COMPUTERIZED VIDEO IMAGE REDUCTION STUDIES

By

WEN-SHYONG SU

Diploma
Taipei Institute of Technology
Taipei, Taiwan, Republic of China
1964

Master of Science
Wichita State University
Wichita, Kansas
1968

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1971

COMPUTERIZED VIDEO IMAGE REDUCTION STUDIES

Thesis Approved:

Victor W. Bolie
Thesis Adviser

Paul A. McCollum

Joy L. Howard

Bennett Basore

Durham
Dean of the Graduate College

ii

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to Professor Victor W. Bolie, chairman of my committee and thesis adviser, for his original suggestion of the problem. For his encouragement, friendship, many hours of assistance, stimulating discussions, and patient guidance throughout the course of this research, I am truly grateful.

I also wish to express my thanks to Professor Bennett L. Basore, Professor Paul McCollum, and Professor J. L. Howard for their consenting to serve on my committee, for their advice and encouragement, and especially for their teaching excellence.

Further, I wish to express my appreciation to my wife, Shu-Tze, and our child, Juliana, for their patience and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Research on the problem of efficient encoding of the information
content of visual images has been gaining momentum in recent years.
Part of the impetus for this research is due to the increasing accumula-
tion of photographic data such as library information, weather satellite
photographs, fingerprint files, hand-addressed envelopes, medical X-
rays; biomedical photomicrographs, bubble chamber photographs, and im-
ages televised from deep-space vehicles. Continuing efforts in pattern
recognition, pattern classification, photointerpretation, image acquisi-
tion, image processing, image coding, learning machine theory, and video
feature abstraction constitute a significant part of the current growing
subject of artificial intelligence.

A perceptron simulation study performed by Rosenblatt (1) in 1960
represents one of the early works of visual pattern recognition in arti-
ficial intelligence. The perceptron constructed by Rosenblatt is a
nerve-net-like device which consists of a fully random network of signal
generating units and connections between them. The organization of a
perceptron is comparable to a simplified version of the known features
of a mammalian visual system. The perceptron has retinal points direct-
ly coupled to association cells, each of which is called an "A" unit.
Each sensory retinal point is connected to an "A" unit, which is in
either an excitatory (+1) or inhibitory (-1) mode. Each "A" unit has a

fixed threshold and delivers an output pulse if the algebraic sum of the input signals is equal to or greater than a given level. During a training phase, the weighting factors assigned to each "A" unit are adjusted such that during the test phase the total signal output delivered by a set of the "A" units approaches the desired output. The total value distributed over all the "A" units is constrained to be a constant so that one unit can only gain parasitically at the expense of the other units. The perceptron system model was simulated by a digital computer. For a suitable control against a trivial or ambiguous result, the perceptron system was required to respond only to the "actual form" of the stimulus, not its location on the retina.

Widrow and Hoff (2) described in 1960 an adaptive pattern classifier, namely "Adaline" for adaptive linear, which is mainly composed of a combinatorial logical circuit. The machine was used to illustrate adaptive behaviour and artificial learning. The authors' objective was the minimization of the average number of errors, not the number of logical components used. As a resemblance of a "neuron" model, this system has n binary signal inputs (each having a value of +1 or -1) each with an adjustable weight or gain, and each furnishing an analog input to a summing amplifier which also receives a subtractive bias-level input. The output of the summing amplifier serves as the input to a quantizer which has only a single two-level output, the binary value of which is determined by an assigned quantizing threshold. For a given pattern, the desired output is forced by changing in a stepwise iterative searching manner all gains, including the bias-level, by the same absolute magnitude to gradually bring the error (between the desired output and the non-quantized summer output) to zero. The manual

iterative searching routine is repeated for all the training patterns, and as the authors claimed, this process would always converge statistically. The adaptation amounts to searching an n-dimensional parabolic stochastic surface for a minimum. The experiments with the "Adaline", conducted with several 3 by 3 patterns with ten percent noise, showed an average measured misadjustment (the ratio of the mean increase in mean square error to the minimum mean square error) of 91.5 percent. The authors concluded that the number of patterns required to train an adaptive classifier is equal to several times the number of bits per pattern.

Kazmierczak and Steinbuch (3) mentioned in 1963 that the existing knowledge of the human visual system should be considered in designing a mechanical perceptual system for the task of complex pattern recognition. They compared the properties of both systems, such as learning, recognition, motion, sensing, and structure. The visual system can readily respond to external teaching and can perform internal adaptation, abstraction, and learning. The authors felt that a pattern recognition system must generate a feature set (x_i) from a pattern, and classify this feature set into the pattern class code (c_j). However, the realization of internal learning processes by machines was found to be much more difficult than that of the external learning process. The reason for this is that no effective network exists at present to generate features which are invariant with respect to position, skew, size, contrast, and deformation. For the purpose of pattern classification, the authors introduced a learning matrix which, in the form of circuit structure, consists of column inputs (a feature vector) and row outputs (the category) each connected with positive or negative variable weights

as commonly used in adaptive network schemes. The performance of the learning matrix is that either the most similar class (c_i) will be excited upon applying an arbitrary set of features (x_j), or, conversely, the stored feature set (x_j) will be excited upon application of a pattern class signal (c_i). Several examples of two-dimensional features were illustrated to compare the results of pattern classification obtained by using different types of adaptive categoriziers. Kazmierczak and Steinbuch commented that the problem of pattern classification seems to be solved by means of the learning matrix, a linear decision function, or other techniques, but the problem of feature generation is still unsolved.

Bolie (4) developed in 1965 a cognitive learning algorithm based on "alphabets" of linearly independent "symbols". According to the algorithm, a machine can learn, within prescribed limits, to recognize and respond properly to various stimuli. The machine stores an "alphabet" in the form of a square matrix of column vectors each vector of which represents a linearly independent symbol. During the training program, the matrix is inverted. The matrix-inverse is used to premultiply (multiply on the left) any signal vector to be identified during the test program. An averaging process for a series of exposures to a distorted alphabet improves the signal-to-noise ratio, by an amount which is proportional to the square root of the number of independent exposures. The author also derived a formula which showed the errors due to insufficient training and the errors due to noise in the test exposure. Various cases with noise are illustrated with a specific example indicating that the deviation from the ideal result is less than ten percent in the case of moderate noise. It was found that a

near-perfect separation of the signals can be achieved if a sufficient number of samples are employed during training, if the various signal vectors are reasonably orthogonal, and if a large number of input stimuli are used in the test program.

Bolie (5) also developed, from a more generalized point of view, a machine learning algorithm by which a machine, designated by the author as "the $\Omega(M,N)$ machine", can be trained to adaptively search for the most probable category of a given input stimulus. The input stimulus is represented in a vector form of $\underline{n}$ analog components while the output response is generated in a similar vector form of $\underline{m}$ analog components, with the relationship between $\underline{m}$ and $\underline{n}$ being immaterial. It makes no difference as to whether the input stimuli are linearly independent. The process includes provision of a set of $\underline{m}$ different "reference vectors", each of which begins as a set of random numbers but is ultimately intended to become the mean of a statistical sampling of a given category of input stimulus vectors. Associated with each reference vector, a "tolerance vector" is used to allow for moderate imperfection in the input vector. A discriminant is made by performing the "analog NOR transform" on the elements of a "discrepancy vector" generated by the stimulus, reference, and tolerance vectors. The adaptation of the machine amounts to, after each exposure to an input stimulus vector, algebraically adding a suitably computed incremental vector to a certain one of the reference vectors, and adding another incremental vector to its corresponding tolerance vector. The experiments were conducted with 13 different categories of patterns, each of which consisted of 11 samples, and they demonstrated that the output response in the training phase tended to mutate toward a corresponding unit vector as successive

stimulus samples were sensed. The experimental results of exposing the trained machine to a sequence of unidentified input stimulus vectors also demonstrated the performance of automatic identification of the stimuli, and of automatic adaptation of the reference, and tolerance vectors. According to the author, the $\Omega(M,N)$ machine can serve effectively as the "core" of a more intelligent machine, particularly if some improvements can be made, for example, in the saving of computer time.

The application of spatial filtering theory in the field of pattern recognition has been developed by many investigators in recent years. Cutrona, et al. (6) employed in 1959 a coherent optical system, which might be composed of lenses and film planes, to obtain the two-dimensional Fourier transform of any given light amplitude distribution. The system has the property that the relative phases of the light waves in various parts of the system are invariant with time. An optical filter is made by inserting transparencies at appropriate positions in the system. As amplitude and phase control is available over two-dimensional regions in both the spatial and frequency domains, a wide variety of filter functions can be synthesized. In 1962, Montgomery and Broome (7) applied the spatial filtering theory to the enhancement of the detection of two-dimensional localized objects superimposed on a noisy background. With the help of the so-called "neighborhood modification process", they derived four types of matched spatial filters such as the linear, quadratic, general statistical, and decision filters. The spatial filters discussed were simulated by a digital computer. Vander Lugt (8) made in 1964 a further study of the spatial filtering system by considering the case of detecting isolated signals

in a variety of noise backgrounds (non-uniform distributed noise backgrounds). His technique was based on the use of a Mach-Zehnder interferometer in the realization of an optimum filter. The mathematical theories of spatial filtering were fully derived by the aforementioned three groups of investigators. The use of spatial filtering played a major role in the development of the modern laser-projected hologram patterns.

Tretiak, Eden, and Simon (9) described a technique for obtaining the internal structure of an object from its X-ray images. The method is mainly based on the possibility of evaluating the three-dimensional Fourier transform of the extinction function associated with an object by taking a collection of X-ray photographs while the object is rotating about a fixed axis, and by using the Fourier transform function thus obtained to find the two-dimensional Fourier transform of a function linearly related to the extinction function.

Howell (10) investigated an optical cross-correlation method for two-dimensional pattern recognition. In this method, two variably transparent planes, each with a pattern of interest, are separately located between an ideal planar diffused light source and an observation transparency plane. The entire correlation function between a pair of patterns is obtained in the form of a light intensity distribution on the observation plane. The author made a detailed comparison of theoretical predictions with experimental cross-correlation results for numerous simple geometric patterns, and concluded that this technique may be used in the implementation of an automatic reading machine.

Morris (11) developed a computer-controlled video system for video image acquisition. An image interrogator was designed to serve as the

interface between a standard television system and a modern digital computer. Software programs were generated for image scanning, data collection, and to control the size and location of a sampling "window". The system was used to obtain the quantized and filtered images from the photographic images of human faces.

Nagy (12) presented in 1968 a wide-range survey which reviewed the statistical, adaptive, and heuristic techniques used in laboratory investigations of pattern recognition problems. Levine (13) made a survey of computer algorithms and philosophies applied to problems of feature extraction and pattern recognition in conjunction with image analysis.

A collection of recent works contributed by many investigators in the areas of character recognition, pattern perception and analysis, and statistical classification procedures can be found in the book edited by Kanal (14). Another book edited by Cheng, et al. (15) consists of many papers in the areas of picture processing, learning machines, automatic target recognition, and automatic photointerpretation in the biomedical sciences, in the environmental sciences, and in high energy physics.

The literature review (1-15) shows that little has been effectively done in the task of reducing the number of bits in an image without sacrificing its information content. However, in spite of the fact that normally a large number of bits of computer memory is required to store a whole image in a computer, there exists the possibility of minimizing the image dimensionality by means of preprocessors which incorporate high-speed video systems.

The purpose of the research reported in this thesis has been to conduct theoretical and experimental investigations on the reduction of

video images with aid of portable television camera-and-recorder equipment, supplemented by specially designed auxiliary electronic subsystems. Following a detailed description of the basic video equipment, the design and operating characteristics of a video cross-correlating sub-system and a video waveform-agreement-detection sub-system are presented. Each of these two sub-systems has the property of converting a camera-sensed input image into a corresponding feature vector. A mathematical algorithm for the reconstruction of a video image from its feature vector is then presented, following which some experimental results are shown and discussed.

CHAPTER II

THE BASIC VIDEO APPARATUS

The basic video apparatus used in this investigation is a portable commerical closed-circuit television system, which includes the Sony AV-3200DX Video Camera Ensemble, the Sony AV-3600 Video Tape Recorder (Video-corder), and the Sony CVM-110U Video Monitor. Accompanied by a special synchronization unit, this composite ensemble of television equipment is used as the source of the video signals employed in the research reported here. This chapter is devoted to the description of the television equipment, the composite television signals, the special synchronization unit, and the video-apparatus connections.

