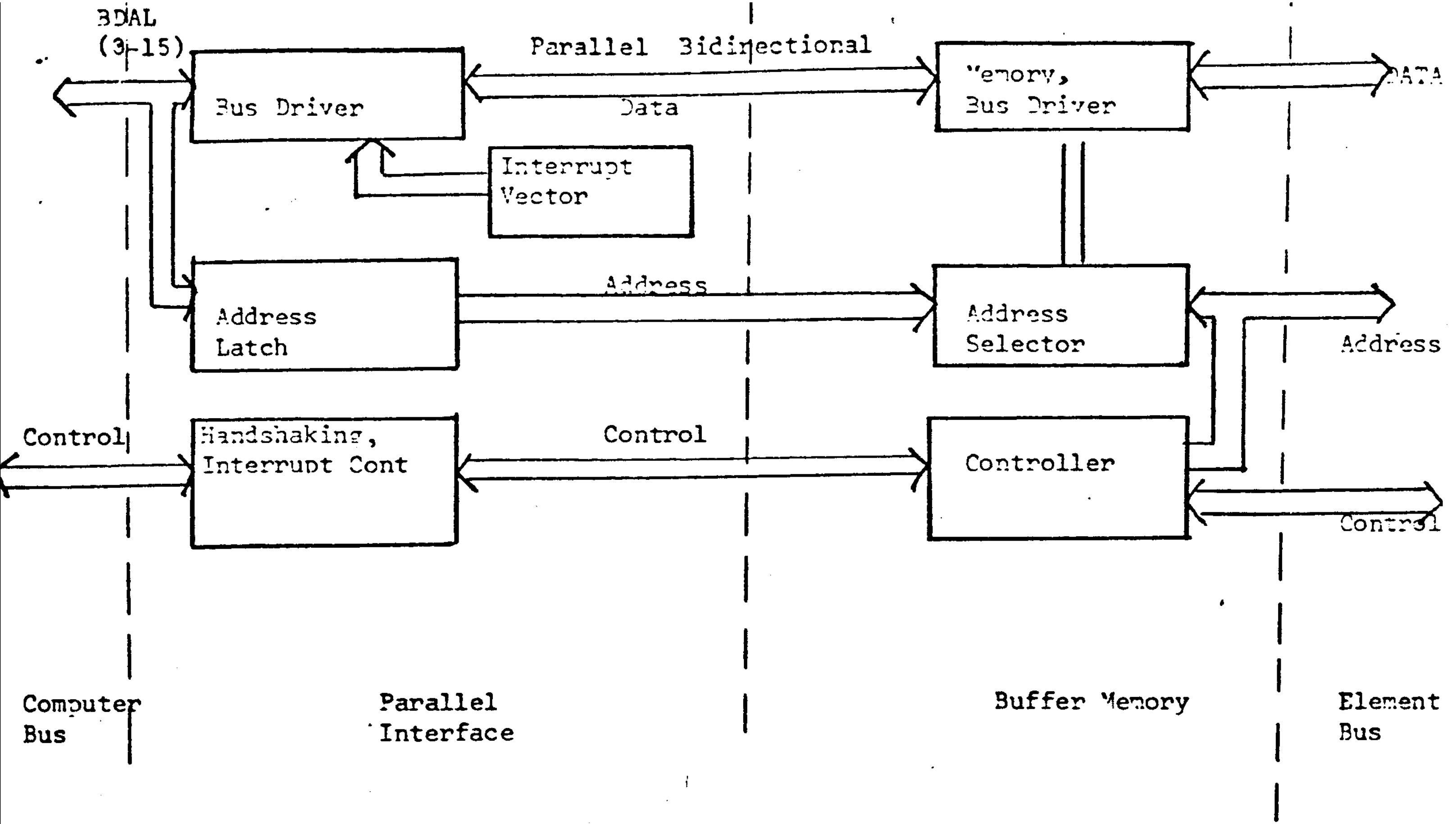
y_8 4.06 mSec.

y_9 8.13 mSec.

