DESCRIPTION OF PLANS FOR ROCKEFELLER FELLOWSHIP

BASIC DESIGN GOALS

The primary design goal of the Digital Image Processor (DIP) is to provide a comprehensive video image processing device for the individual video worker, that is to say, to provide in a single instrument all of the control that a video worker requires to create or, as some say, produce work. This would include standard special effects that a TV studio can produce like fades and switches, color correction for errors in illumination, noise reduction and image enhancement techniques, time base correction, squeeze zooms, multiple images, computer control of tape recorders for editing, flip-flops, zaps, and, of course, color chromakey. In addition to effects that have names, presently used in studios, it would be able to implement an immense variety of processes and patterns that the user will be able to create. The basic concept here is the creation of a generalized device rather than a special purpose device. A modern TV studio uses the organizing principle of a large number of special purpose devices which are interconnected to produce the required processing for the TV signal. The Digital Image Processor is a general purpose device which can implement a larger range of processing than combinations of special purpose devices in a color TV studio.

The second major design goal is a user oriented or habitable system. It is clear that a general purpose device, a device that can do an immense variety of things, is inherently more difficult to control than a set of special purpose devices. Great pains will be taken to make the DIP accessible and easy to use by people without a large amount of technical training. Critical to the implementation of this design goal is the use of a small computer graphics system, the Z-Grass machine developed by Dr. Thomas DeFanti, to control the DIP. The primary function of this computer graphics system will be to accept control information from the user and to inform the user what the state of the instrument is. The input control can come in many forms. It can be commands on a typewriter keyboard, dial turning, joy stick waggling, various touch sensitive devices, or even sensitive clothing. The concept here is that the general purpose digital computer forms an intelligent interface between the control devices which are run by the user of the instrument and the actual control of the Digital Image Processor. It is hoped that this will accomplish the goals of both a highly facile performance instrument in the sense of a musical performing instrument and a program controllable studio instrument.

HISTORY, PRESENT AND FUTURE OF THE PROJECT

I began to consider building this instrument about five years ago. Two years ago, I received a Guggenheim Fellowship to design and construct what turns out to be the conceptual father of this design. The difficulty was that the design was of such large scale and of such expense that near the end of the fellowship period when I had completed a great deal of the design and had actually begun construction of one of the modules, it became clear that I couldn't continue without large scale industrial support and a veritable research laboratory of electronic designers. In virtual despair over this situation, the idea of time division multiplexing the signal and lowering the number of bits necessary to represent the signal occurred to me. This ephemeralized the design to between one-third and one-ninth of its previous size. This brought the design back into a cost range that would be affordable by dedicated individuals and by small institutions. Fortunately, only about a third of the work completed on the Guggenheim Fellowship is lost. The rest applies rather directly to this design, including a large parts stock.

At this point, I am ready to start construction of the first modules for the new Digital Image Processor (DIP). Financed by my own money, progress will be very slow. The major virtues for the video community of supporting this work with a Rockefeller Fellowship would be as follows:

The instrument would be completed years before it would be under my own individual resources. The free distribution of the plans to individuals and not-for-profit institutions would be assured and the reduction to printed circuit board instead of wirewrap technology would make the instrument twice as compact, which really means twice as powerful in the same physical size, and would greatly ease the effort of copying the instrument. I would not be able to afford this design step, the reduction from wirewrap to printed circuit board with my own resources.

DISTRIBUTION

My plans for distribution of this design, should I receive a Rockefeller Fellowship, would be as follows:

I will complete sufficient documentation of the instrument so that an individual without immense technical background would be able to successfully copy the instrument and then I will give away the plans for the cost of postage and printing cost or some nominal charge associated with my cost of distributing them.

This distribution technique may seem ill-conceived at first look, but in the case of the Analog Image Processor, it turned out to be an extremely effective form of distribution. First of all, it has resulted in over 20 copies of the Analog Image Processor (AIP) being constructed, mostly by individual artists who actually use the instrument to do a large range of productions. A considerable percentage of video art broadcast in Chicago has been performed on the AIP and there is a large number of artists who would be doing an entirely different kind of expressive work without the availability of this instrument at low financial cost. An example of another distribution technique would be to try and manufacture and sell the instrument. A comparable instrument in structure to the AIP is the Rutt-Etra Scan Processing Video Synthesizer. It turns out that there are approximately equal numbers of the Rutt-Etra Synthesizer and the AIP indicating rough parity in terms of success of the two distribution techniques. The AIP, however, is on its own. It is still being copied and constructed by individuals with relatively little effort on my own part, so that it is likely that the number of Image Processors will continue to grow in the near future.

A second aspect of this distribution technique is that because it requires a relatively low financial commitment and a relatively large personal time commitment, it tends to lend itself to the individual practicing video worker, who typically has relatively small income but considerable dedication, so I consider it to be an optimum distribution technique for the designs of the DIP.

BUDGET SUMMARY

The $10,000 stipend, I would use to take a leave of absence from the University of Illinois at Chicago Circle for between 6 and 9 months to focus on the implementation and documentation of the design of the DIP. What follows is an approximate budget for the $25,000 expense category.