The Commercial Television Equipment

The Sony AV-3200DX Video Camera Ensemble mainly consists of a video camera (AVC-3200) with a zoom lens, and an electronic viewfinder (AVF-3200) mounted atop the camera to permit the operator to see the actual picture being produced. Figure 1 shows the actual video camera and viewfinder as interconnected. The corresponding block diagrams of the camera and viewfinder are shown in Figures 2 and 3, respectively.

The Sony AVC-3200 camera has a horizontal resolution of more than 400 lines. Its synchronism system can operate with either internal or external synchronization (abbreviated as sync) pulses. The internal synchronism is vertically locked into the 60 Hz power-line source, and
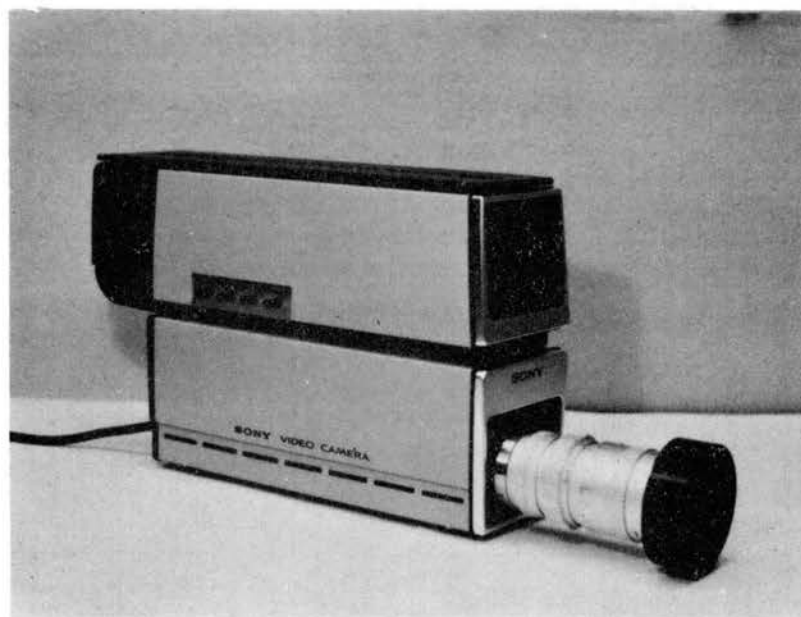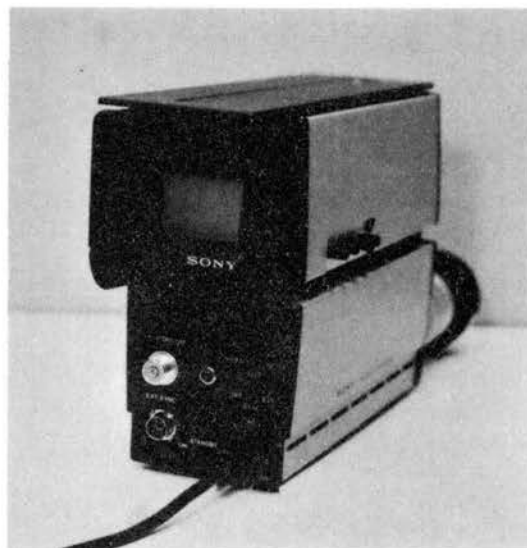
Figure 1.   Sony AVC-3200DX Video Camera Ensemble

Figure 2. Sony AVC-3200 Video Camera Block Diagram
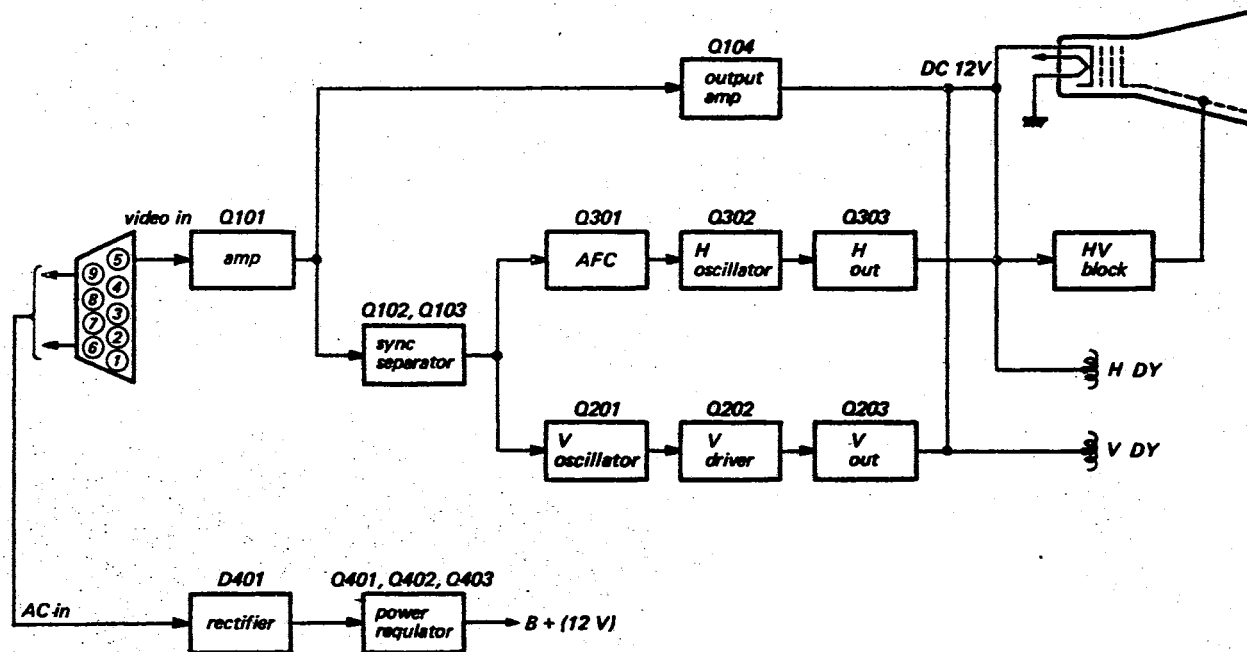
Figure 3.   Sony AVF-3200 Video Viewfinder Block Diagram

performs with random horizontal interlace. In case the external vertical and horizontal sync pulses are supplied to the camera, the camera synchronism performs with two-to-one interlace. The feasibility of the camera external synchronism is a significant feature of this equipment. As it will be seen later in the next chapter, the experimental studies reported here take significant advantage of the capability of maintaining the synchronism between the video camera and the videocorder.

The camera video output is supplied in the form of a one-volt (peak-to-peak) composite video signal from a 75-ohm source. The camera video signal amplitude for a fixed image can be adjusted by properly changing the aperture of the zoom lens. A selected part of the characteristic specifications of the Sony AVC-3200 video camera is listed in Table I. A detailed description of the video camera is referred to in the service manual supplied by the manufacturer (16).

The Sony AV-3600 Videocorder permits automatic recording of the picture and its associated sound through the use of the Sony video camera and microphone. It can also record the video and audio signals directly from the video monitor, when the monitor is used as an ordinary television receiver. When operating in the playback mode, the videocorder provides the video and audio information of the picture and sound pre-recorded on the tape. Figure 4 shows the photographs of the Sony AV-3600 videocorder; its corresponding block diagram is shown in Figure 5.

The video information is recorded on the tape through two rotating video heads in the form of a frequency-modulated carrier. When the videocorder operates in the playback mode, the video image is recovered by means of an FM demodulator. There is a magnetic-brake servo system

TABLE I

SONY AVC-3200 VIDEO CAMERA CHARACTERISTIC SPECIFICATIONS

| Characteristic	Specification
Scanning System	525 lines per frame, 30 frames per second.
Sync System	Internal sync: vertical line-lock (60 Hz) sync with random interlace.
	External sync: EIA standard, 2:1 interlace.
Horizontal Resolution	More than 400 lines at center.
Horizontal Frequency	15.75 K Hz.
Vertical Frequency	60 Hz.
Signal-to-Noise Ratio	Greater than 42 db.
Video Bandwidth	6 M Hz.
Video Output	1 volt (p-p) composite video signal, synce negative, 75 ohms, unbalanced.
Automatic Sensitivity Control Range	30 to 10,000 footcandles (with f 1.8 lens opening).
Power Requirements	117 volts, ac, 60 Hz.
Zoom Lens	f/2, 16-64 mm
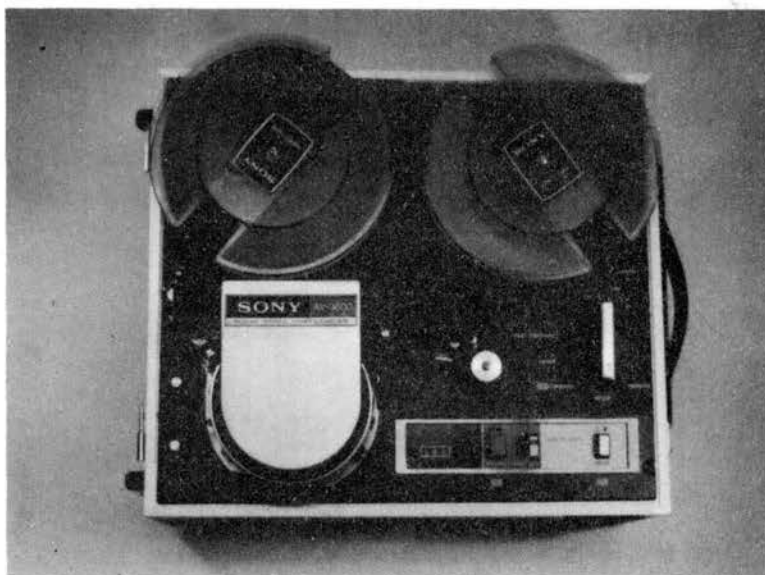
Figure 4.   Sony AV-3600 Video Tape Recorder
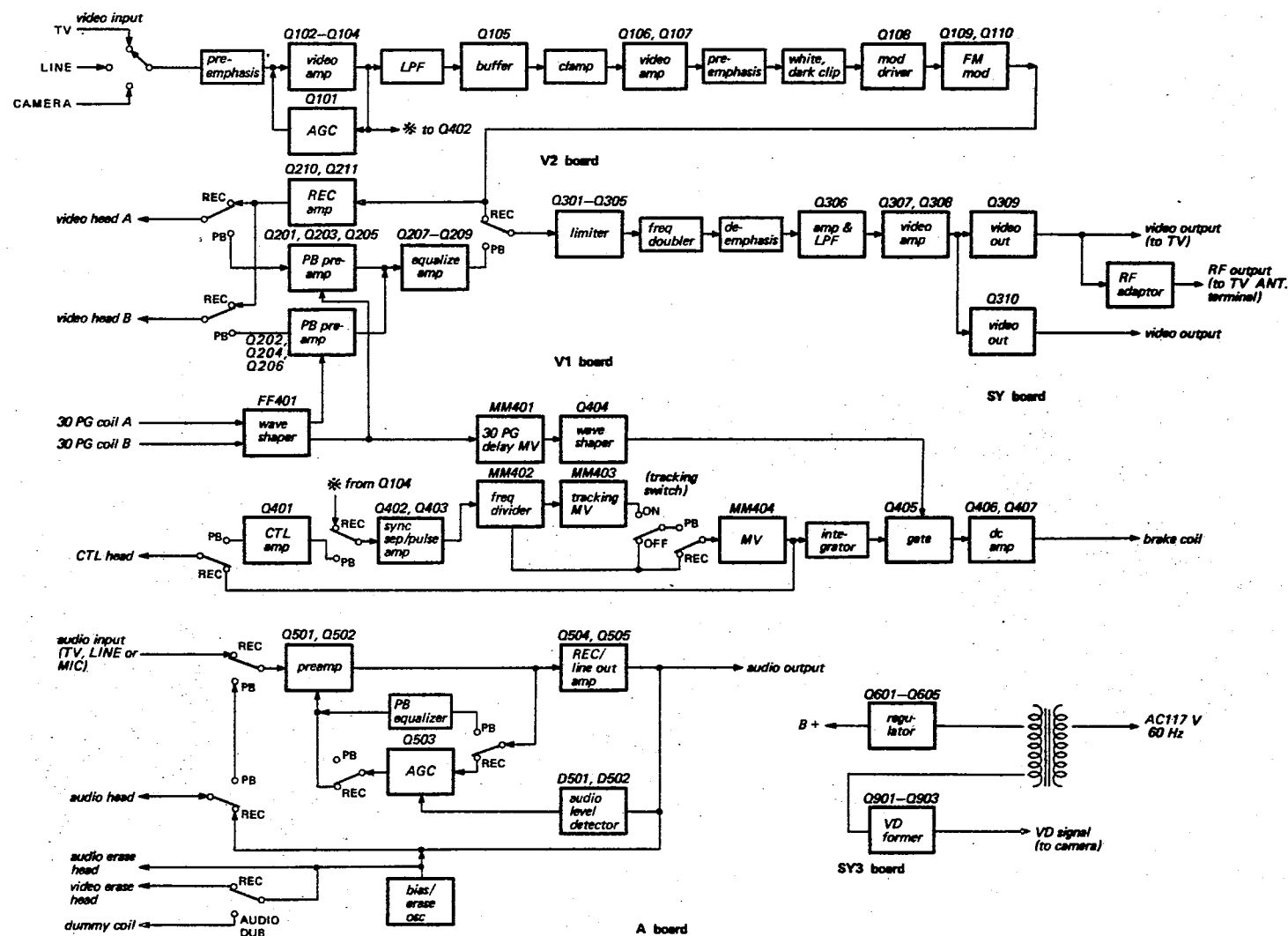
Figure 5. Sony AV-3600 Videocorder Block Diagram