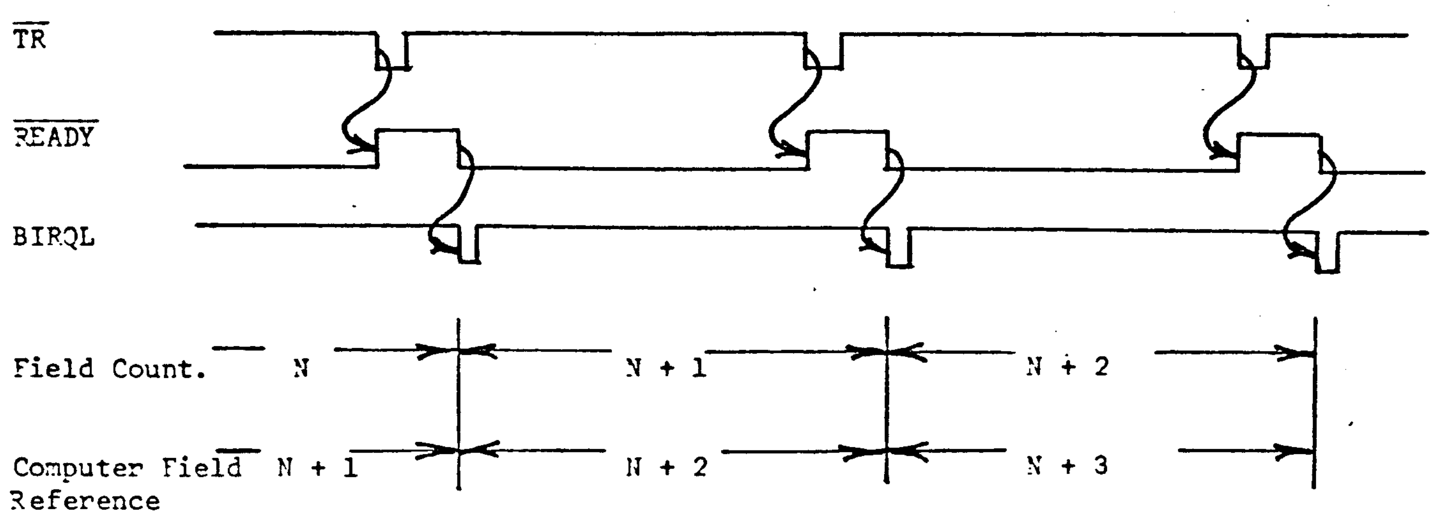
y_{10} 16.26 mSec.

Vertical Blanking





BDAL (3-15)
Parallel Bidirectional
Data
Bus Driver
Interrupt Vector
Memory, Bus Driver
DATA
Address
Address Latch
Address Selector
Address
Control
Handshaking, Interrupt Cont
Control
Controller
Control
Computer Bus
Parallel Interface
Buffer Memory
Element Bus



TR
READY
BIRQL
Field Count.
N
N + 1
N + 2
Computer Field Reference
N + 1
N + 2
N + 3