Test Equipment	$3,000
Electronic and Printed Circuit Design Services	5,000
Printed Circuit Production Services	4,000
Z-Grass Computer	5,000
Specialized Software for Computer	5,000
Electronic Parts	5,000
Help on Documentation	1,000
Printing of Documentation	1,000
Telecommunications	750
Miscellaneous Office Supplies, Videotape, Digital Storage Medium	1,000
Travel (My collaborators on this project live on the West Coast.)	3,000
Administrative Expense (probably scarfed up by the sponsoring institution) at least	3,000

This totals to more than $25,000. The difference plus probably much more will come from my personal savings.

TECHNICAL DESCRIPTION OF INSTRUMENT

Figure 1 is the first systems block diagram. As indicated above, there is a general purpose computer graphics system which controls a special purpose video signal processor called the Digital Image Processor and also receives input from conventional computer terminals, dials, joysticks, pressure sensitive surfaces, keyboards, and a variety of input devices, many of which I have already developed under a National Endowment for the Arts grant. The report of that grant is included in the Appendices.

The communication protocol used between the general purpose computer and the digital signal processor is a standard parallel I-O port configuration, so that as new developments in low cost computing become available, they can be incorporated for controlling digital image processors. The computer chosen for development is the Z-Grass machine. The hardware was developed by Dave Nutting and Associates and it will be manufactured and distributed by Datamax Corporation. It is a Z80 based computer graphics system which is exquisitely well adapted to the control of visual and sonic entities. The software for that system was developed by Dr. Thomas DeFanti of the University of Illinois at Chicago Circle and represents the state of the art in interactive real-time control of images and events. Dr. DeFanti has committed himself to extensions to the Z-Grass language explicitly for the control of the Digital Image Processor (DIP). It is also likely that Terry Diz will complete a doctoral dissertation in Information Engineering on the implementation of a user oriented control system for the Digital Image Processor in Z-Grass. It is most fortunate that these two highly skilled and creative workers are able to contribute to this project.

The DIP will be constructed in four successive phases. The block diagrams 1 through 8 represent the Phase 1 instrument. As indicated in Figure 1, the digital signal processor contains two main subsections, the analog box and the digital box. Figure 2 is a block diagram to card level of the analog box. The analog box takes care of all of the video protocol. It genlocks to an external source, does the color decoding and analog to digital conversion for color cameras and the digital to analog conversion for output. Figure 3 is a diagram to card level of the digital box. The digital box contains the digital processing modules, an 8x8 matrix for interconnecting the digital processing moduels, a digital color encoder, two pattern generation modules and a test generator and receiver module. Figure 4 shows the interconnection diagram of the processing modules, camera and pattern generator inputs and the 8x8 matrix. The matrix allows connection from any input on the left to any set of outputs on the bottom. The digital signals routed by that matrix to the processing modules are parallel pipeline time division multiplexed signals. Parallel pipeline architecture is similar to a bucket brigade. Figure 5 is a diagram illustrating this. The signal coming from the left which, say, represents a gray value, is latched at the first latch, processed and then passed on to the next stage where it is latched again. At the same time a new piece of information is coming in from the left. From the middle latch the information is processed and sent ot the next latch aed on the next clock tick, that information is latched. So that on each clock tick, information is stored, processed, and transferred to the next latch. The instrument runs at 14.31818 megahertz, meaning a 69 nanosecond processing time from latch to latch. This is a digital architecture that is capable of passing and processing information fast enough to do real-time video. The time division multiplexing means that on succeeding clock ticks, different kinds of information are transferred. A color experience is characterized by three dimensions of information, normally thought of as red, green, and blue, the amount of primaries associated with any color experience. Time division

multiplexing, for instance, could send a red signal, then a green signal, then a blue signal. But to take advantage of inherent encoding efficiencies of NTSC video, the time division multiplexing is as follows. A sample of Y luminence is sent and then a sample of R-Y, which is one of the chromance signals, another sample of luminence, and then a sample of B-Y is sent, and then it repeats again with Y. In this form luminence information or gray value information is sampled twice as often as the chromance information. This is consistent with the amount of information available in eahc of those categories with NTSC TV.

Another efficiency of the design is to send Y as 7 bits worth of information, meaning 128 possible values and to send R-Y and B-Y with 6 bits worth of information or 64 possibilities each. This is also consistent with the amount of information available in NTSC TV.

The basic processing modules are illustrated in Figure 6. The physical module is enclosed in a dotted line, but it is best to think of the process as the logical module illustrated. The lookup table outside of the dotted lines is actually present in the next physical module. The lookup tables illustrated are high speed memories where the data coming in is the address and the data going out is the data stored at that address. This forms an arbitrary lookup table so that for every value coming in there is an arbitrary value going out. Because of the time division multiplexing, each box can be thought of as three independent lookup tables, one for Y, one for R-Y, and one for B-Y. In addition to the lookup tables, there are delay elements which can generate an arbitrary delay of the signal. It essentially is a pipeline with no processing, and an adder indicated by a box with a + sign. This form of module is capable of doing colorization, color correction, addition (which is superimposition), multiplication (which is keying), and utilizing the delay elements can form both transversal and recursive filters that are useful in controlling the amount of detail present in the image and can be used for various image enhancement and noise reduction algorithms. There are 3 of these logical modules in the Phase 1 DIP. The pattern generators are illustrated in Figure 7. They are composed of phase accumulators, memory lookup tables, and adders. In addition to generating the classic wipe and split screen patterns and varying them dynamically, it can generate a very large range of interesting images, one of the more bizarre of which is illustrated in Figure 8. These patterns can be used as images themselves or can be used to control the processing of other images.