which controls the rotational speed and angular position of the video heads. During the record operation, separated 60-Hz sync signals serve as the timing reference for the servo system. However, in the playback mode, the timing reference for the servo system is obtained directly from the recorded control-track pulses.

The Sony AV-3600 videocorder has a horizontal resolution of more than 300 lines. When the videocorder operates either in the record mode or in the playback mode, the one-volt (peak-to-peak) video output composite signal is available from a 75-ohm source. A selected part of the characteristic specifications of the videocorder is listed in Table II. The detailed circuit description, characteristic specifications, and operating instructions can also be found in the Sony service manual (17). A useful feature of the videocorder is that the RECORD button is not released in the PAUSE position. This feature provides a convenient way to sequentially record stationary images of interest.

The Sony CVM-110U Video Monitor can display the picture and sound information during both record and playback operations. It can also receive both VHF and UHF channels defined by American TV Standards. In the studies reported here, the CVM-110U monitor is used, not only as a display apparatus, but also as a source to obtain the sync pulses which are needed for the purpose of establishing and maintaining synhcronism between the videocorder and the video camera. Figures 6 and 7 show respectively the photographs and the block diagram of the monitor. The circuit details and characteristic specifications of the CVM-110U monitor are referred to in the Sony service manual (18).

TABLE II

SONY AV-3600 VIDEOCORDER CHARACTERISTIC SPECIFICATIONS

Characteristic	Specification
Video Signals	
Input	0.5 - 2 volt (p-p), sync negative, 75 ohms unbalanced.
Output	1.0 volt (p-p), sync negative, for a 75-ohm load, unbalanced.
Resolution	More than 300 lines.
Signal-to-Noise Ratio	More than 40 db.
Video Recording System	Rotary two-head slant-track scanning, 2:1 interlace composite video signal, FM recording.
Tape Transport	
Tape Speed	7.5 inches per second.
Recording Time	60 minutes using SONY V-32 tape. 30 minutes using SONY V-31 tape.
Audio Signals	
Input	Mic: -65 db, 600 ohms, unbalanced.
Output	Line: 0 db, 10 K ohms, unbalanced.
Frequence Range	80 - 10,000 Hz.
Signal-to-Noise Ratio	More than 40 db.
Power Requirements	117-volt, ac ($\pm$ 10%). 60 Hz ($\pm$ 0.5%).

Figure 6. Sony CVM-110U Video
Monitor

TV → Q BOX

CRT

Q501 EF
VIDEO OUT

Q601
SYNC SEP

SYNC OUT

Q2651
VIDEO AMP

Q2652
PULSE AMP

Q2653
PULSE AMP

Q2654
PULSE AMP

Q2655, 2656
FLIP FLOP

GATE SIGNAL OUT

Q2601
VIDEO AMP

Q2602
SYNC SEP

EXT TV VIDEO IN

D301
VIDEO DET

Q401
VIDEO DRIVE

TV AUDIO IN

Q2502, 2505
VIDEO GATE

GATE SIGNAL IN

Q2503, 2504
VIDEO AMP

Aperture correction circuit

Q2506
VIDEO OUT

Q2501
VIDEO AMP

VIDEO OUT

D402

BC

D403

FM DET

VIDEO OUT

AUDIO OUT

TV AUDIO IN

Q2551
AUDIO AMP

Q2552, 2563
AUDIO GATE

SP

Q551 S
AUDIO AMP

EXT AUDIO IN

GATE SIGNAL IN
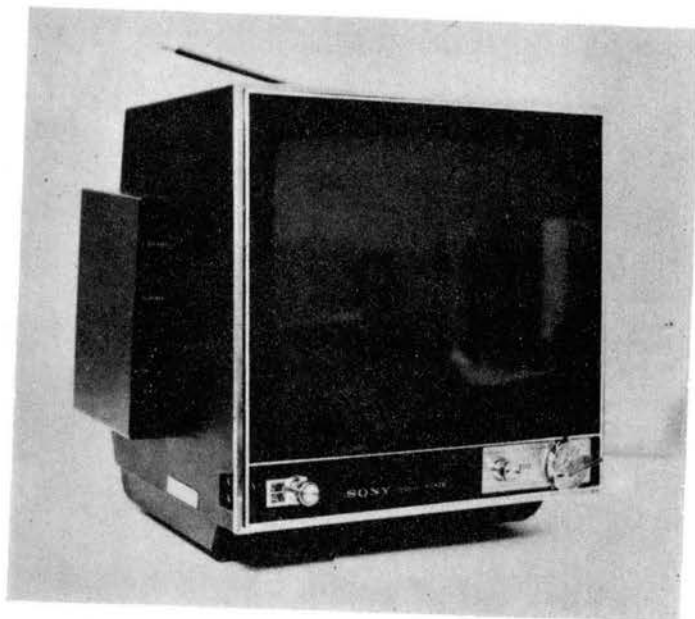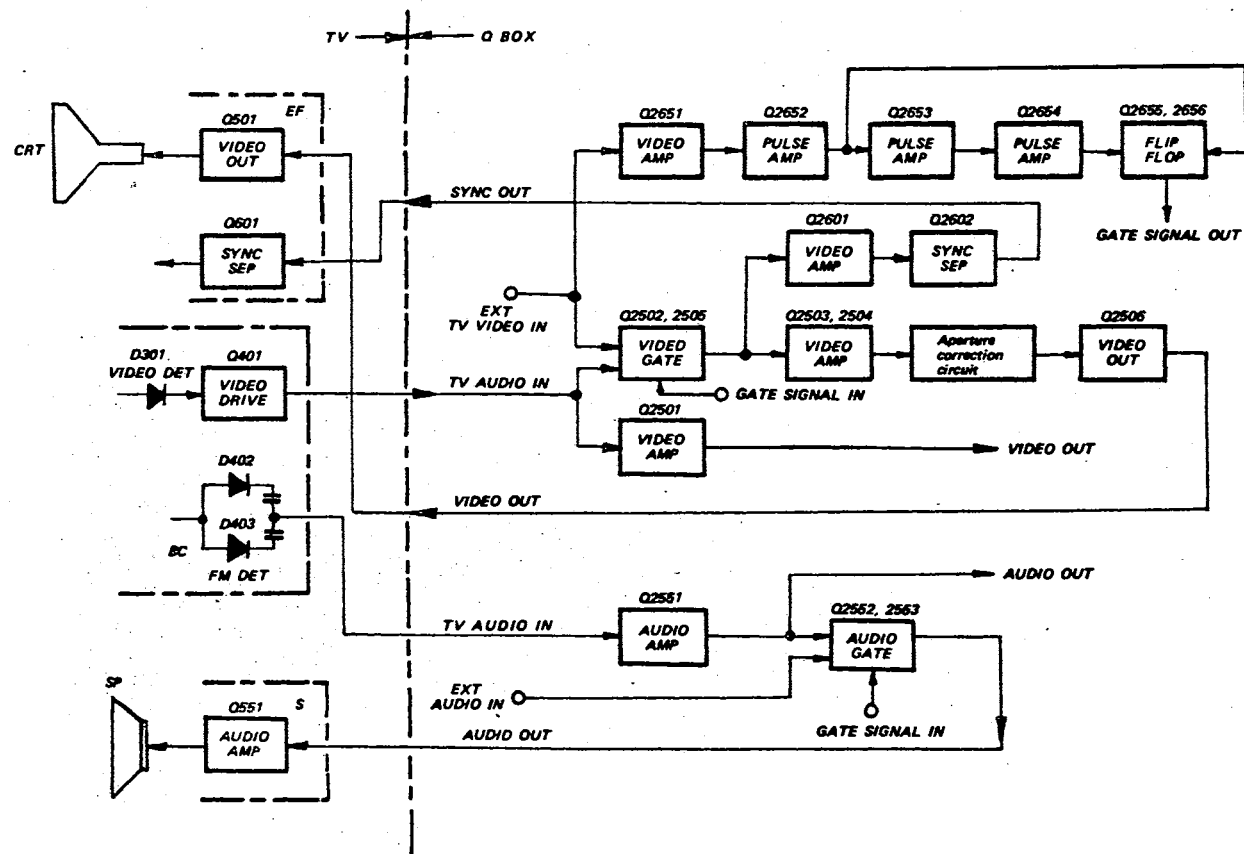
AUDIO OUT

Figure 7. Sony CVM-110U Video Monitor Block Diagram

The Composite Video Signals

According to the basic principle of television systems (19), the coding of two-dimensional images into electronically represented information is accomplished by time sampling schemes. Of all the time sampling schemes which have been developed, the "two-to-one" interlaced scanning technique is commonly used in television systems. Figure 8 shows a simplified presentation of the two-to-one interlaced scanning.

The composite video signal of an image is composed of its electronically represented illumination signal (a function of time) superimposed by both vertical and horizontal periodic sync and blanking pulses. A simplified waveform for a television signal is shown in Figure 9.

In the studies reported here, a composite video signal is available from each of the two 75-ohm sources of the video camera and the videocorder. The following description of the composite signal is based on the use of the Sony video apparatus. Two different images are illustrated. The corresponding waveforms are measured by using an oscilloscope.

A typical image and one part of the corresponding horizontal waveform are shown in Figure 10. This special image contains the information of white in the left-half part and black in the right-half part. The white part of the image corresponds to the white level shown in the waveform, which is about +0.5 volt with respect to the blanking level. The blanking level is actually the ideal black level. Since the right-half part of the image is not purely black, its signal level is +0.08 volt more than the blanking level.