Figure 7

Element Bus

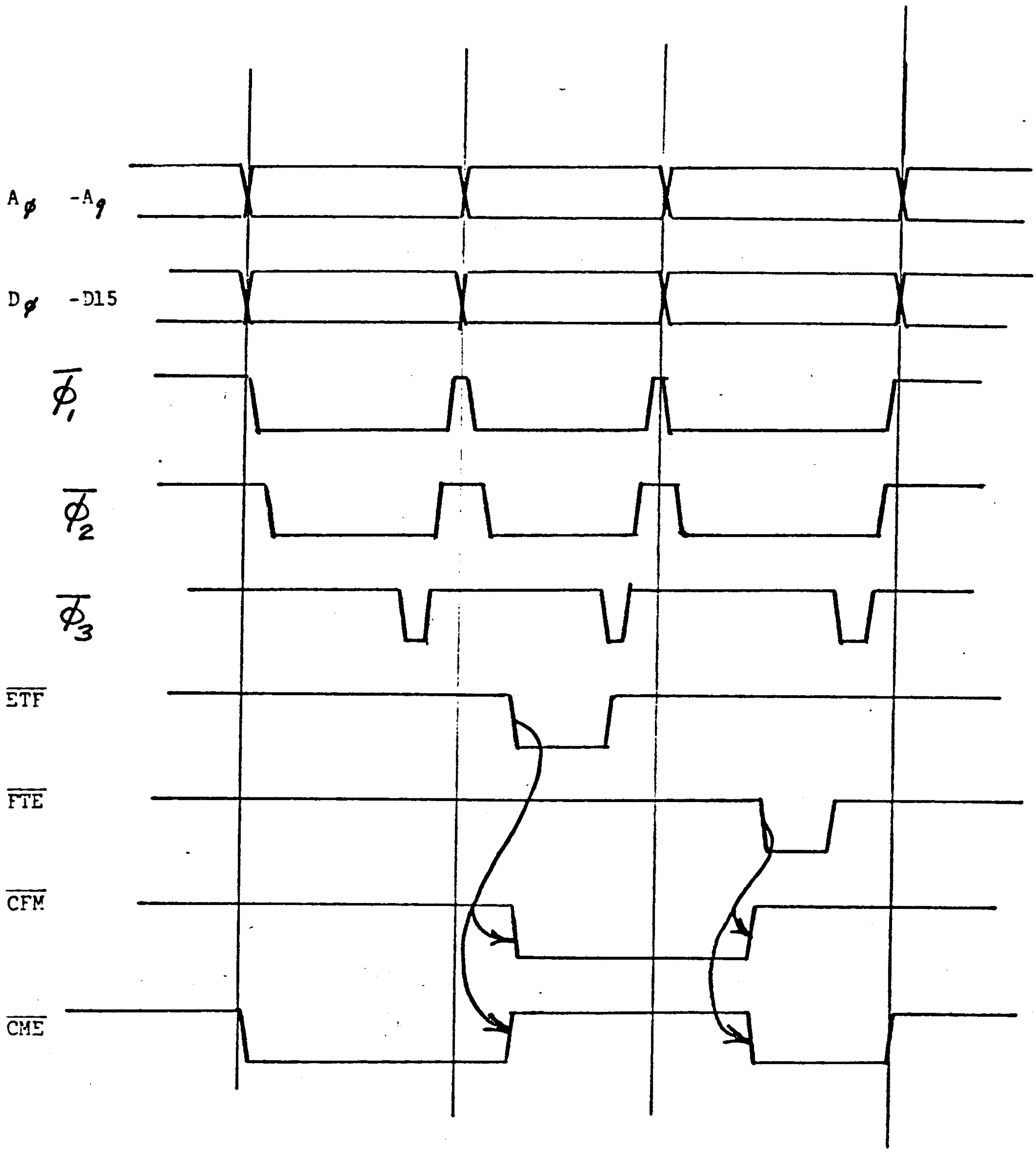


Figure 8

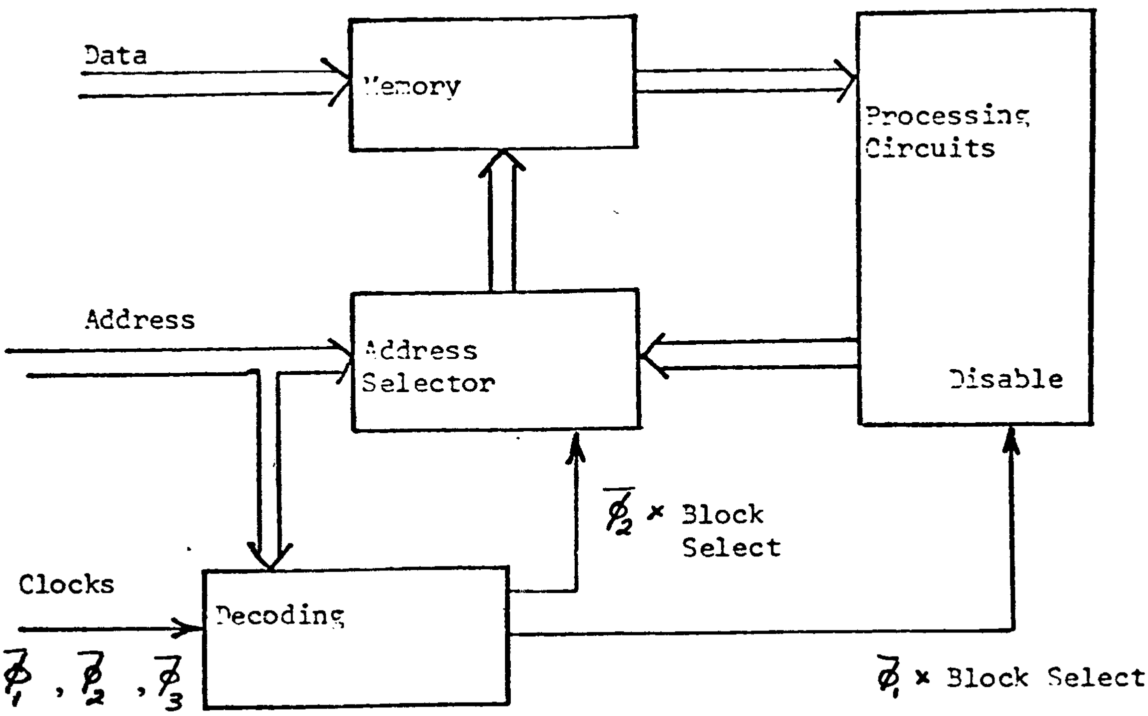

Data
Memory
Processing Circuits
Address
Address Selector
Disable
$\overline{\phi}_2$ × Block Select
Clocks
Decoding
$\overline{\phi}_1$, $\overline{\phi}_2$, $\overline{\phi}_3$
$\overline{\phi}_1$ × Block Select