The Phase 1 DIP is expected to cost about $5,000 in parts. In addition to that the computer system required to control it will cost between $2,500 and $5,000 depending on peripherals. The chief limitation of the Phase 1 instrument is that it will not be able to utilize more than one videotape input. In order to use more than one videotape input, multiple time base correctors are required. These can be added to the design for approximately $1000 a channel in parts.

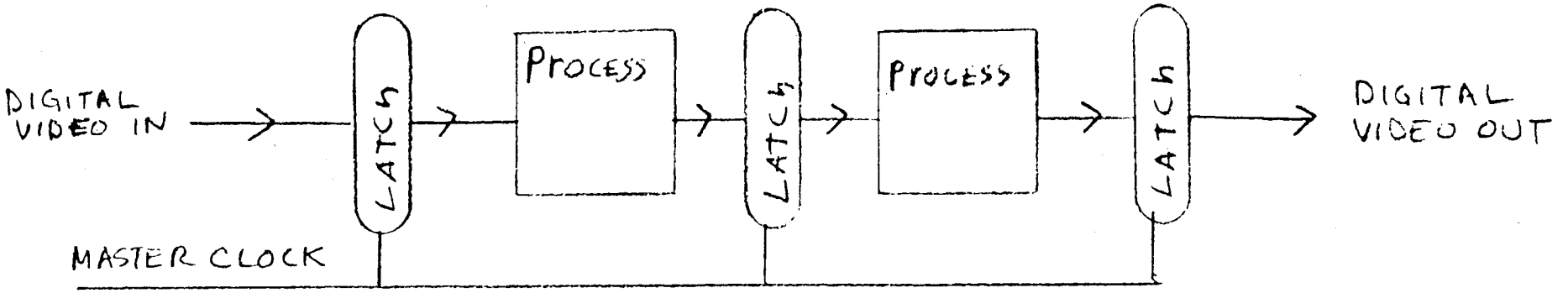
The Phase 2 instrument would be able to do A-B roll editing through an image processor and computer control of tape recorders forming a complete postproduction facility capable of wipes and fades and keys along with more interesting combinations of two or more input sources on tape.

The Phase 3 DIP would include a new structure called the convolution and geometric processing module. Since all the inputs and outputs of the matrix are already used in the Phase 1 and 2 instrument, this would require enlarging the switching matrix and adding more mechanical boxes to the system. The convolution function is extremely powerful in doing advanced image enhancement, feature extraction, and noise reduction techniques in addition to giving the user very detailed control over the amount of spatial information in the image being processed. The geometric processing element of the convolution processor would allow all of the effects which are normally associated with a fram buffer such as squeezing and

flipping and multiple images, but only in the horizontal direction. In other words, you could create a bilaterally symmetric image or the same image repeated twice in the horizontal direction. The price of the convolution processor at this point is uncalculated. I guess it to be around $1000 a channel in parts.