The horizontal sync pulse, supplied by the Sony apparatus is 0.5
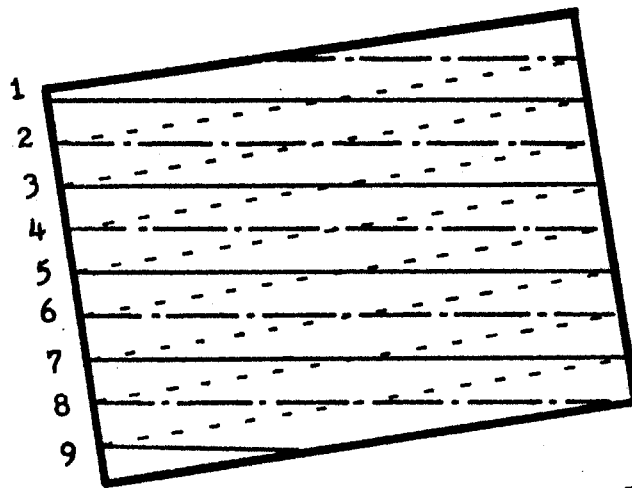
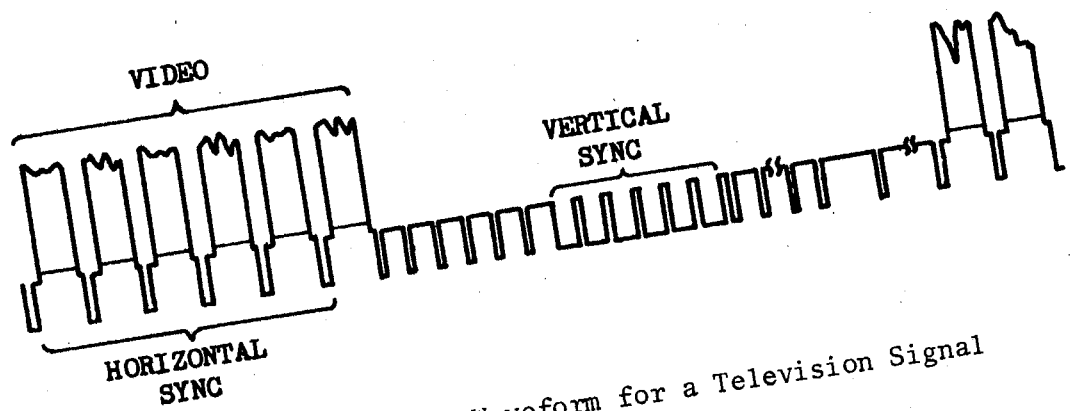Figure 8. Simplified Presentation of Interlaced
           Scanning

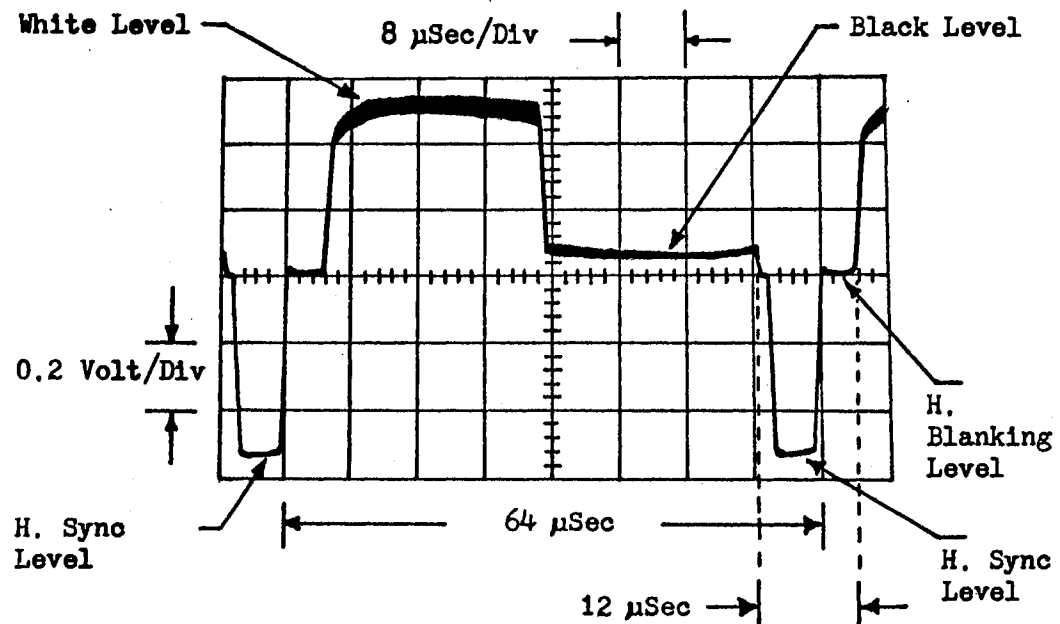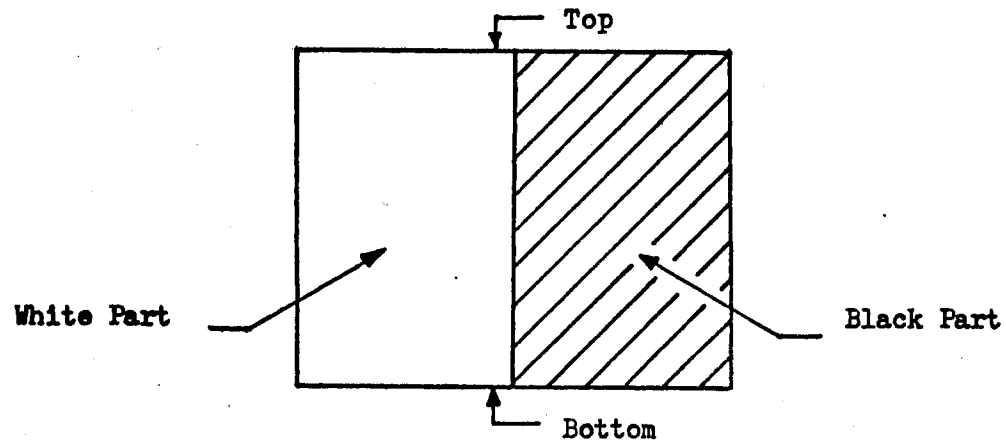Figure 9. Simplified Waveform for a Television Signal

Figure 10. An Image and Its Horizontal Waveform

volt negative with respect to the blanking level. The time period of the horizontal sync pulses is approximately 64 microseconds, and corresponds to the horizontal frequency of 15,750 Hz. Of the 64 microseconds, only 52 microseconds are used to transmit information; the rest of the time (12 microseconds) is employed for the horizontal synchronism and blanking. The horizontal sync pulse is 5 microseconds in duration.

Another example for illustrating the vertical signal representation is shown in Figure 11. This typical image consists of two parts, white in the upper half and black in the lower half. The white level and the black level shown in the signal waveform correspond to the white part and the black part of the image, respectively. The negative vertical sync pulse is also 0.5 volt in amplitude with respect to the blanking level. The vertical period is 16.7 milliseconds, of which 1.0 millisecond is used for the vertical synchronism and blanking. A typical vertical sync pulse is 0.2 millisecond in duration. In Figure 11, the negative peaks of the many (262.5) individual horizontal sync pulses are represented by the almost continuous thin horizontal line having the same level as the peaks of the vertical sync pulses.

For both cases of the videocorder and the video camera, the sync level (peak negative excursion of the sync pulses) is fixed regardless of the variation of different video signal levels.

The Video Apparatus Connections

The cabling diagram of the video apparatus operating in the record mode is shown in Figure 12. For the recording of images, a six-conductor cable, namely the camera cable, is used to connect the connector CN4 on the videocorder to the connector CN303 on the video
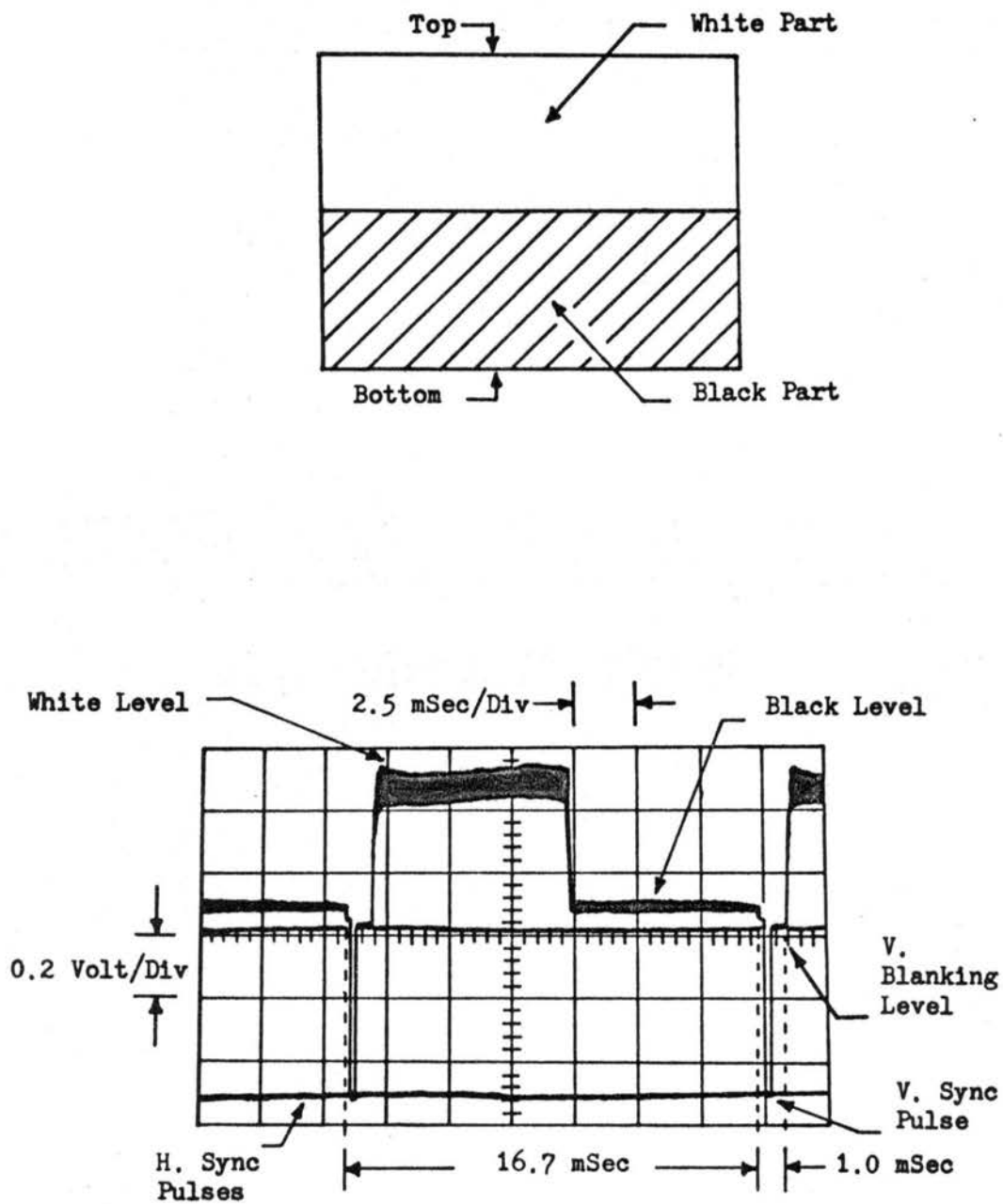
Figure 11. An Image and Its Vertical Waveform