Figure 3

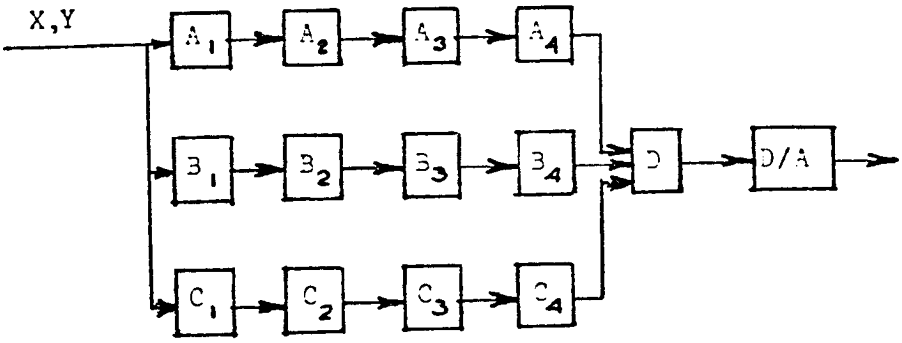
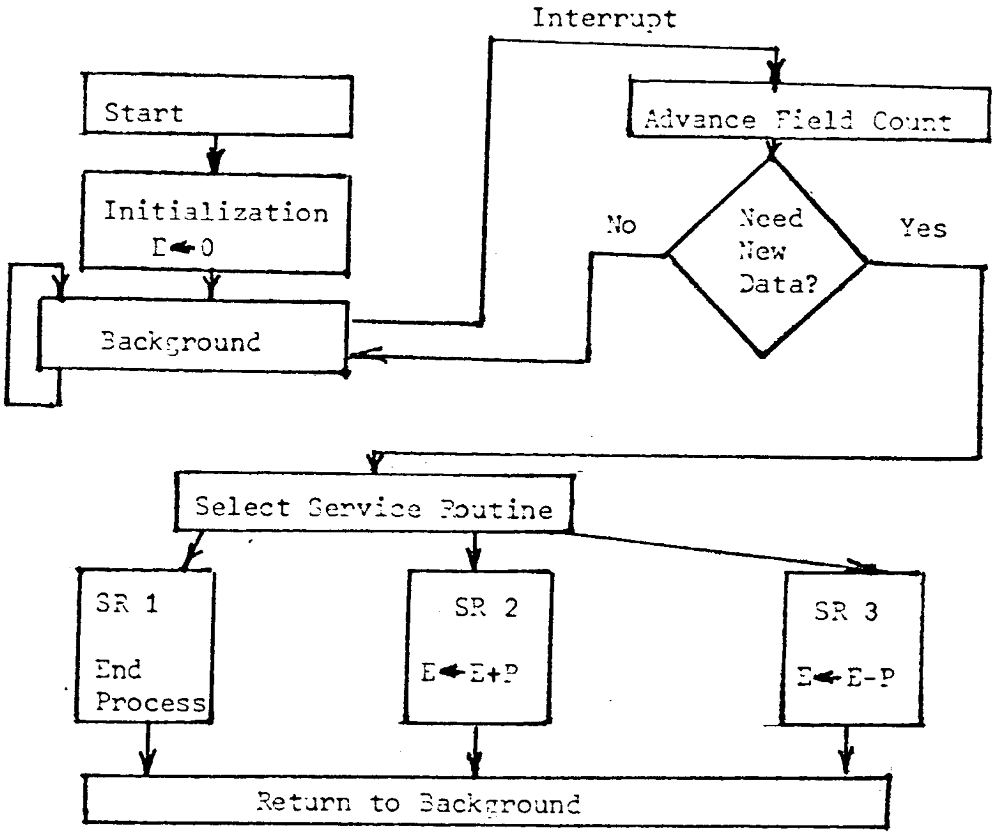


Figure 10



Data Set

Frame count	Service Routine Lable	Parameter P
600	2	8
700	1	0
2000	3	4
2200	1	0

Figure 11