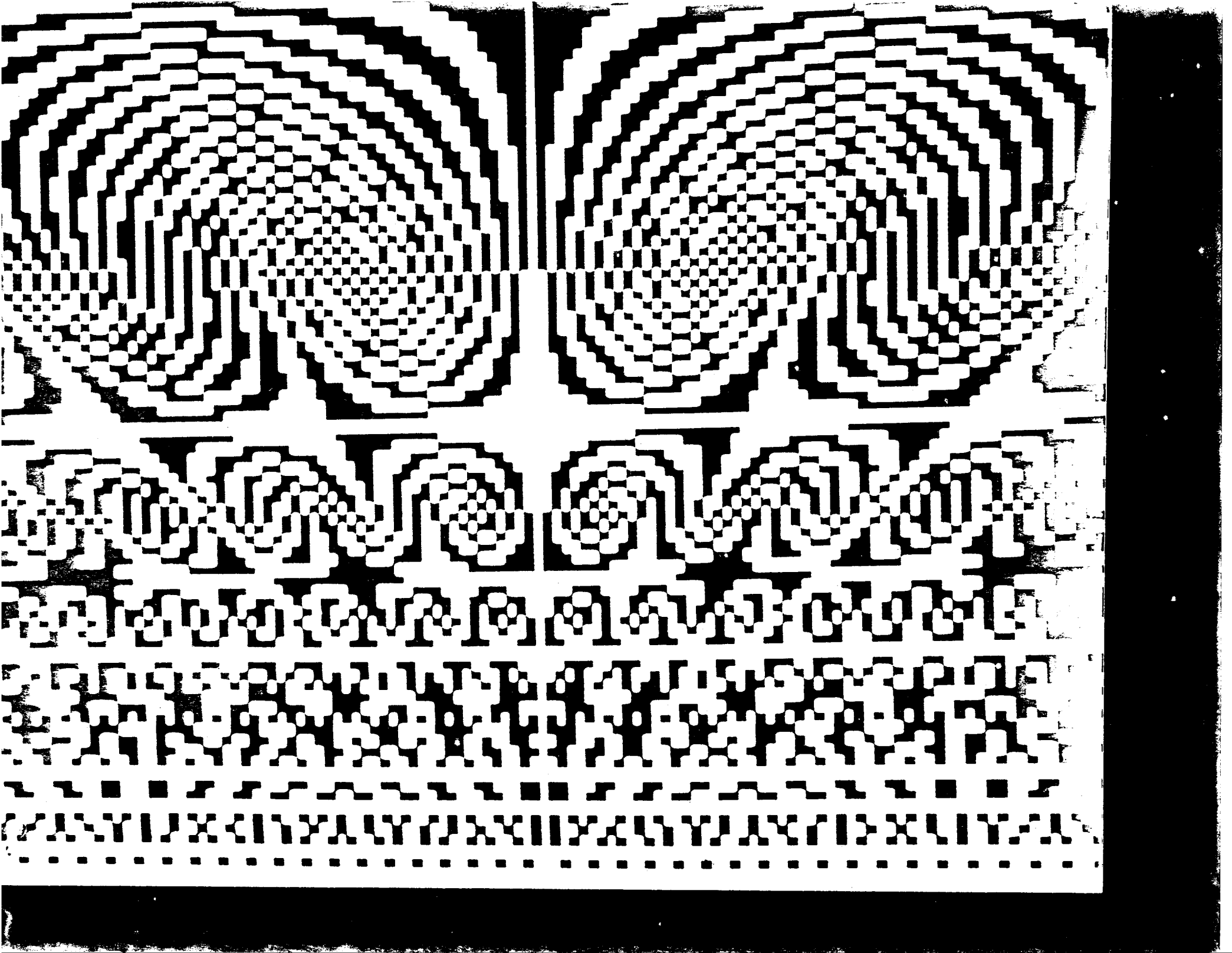
The Phase 4 instrument would utlize a full frame buffer. This frame buffer would be external to the instrument and probably of someone else's design. There are a number of excellent people in the country working on versatile frame buffers. With this element, one would be able to do functions that are euphemistically called electronic opticals, such as zoom-ins and zoom-outs (called squeezes), pans, rotations, and multiple images.