SONY AV-3200DX
CAMERA ENSEMBLE

CN302

CN303

75-OHM
COAX

6-WIRE
CAMERA
CABLE

○ CAMERA VIDEO OUTPUT

○ RECORDER VIDEO OUTPUT

75-OHM COAX

8-WIRE MONITOR
CONNECTING CABLE

SONY CVM-110U
VIDEO MONITOR

CN2502

MICROPHONE

CN5   CN4   CN9   CN3

SONY AV-3600
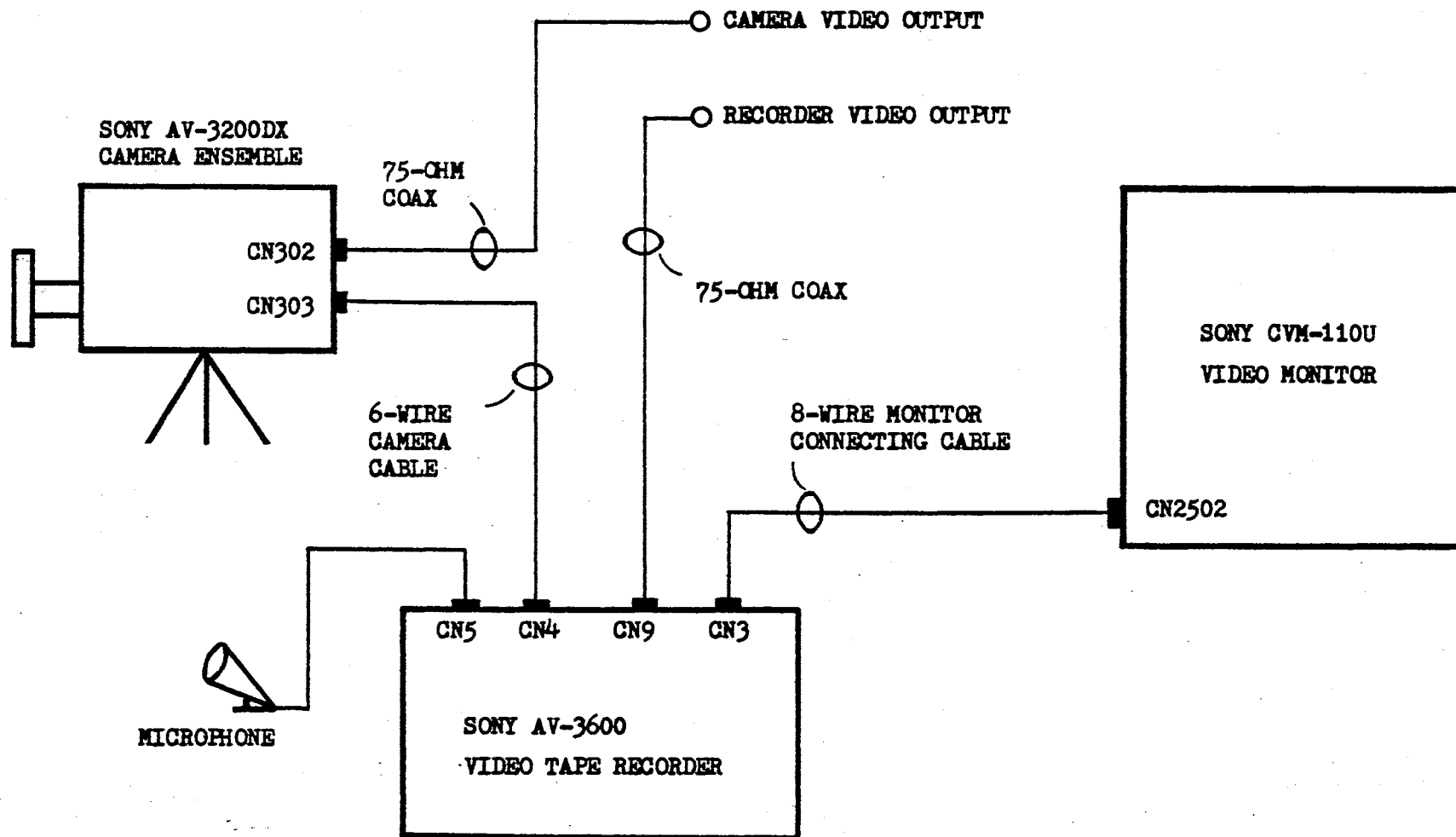·VIDEO TAPE RECORDER

Figure 12. Video-Apparatus Connections for the Record Mode

camera. The video signal is supplied from the video camera to the videocorder through the first conductor in the camera cable; the external vertical and horizontal sync signals are each supplied from the videocorder to the camera through the second and fifth conductors in the same cable. For the recording of sound, a microphone is connected to the mini jack CN5 on the videocorder. A monitor connecting cable, which is composed of eight conductors, connects the connector CN3 on the videocorder to the connector CN2502 on the monitor.

For the studies reported here, it is necessary to observe simultaneously both the video signal of any image pre-recorded on the videocorder tape and the video signal of a test image being sensed by the video camera. To meet this requirement, it is necessary that the synchronism between the video camera and the videocorder be maintained when the videocorder operates in the playback mode, and that the video camera be set to operate with external sync pulses. Such a video apparatus operation will be called herein the special playback mode.

Figure 13 shows the cabling diagram for the video apparatus operating in the special playback mode. In addition to those components shown in Figure 12, Figure 13 has a special sync unit which will be described in the next section. The 8-conductor monitor-connecting cable is still in the same position. However, the 6-conductor camera cable in which only three conductors (vertical sync, horizontal sync, and common ground) are actually used, connects the video camera to the output of the special sync unit, rather than to the connector CN4 on the videocorder. This means that the external sync signals are supplied to the video camera from the special sync unit.

As shown in Figure 13, the three banana plugs identified by the
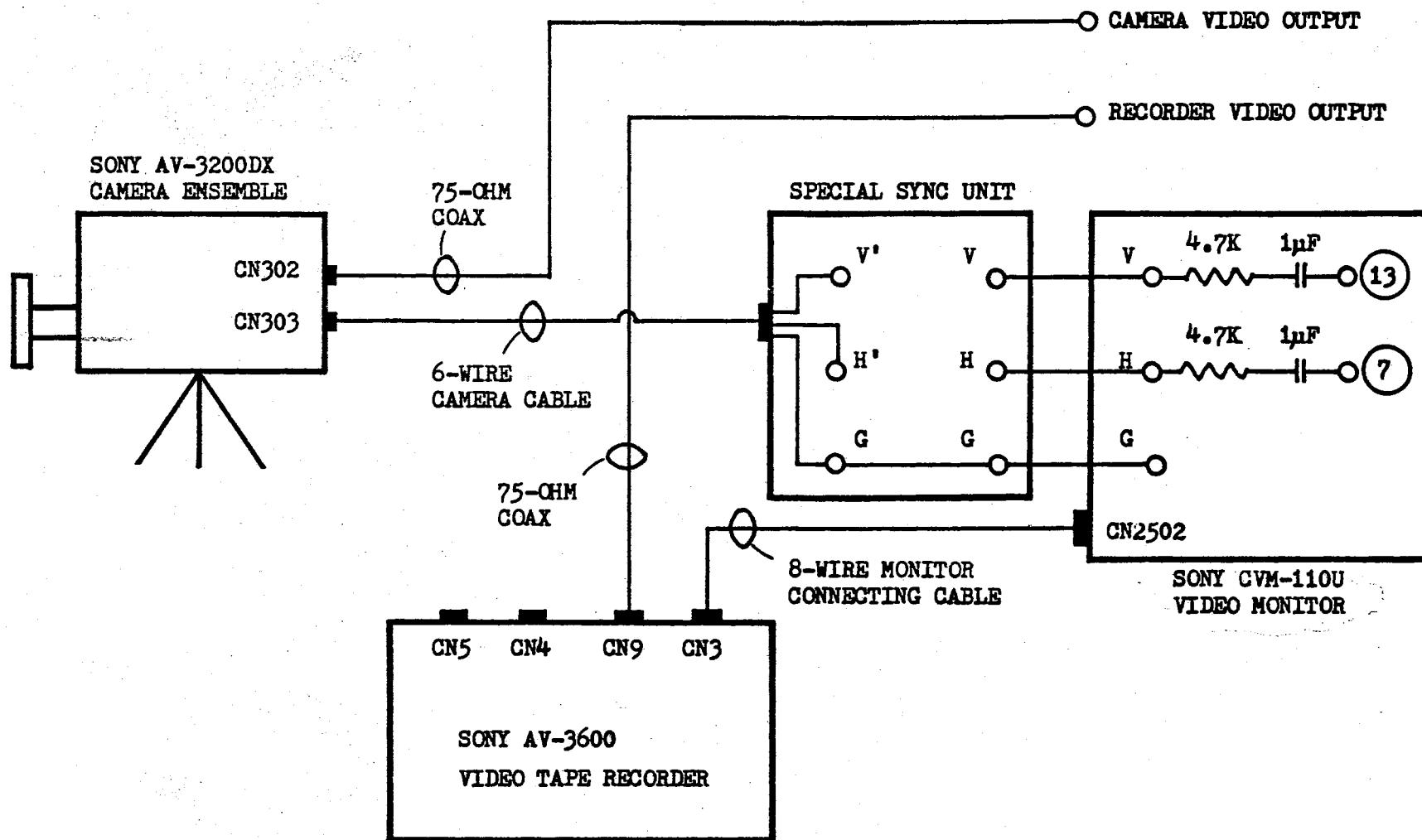
Figure 13. Video-Apparatus Connections for the Special Playback Mode

capital letters V, H, and G are additionally mounted on the back cover of the monitor, as a modification. The V refers to the vertical sync, H to the horizontal sync, and G to the common ground. The V and H receive their excitations from the test points 13 and 7, respectively, on the EF board of the monitor, each through an added 4.7-kilohm resistor in series with an added 1.0 micro-farad capacitor, as shown in Figure 13. The signals supplied by the monitor plugs V and H are each fed to the special sync unit through two individual conductors. The other conductor connects the ground terminals of the special sync unit and the monitor. To prevent electric shock, the 750-kilohm resistor R903 located inside the monitor immediately behind the 117-volt power source plug was removed (the reason for its presence is unknown).

In both Figures 12 and 13, the camera video output is supplied through a 75-ohm coaxial cable extending from the connector CN302 on the camera, and the videocorder video output is obtained through a similar 75-ohm coaxial cable extending from the connector CN9 on the videocorder.

In Figure 12, when the videocorder operates in the record mode, its output waveform is identical to that of the video camera. However, during playback operation, the videocorder video output is not synchronized with that of the camera if the entire apparatus is cabled as shown in Figure 12. The synchronism between the videocorder video output and the camera video output can be established if the apparatus is cabled as shown in Figure 13.

The Special Synchronization Unit

A preliminary experiment showed that if the video appartus connections are made as shown in Figure 12, if the videocorder operates in the playback mode, and if the camera operates with either external or internal sync pulses, the camera video output signals are not synchronized with the videocorder video output signals. The signals are not synchronized because on the videocorder the fifth pin of the CN4 connector, which is supposed to be supplied with the horizontal sync pulses, is open without any connection to the inside sync circuit of the videocorder, as shown in the videocorder schematic diagram (17).

As described in the last section, the separated vertical and horizontal sync pulses are available from the V and H plugs installed on the videocorder. The vertical and horizontal sync signals are each separated from a composite sync signal extracted from the composite video signal supplied by the videocorder through the 8-conductor monitor connecting cable. However, if the horizontal sync signal obtained from the H plug is directly fed into the fifth pin of the connector CN303 on the video camera, then the horizontal sync signals contained in the camera output composite signal will be delayed by 10 microseconds with respect to the original horizontal sync signals supplied by the videocorder. This 10-microsecond delay is due to the function of the Q201 horizontal former and the Q202 pulse amplifier of the video camera. This difficulty is overcome by means of the specially designed sync unit.

The special sync unit consists of a horizontal sync delay circuit and a vertical sync inverter circuit. The circuit details and detailed description of the special sync unit are given in a recent publication

(20) which is reprinted herein as Appendix A.

The Applicable Video Signals

As previously noted the synchronism between the video camera and the videocorder can be established with aid of the special sync unit. This is a significant function of the video apparatus employed in the research reported here. A series of reference images which are known beforehand can be recorded on the video tape of the videocorder. The diagram of the video apparatus connections for recording images is already shown in Figure 12. After recording, the video apparatus connections are changed to the ones shown in Figure 13. When the entire video apparatus operates in the special playback mode, the pair of synchronized individual video signals are available from the videocorder and the video camera.

A typical example of the synchronized video signals is shown in Figure 14. The upper waveform corresponds to a camera test image of white in the left and black in the right, and the lower waveform corresponds to a videocorder pre-recorded reference image of black in the left and white in the right.

The pair of synchronized signals, available from the video camera and recorder operating in the special playback mode, are used as the fundamental signals in the experiments to be described in the following chapters.

20μSec/Div →| |←

0.5mVolt/Div
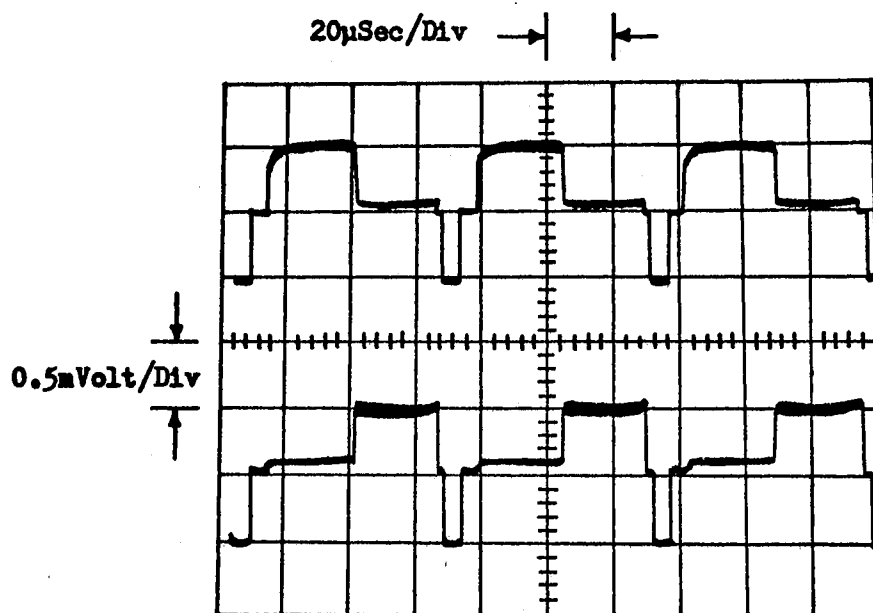
Upper trace: White left, black right.
Lower trace: Black left, white right.

Figure 14. A Pair of Synchronized Video Signals

CHAPTER III

THE VIDEO CROSS-CORRELATING SUBSYSTEM

The previously described commercial television equipment, when accompanied by the special sync unit, can serve effectively as the source of a pair of synchronized video signals which are applicable to the image preprocessing operations. As a typical preprocessor for image reduction, a video cross-correlation subsystem can be used to generate a series of self-normalized coefficients of cross-correlation between an image sensed by the television camera and a series of images retrieved from the television tape recorder. The purpose of this chapter is to define the cross-correlation coefficients of two-dimensional images, to summarize the design and performance of the video cross-correlating subsystem, and to compare the theoretically calculated and experimentally realized coefficients of cross-correlation.

Image Cross-Correlation Coefficients

The information content of a pair of two-dimensional optical images can be represented by $F_i(x,y)$ and $F_j(x,y)$, in which the subscripts i and j are used to distinguish the images, and in which the denotations x and y correspond to the X-coordinate and Y-coordinate, respectively. By means of television scanning, the two-dimensional functions $F_i(x,y)$ and $F_j(x,y)$ are converted to the corresponding time functions $U_i(t)$ and $U_j(t)$, each of which is equivalent to a corresponding video signal. If

the running-time averages of $U_i(t)$ and $U_j(t)$ are expressed by $\overline{U_i(t)}$ and $\overline{U_j(t)}$, and if the time functions $\alpha(t)$ and $\beta(t)$ are arranged in such a way that

$$\alpha(t) = U_i(t) - \overline{U_i(t)} \qquad (3.1)$$

and

$$\beta(t) = U_j(t) - \overline{U_j(t)} \qquad , \qquad (3.2)$$

then, the self-normalized cross-correlation coefficient of the pair of images $F_i(x,y)$ and $F_j(x,y)$ is denoted by ρ_{ij}, and defined as

$$\rho_{ij} = \frac{\overline{\alpha\beta}}{(\overline{\alpha^2} \cdot \overline{\beta^2})^{1/2}} \qquad (3.3)$$

where $\alpha = \alpha(t)$, $\beta = \beta(t)$, and where $\overline{\alpha\beta}$, $\overline{\alpha^2}$, and $\overline{\beta^2}$ are the running-time averages of $\alpha\beta$, α^2, and β^2, respectively. In Equation 3.3, the denominator $(\overline{\alpha^2} \cdot \overline{\beta^2})^{1/2}$ is used to self-normalize the cross-correlation coefficient ρ_{ij} so that ρ_{ij} falls within the range of $-1 \leq \rho_{ij} \leq +1$. For the case of a pair of identical images, the self-correlation coefficient will be +1, and for the case of a pair of completely opposite images, the cross-correlation coefficient will be -1.

Subsystem Design

The video cross-correlating subsystem (20) is designed to perform the functions which are mathematically expressed by Equations 3.1, 3.2, and 3.3. Figure 15 shows the block diagram of the video cross-correlating subsystem. The arrangement of the video recorder, video monitor, the special sync unit, and the video camera is essentially the
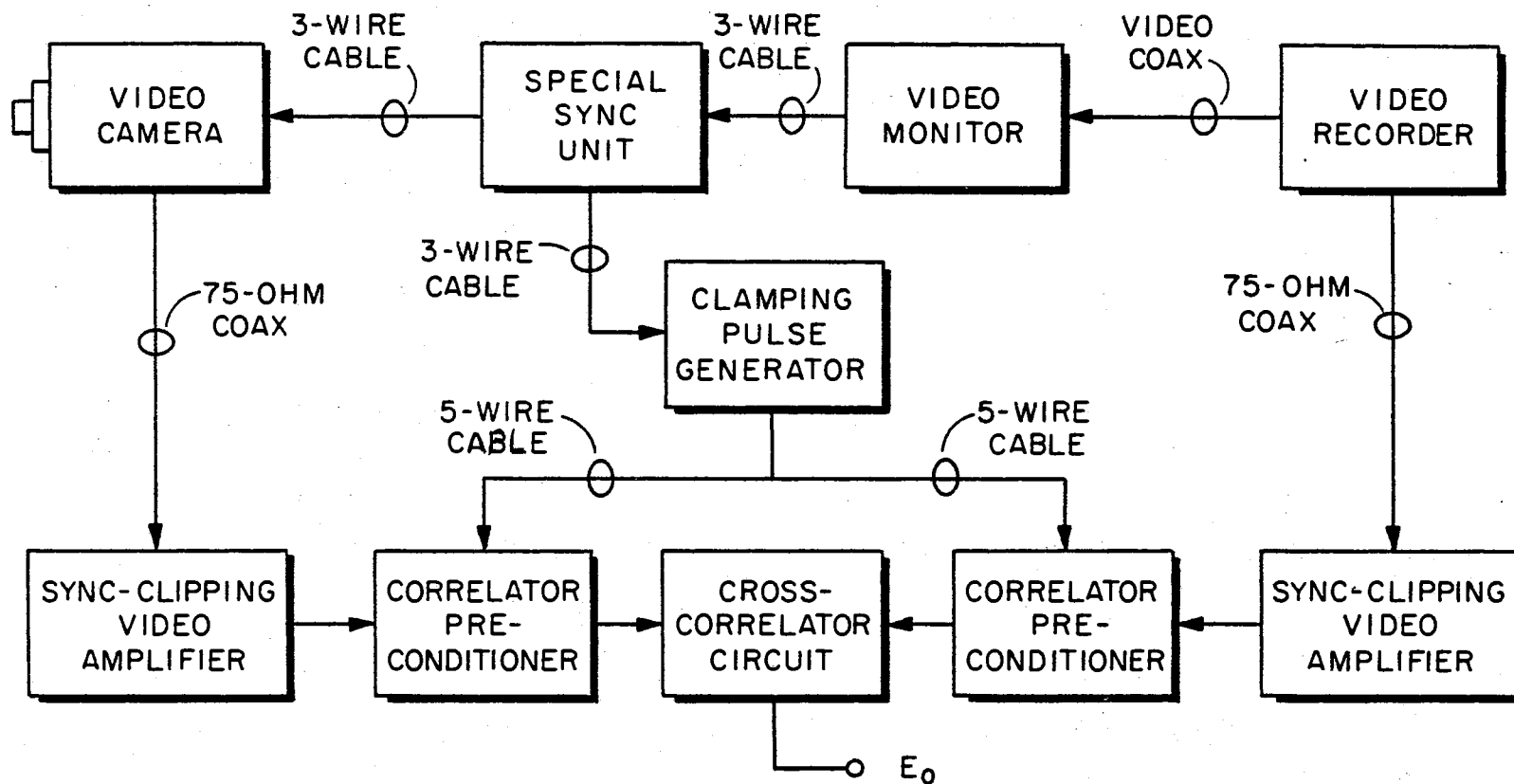
Figure 15. Video Cross-Correlating Subsystem

same as that shown in Figure 13. With this arrangement, the synchronism between the camera scanning beam and the playback video head of the re-recorder can be established and maintained.

The two synchronized ±0.5 volt composite video signals, which are furnished through the 75-ohm coaxial cables from the video recorder and video camera, are each fed to a corresponding sync-clipping video amplifier which removes the sync pulses and amplifies the video to a scale of +15 volts for "black" and -5 volts for "white". The outputs of the sync-clipping video amplifiers are used as the signal inputs to a pair of correlator pre-conditioners. Each correlator pre-conditioner subtracts out the DC-average value of its clipped and amplified video input, and clamps-to-zero any residual signal existing in the blanking-pulse intervals. The clamping pulses used in the correlator pre-conditioners are obtained from a clamping-pulse generator which is controlled by signals generated in the special sync unit. The necessity of the clamping-pulse generator is due to the difference of time-widths between the blanking and sync pulses.

The output signals of the two correlator pre-conditioners are equivalent to the time functions $\alpha(t)$ and $\beta(t)$ as indicated in Equations 3.1 and 3.2. The pre-conditioner output signals α and β are the two signal inputs to the cross-correlator circuit. It is noted that the average values $\bar{\alpha}$ and $\bar{\beta}$ must be zero in order to assure proprer operation of the cross-correlator circuit. The cross-correlator circuit produces the output voltage E_0, which falls within the range of $-5 \leq E_0 \leq +5$ volts, and which is given by the equation

$$E_0 = 5 \cdot \rho_{ij} \quad , \quad \quad (3.4)$$

where ρ_{ij} is the actual cross-correlation coefficient of two corresponding images. The correlator output voltage E_0 thus obtained can be stored into the IBM 1620 computer memory through an analog-to-digital converter, if desired.

As previously noted, the circuit details of the video recorder, video camera, and the video monitor are referred to in the service manuals (16-18). A detailed description of the subsystem design, including the special sync unit, the sync-clipping video amplifiers, the correlator pre-conditioners, the clamping-pulse generator, and the cross-correlator are given in a recent publication (20), which is reprinted herein as Appendix A.

Subsystem Performance

The subsystem performance is summarized by the 8-by-8 matrix arrangement of Figure 16, in which the left column shows the eight different images retrieved sequentially from the recorder and the top row shows the same set of images sensed one at a time by the camera. Each numerical entry in the 64-element matrix is the corresponding subsystem output voltage E_0. Each column vector corresponds to an image sensed by the camera.

If each entry in the 8-by-8 matrix shown in Figure 16 is multiplied by a common factor of 1/5, the matrix thus obtained will be the measured correlation-coefficient matrix which is shown in Figure 17. Figure 18 shows the theoretically calculated correlation-coefficient matrix for the same set of images. The numerical entries on the main diagonals of the two matrices of Figures 17 and 18 are respectively the measured and desired self-correlation coefficients. It is seen that the measured

IMAGE INPUT FROM CAMERA

	4.7	2.6	0.1	-2.6	-4.6	-2.1	0.4	2.9
	2.5	4.9	2.9	0.3	-2.3	-4.6	-2.5	0.1
	0.2	2.9	4.9	2.6	0.2	-2.5	-4.5	-2.3
	-2.4	0.3	3.0	4.9	2.9	0.2	-2.4	-4.5
	-4.6	-2.4	0.3	3.0	5.0	2.8	0.4	-2.2
	-1.8	-4.5	-2.7	-0.1	2.6	5.0	3.3	0.9
	0.4	-2.7	-4.6	-2.2	0.4	3.3	4.9	3.0
	2.8	-0.2	-2.7	-4.5	-2.2	0.7	3.1	4.8