PARALLEL PIPELINE PROCESSING



Y	7 BITS
R-Y	6 BITS
Y	7 BITS
B-Y	6 BITS
Y	7 BITS
ETC	

TIME DIVISION MULTIPLEXING



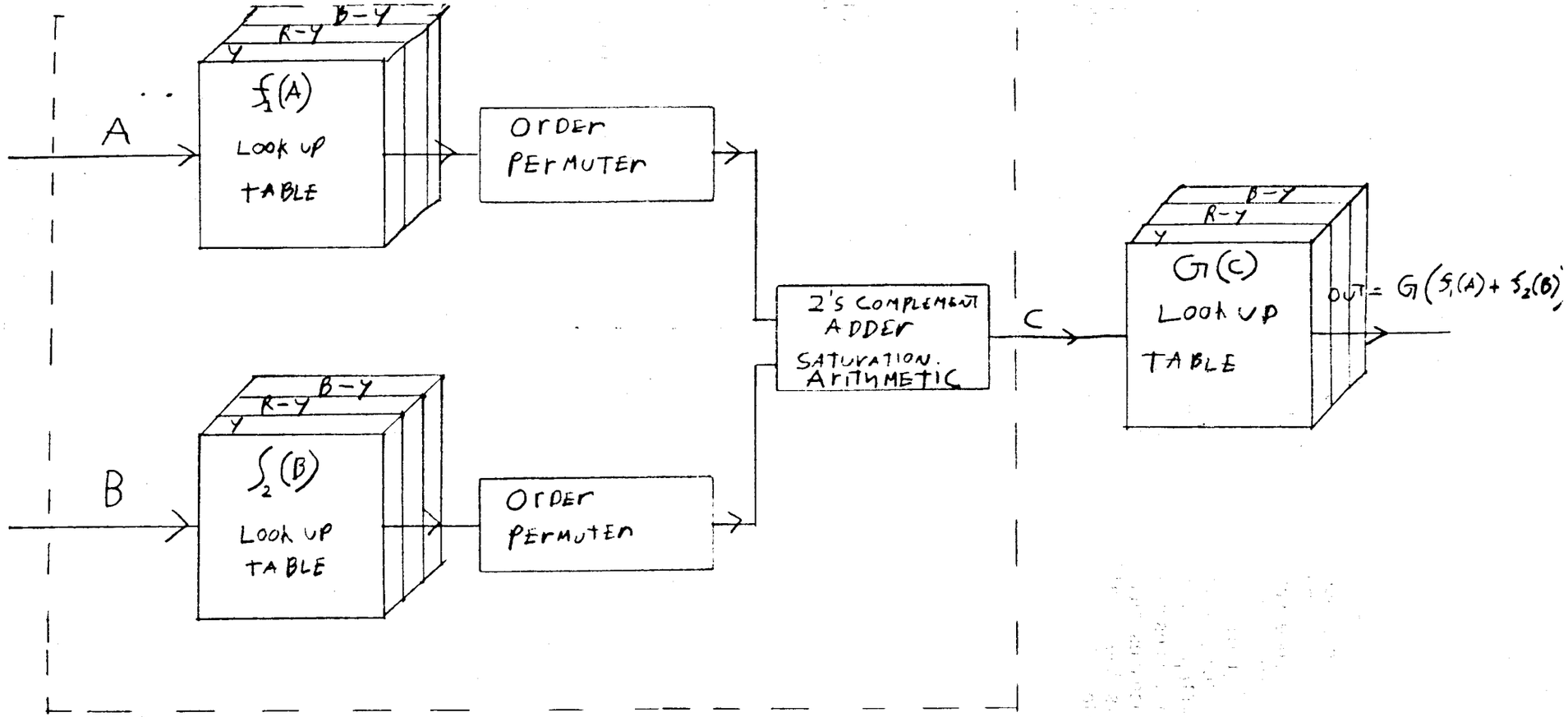
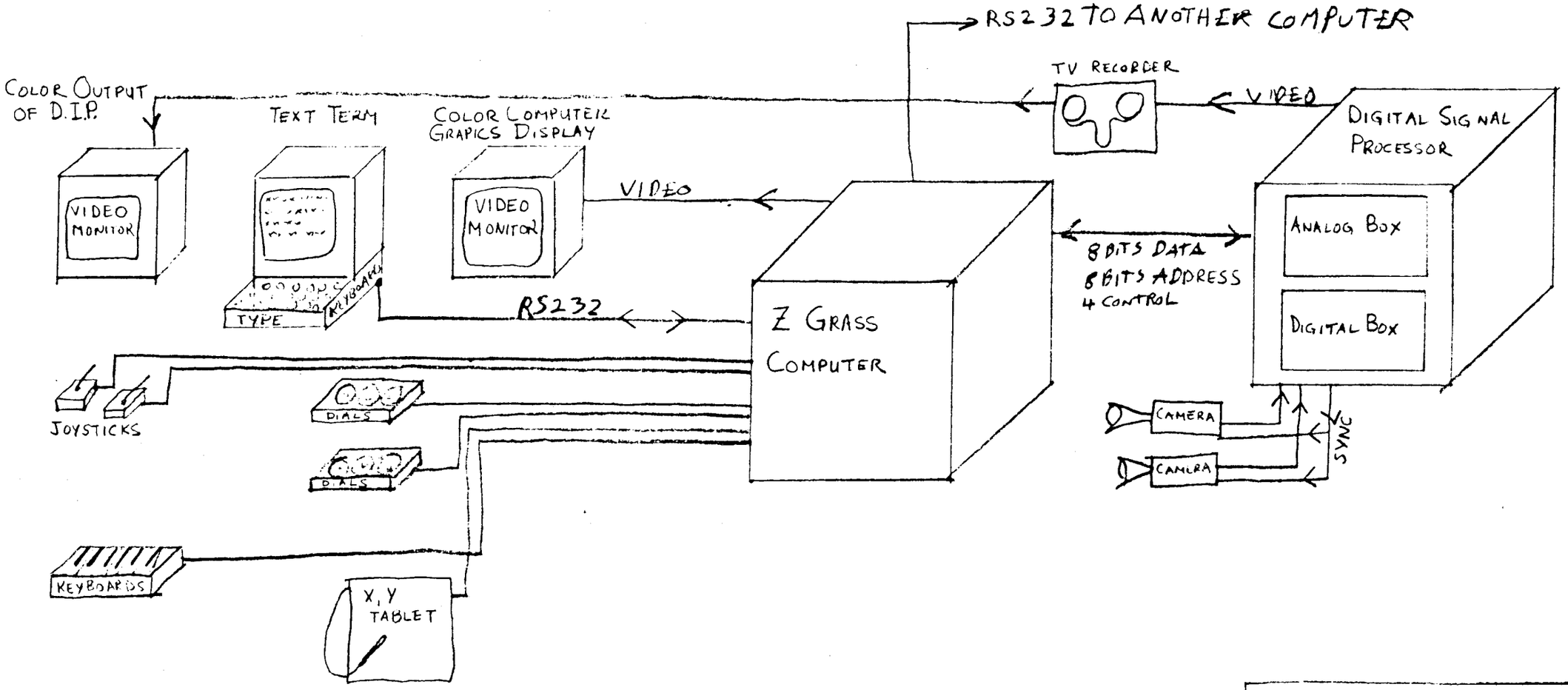
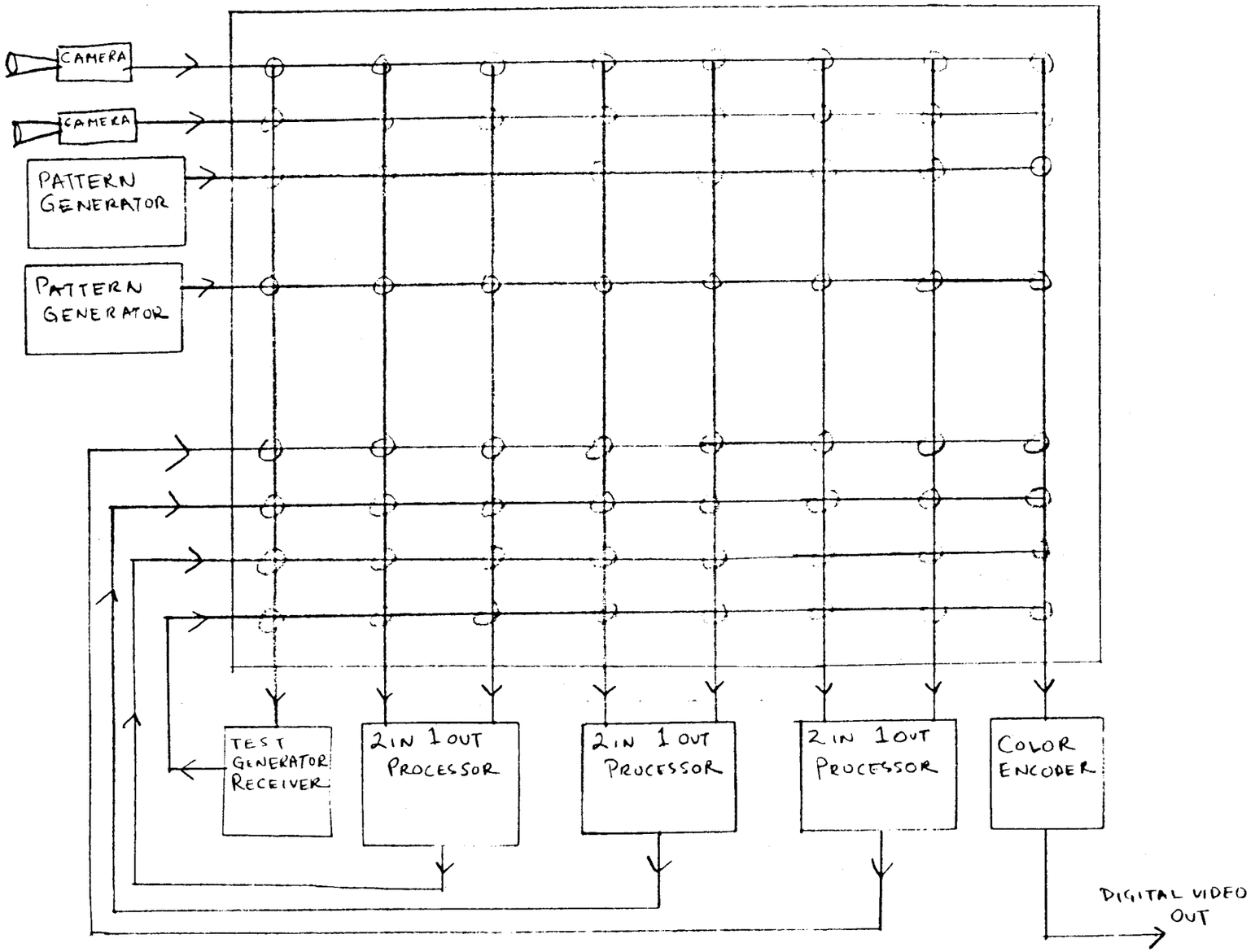