IMAGE RETRIEVED FROM RECORDER

Figure 16. Correlator Output Voltages for Various
Camera and Recorder Images

IMAGE INPUT FROM CAMERA

IMAGE RETRIEVED FROM RECORDER

.94	.52	.02	-.52	-.92	-.42	.08	.58
.50	.98	.58	.06	-.46	-.92	-.50	.02
.04	.58	.98	.52	-.04	-.50	-.90	-.46
-.48	.06	.60	.98	.58	.04	-.60	-.90
-.92	-.48	.06	.60	1	.56	.08	-.44
-.36	-.90	-.54	-.02	.52	1	.66	.18
.08	-.54	-.92	-.44	.08	.66	.98	.60
.56	-.04	-.54	-.90	-.44	.14	.62	.96
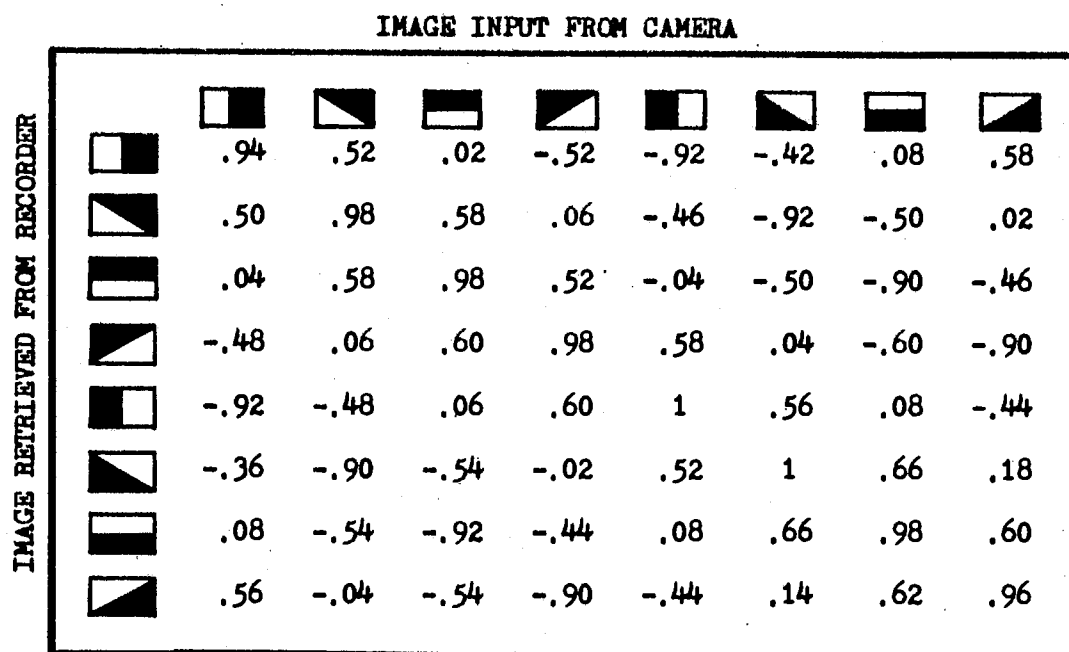
Figure 17. Measured Correlation-Coefficient Matrix

IMAGE INPUT FROM CAMERA

IMAGE RETRIEVED FROM RECORDER

1	.5	0	-.5	-1	-.5	0	.5
.5	1	.5	0	-.5	-1	-.5	0
0	.5	1	.5	0	-.5	-1	-.5
-.5	0	.5	1	.5	0	-.5	-1
-1	-.5	0	.5	1	.5	0	-.5
-.5	-1	-.5	0	.5	1	.5	0
0	-.5	-1	-.5	0	.5	1	.5
.5	0	-.5	-1	-.5	0	.5	1
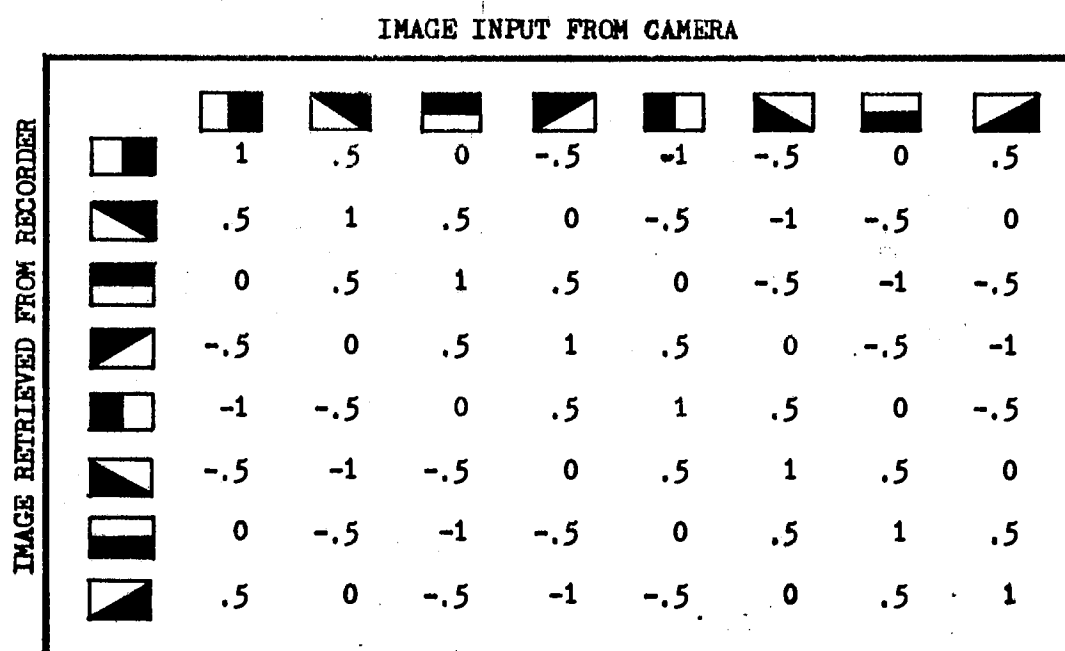
Figure 18. Desired Correlation-Coefficient Matrix

self-correlation coefficients are equal or nearly equal to the desired value (+1). The comparison between the measured and desired correlation-coefficient matrices also shows that the values of the off-diagonal elements of the measured matrix do not differ appreciably from those of the ideal symmetric matrix of Figure 18. In the measured matrix, the greatest error is found in the non-zero entry of 0.18 in row 6 and column 8. In general, the accuracies of the measured cross-correlation coefficients seem to be compatible with the normal tolerances of the electronic components used.

In the experiments performed with the video cross-correlating subsystem, the pre-recording of a series of images was manually controlled by the operator. Each image was recorded on the tape for a period of approximately ten seconds. However, the subsystem output voltage E_0 was ready to be read within three seconds after a retrieved image was stabilized on the monitor screen, regardless of whether the images were composed of simple or of complex patterns.

Thus, it is evident that the video cross-correlating subsystem is capable of generating a series of self-normalized coefficients of cross-correlation between an image sensed by the television camera and a series of images retrieved from the tape recorder. Each optical image is thereby converted into a corresponding feature vector which may be used in the pattern-recognition operations.

CHAPTER IV

THE VIDEO WAVEFORM-AGREEMENT-DETECTION SUBSYSTEM

It has been mentioned that an ideal preprocessor may be used to
minimize the dimensionality of the video patterns without sacrificing
their information content. The first approach to the ideal preproces-
sor, which was realized by use of the video cross-correlator, was des-
cribed in Chapter III. In this chapter, a second approach to the real-
ization of the ideal preprocessor is made by use of a video waveform-
agreement-detector concept developed by V. W. Bolie (21). With this
approach, a video waveform-agreement-detection subsystem of which the
waveform agreement detector is a part can generate, by a high-speed
analog technique, a series of voltages which indicate the degrees of
agreement between an image sensed by the television camera and a series
of reference images retrieved from the television tape recorder.

This chapter is devoted to the construction and testing of the
video waveform-agreement-detection subsystem. The circuit design and
performance sensitivity of the waveform-agreement-detector are des-
cribed. The performance and characteristics of the waveform-agreement-
detection video subsystem are also discussed.

The Overall Subsystem

The video waveform-agreement-detection subsystem is illustrated
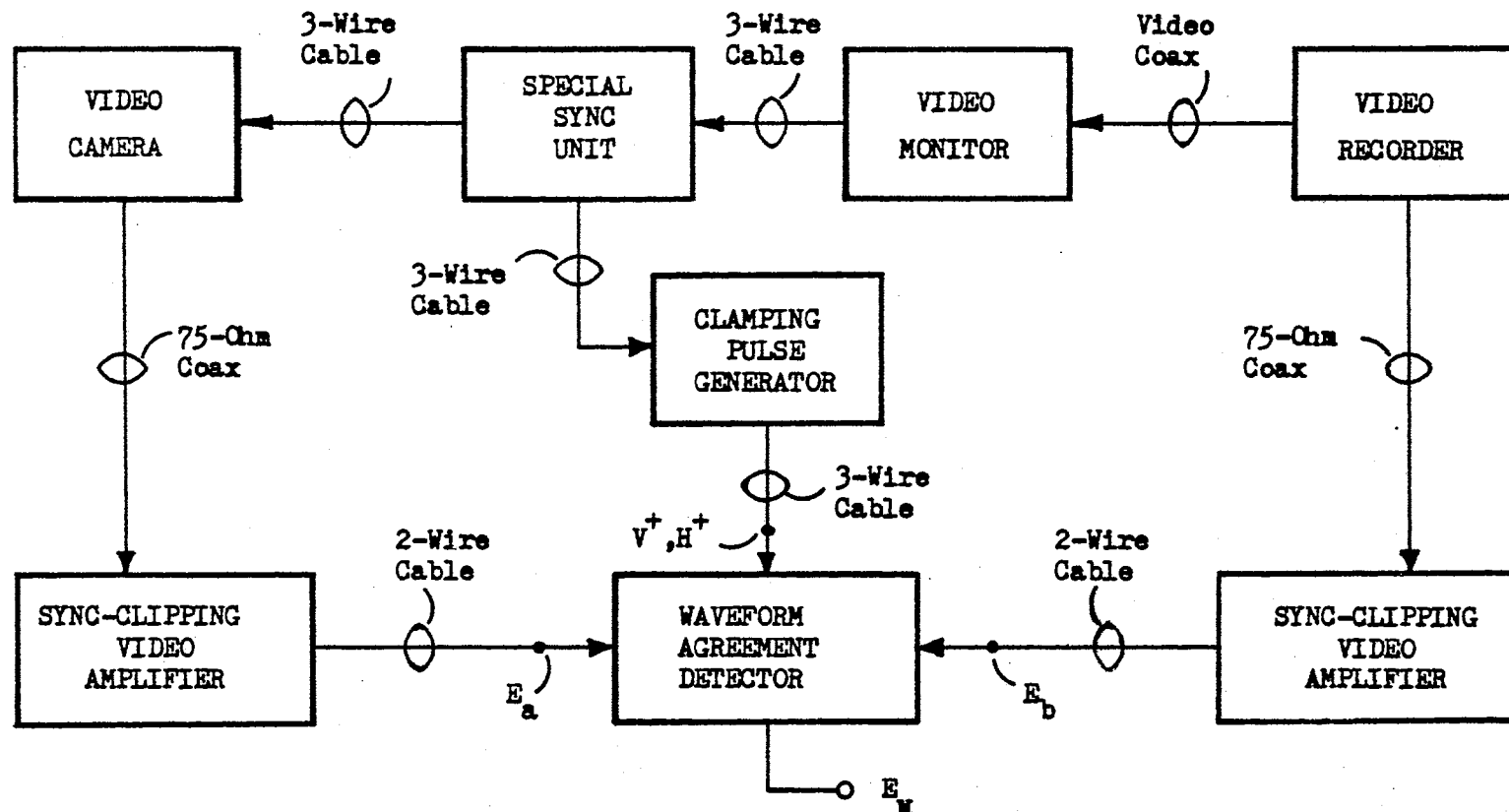schematically in Figure 19. The overall structure of the video

Figure 19. Video Waveform-Agreement-Detection Subsystem

waveform-agreement-detection subsystem is similar to that of the video cross-correlating subsystem which was described in the previous chapter. The video waveform-agreement-detection subsystem, however, contains a waveform agreement detector instead of the circuit ensemble of the cross-correlator and the pair of correlator pre-conditioners used in the video cross-correlating subsystem.

The video camera and the video tape recorder each furnish through a 75-ohm coaxial cable a ± 0.5 volt standard video signal. The camera-recorder synchronism is established by making use of the video monitor and the special sync unit in exactly the same manner as that employed in the cross-correlating video subsystem. Similarly, the sync-clipping video amplifiers, upon receiving the ± 0.5 volt video signal, remove the sync pulses and amplify the video to a scale of 15 volts for "black" and -5 volts for "white".

The sync-clipping video amplifier output voltages E_a and E_b, which vary within the ranges of $-5 \leq E_a \leq +15$ volts and $-5 \leq E_b \leq +15$ volts, are the two signal inputs to the waveform-agreement-detector. The clamping pulses V^+ and H^+ are fed from the clamping-pulse generator to the waveform-agreement-detector for the purpose of clamping-to-zero any "residual signal" existing in either a horizontal or vertical blanking-pulse interval. The waveform-agreement-detector produces the output voltage E_w which is actually a function of the voltage difference $(E_a - E_b)$.