A
B
B-Y
R-Y
Y
$f_1(A)$
LOOK UP
TABLE
$f_2(B)$
LOOK UP
TABLE
ORDER
PERMUTER
ORDER
PERMUTER
2's COMPLEMENT
ADDER
SATURATION.
ARITHMETIC
C
G(C)
LOOK UP
TABLE
OUT = $G(f_1(A) + f_2(B))$

FIG C



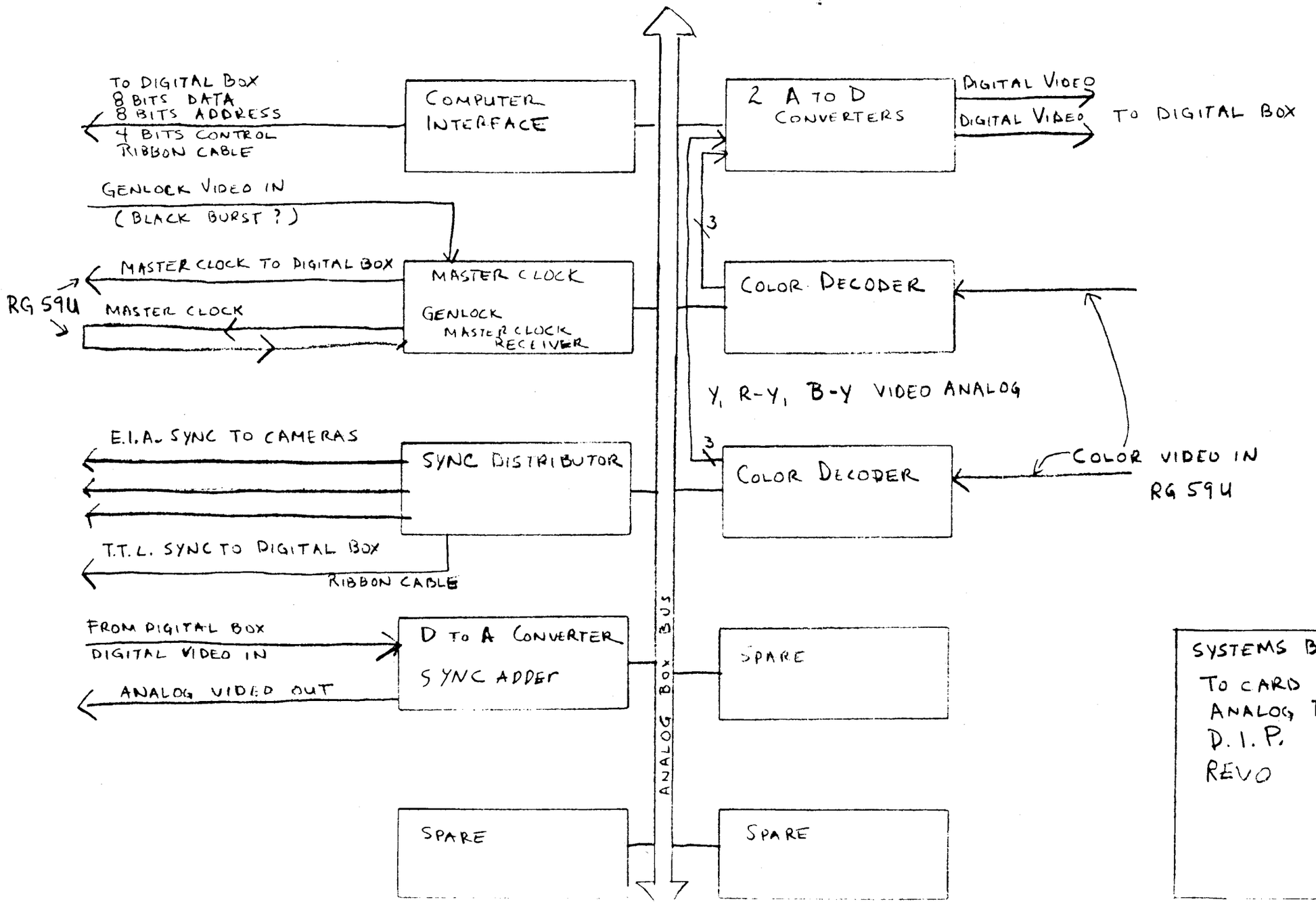
LARGEST SCALE BLOCK SYSTEMS DIAGRAM D.I.P.

REVISION O



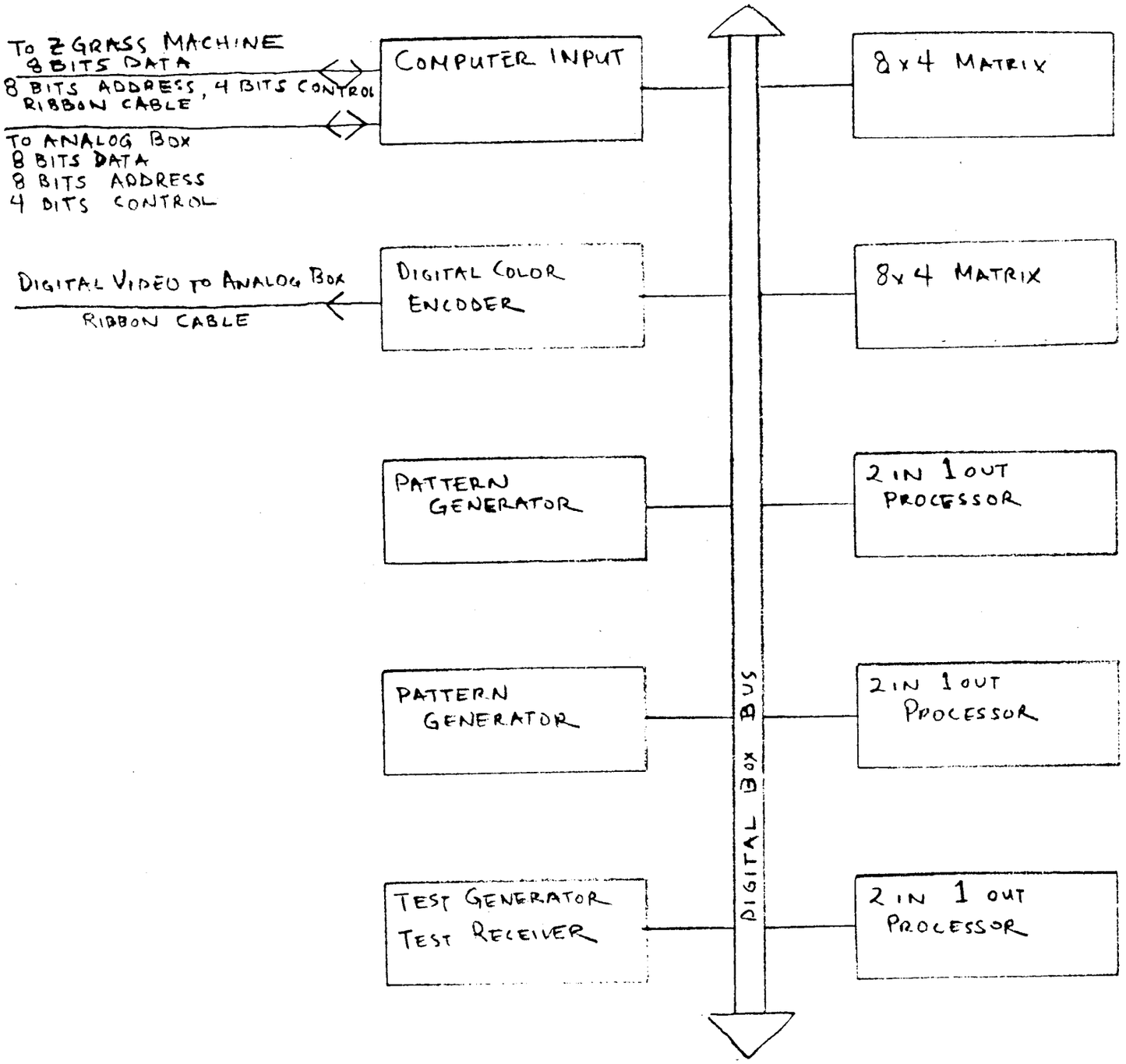
MATRIX INTERCONNECTION DIAGRAM
PHASE 1 D.I.P.

RIVO

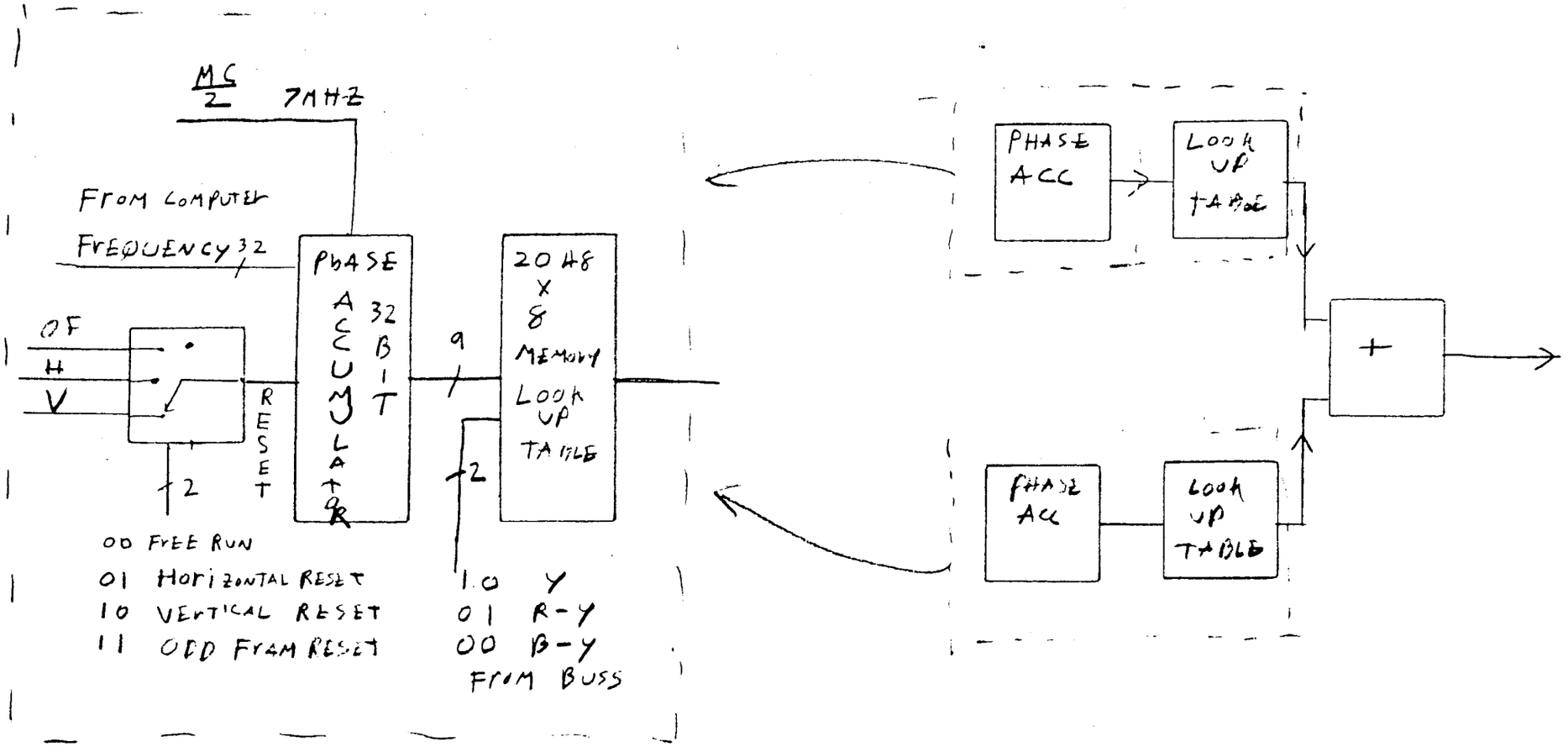


SYSTEMS BLOCK DIAGRAM

TO CARD LEVEL
ANALOG BOX
D. I. P.
REVO



SYSTEM BLOCK DIAGRAM
TO CARD LEVEL
DIGITAL BOX
D. I. P.
RIVO



PATTERN GENERATOR CARD BLOCK DIAGRAM

RIVO