The television apparatus used in the video waveform-agreement-detection subsystem was described in Chapter II. The circuit description of the special sync unit, the sync-clipping video amplifiers, and the clamping-pulse generator is given in Appendix A. The following

discussion will be confined to the waveform-agreement-detector.

Waveform-Agreement-Detector Circuit

For a clear observance of the circuit performance, the waveform-agreement-detector (WAD) circuit is divided into two sections, namely, the first section and the second section of the WAD circuit. The structure of the WAD circuit is shown schematically in Figure 20. The amplified sync-free signals E_a and E_b are fed into the first WAD section which produces the internal voltage E_f. The voltage E_f is the video input to the second WAD section which is also supplied with the clamping control signals V^+ and H^+. The output of the second WAD section is the voltage E_w.

Figures 21 and 22 show the details for the first and second sections of the WAD circuit, respectively. The first WAD section mainly consists of an absolute-difference amplifier and a signal inverter. The transistors Q1, Q2, Q3, and Q4, and their associated resistors are arranged so that when connected for operation the voltage drop across the R4 resistance is approximately equal to one-half of the absolute value of the voltage difference $(E_a - E_b)$. Therefore, when the input voltages E_a and E_b fall within the ranges of $-5 \leq E_a \leq +15$ volts and $-5 \leq E_b \leq +15$ volts, the voltage $e_1 = 0.5|E_a - E_b| - E_{cc}$ at the common node of the R4 resistor and the collectors of the Q3 and Q4 transistors varies within the range of $-15 \leq e_1 \leq -5$ volts. The circuit arrangement of the Q5 transistor and the R5 and R6 resistors is made so as to produce at the Q5 collector the inverted and amplified output voltage E_f. The diode D1 is used to prevent the Q5 transistor from entering the saturation region, and thus to restrict the voltage E_f varying
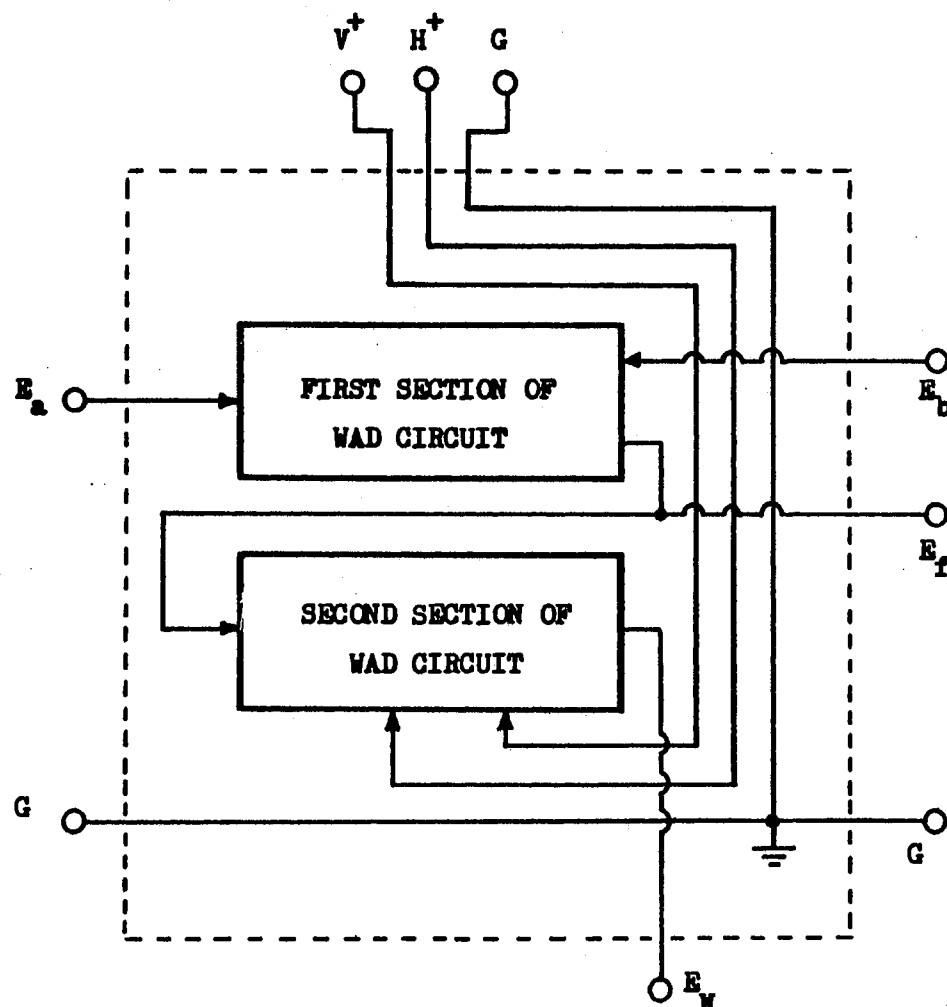
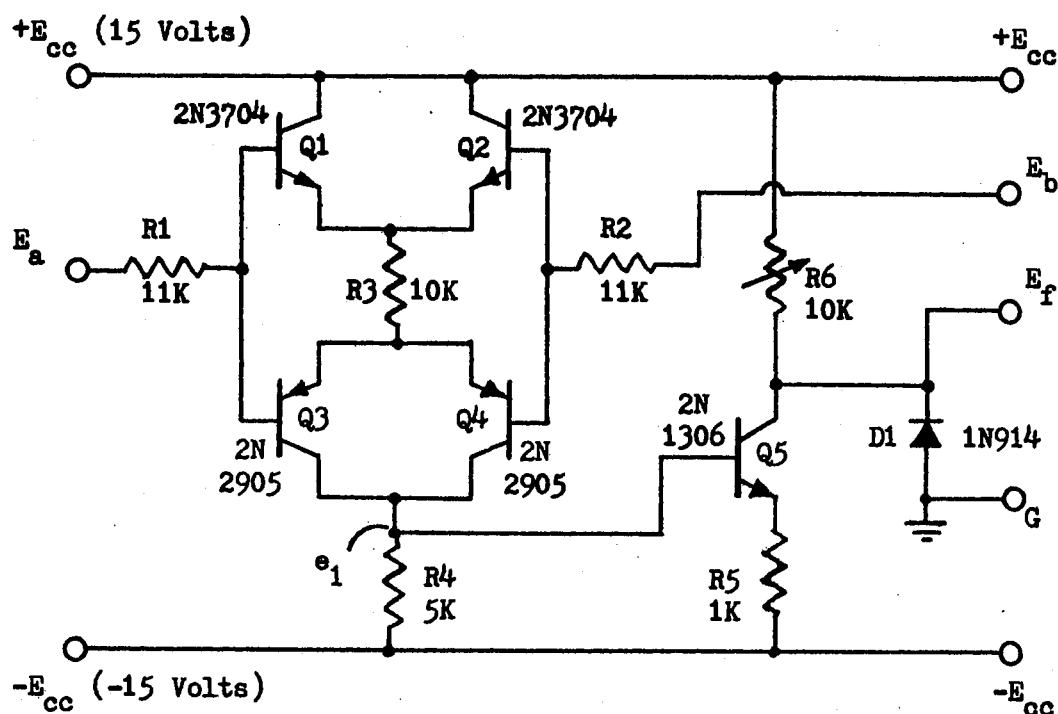Figure 20. Scheme of Waveform Agreement Detector

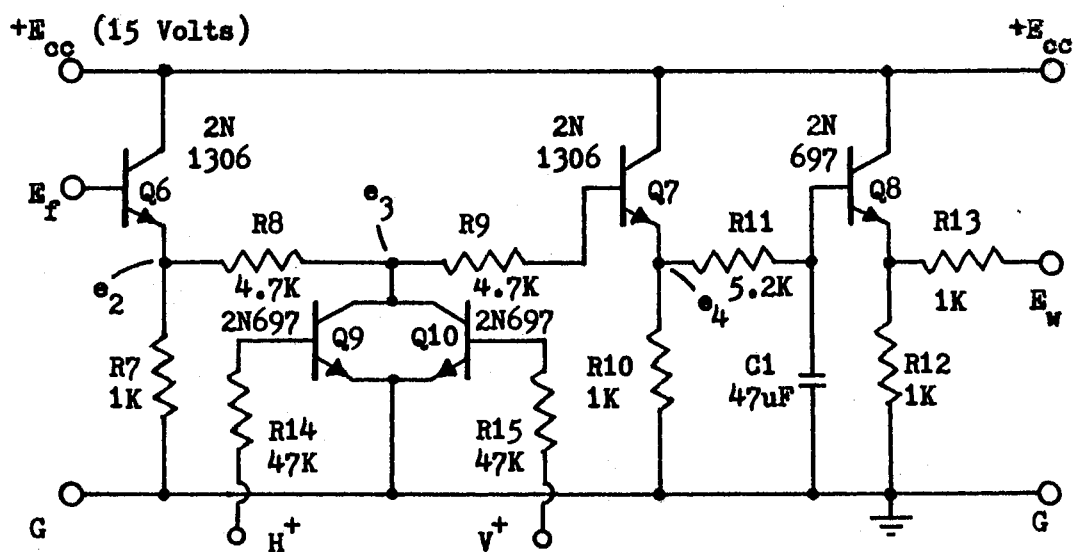Figure 21. First Section of Waveform Agreement Detector

Figure 22. Second Section of Waveform Agreement Detector

within the range of $0 \leq E_f \leq +15$ volts.

The performance sensitivity of the first WAD section can be improved by proper adjustment of the R6 resistance. The desired performance of the first WAD section is shown in Figure 23 in which the horizontal axis is the voltage difference $(E_a - E_b)$ and the vertical axis is the output voltage E_f. In Figure 23, the solid lines, the alternating long and short dashed lines, and the dotted lines refer to Cases I, II, and III, respectively. Case I corresponds to the case of R6 = 1.5 kilohms; Case II to that of R6 = 3 kilohms; Case III to that of R6 = 6 kilohms. According to the property of the absolute-difference amplifier, the characteristic lines are symmetric with respect to the vertical coordinate. It can be seen that the circuit operates with less sensitivity in Case I, moderate sensitivity in Case II, and higher sensitivity in Case III. The voltage E_f is zero within the subintervals of $|(E_a - E_b)| \geq 5$ volts for Case III and $|(E_a - E_b)| \geq 10$ volts for Case II. The remaining subintervals correspond to linear relationships between the input voltage difference $(E_a - E_b)$ and the output voltage E_f. Case I is of linear operation over the whole interval.

Figure 24 shows the measured performance of the first WAD section. For this measurement, two oppositely phased 20-volt (peak-to-peak) 1000 Hz sinusoidal signals, each superimposed on a 5-volt DC level, were used as the two inputs E_a and E_b to the first WAD section. The measured characteristics of the first WAD section is seen to be a reasonably good approximation of the desired one shown in Figure 23. The minor difference between the measured and desired characteristics is that the measured one has a 14-volt flat top existing within the subinterval of $|(E_a - E_b)| \leq 1.4$ volts. The flat top is caused by the two
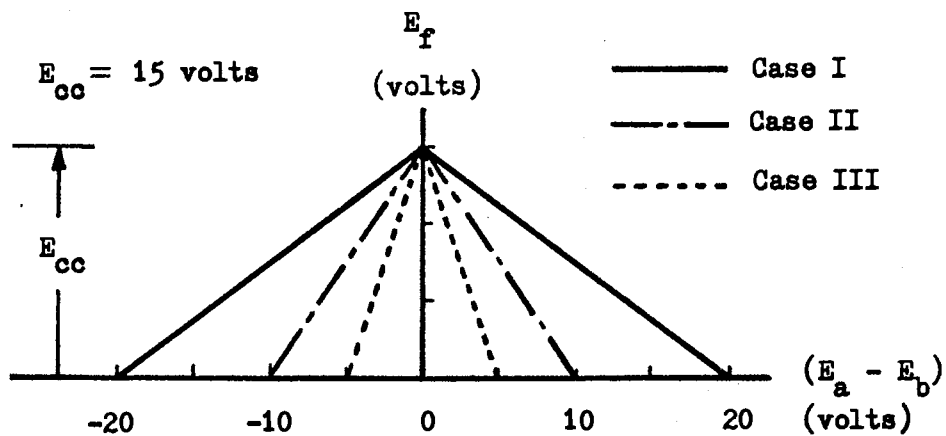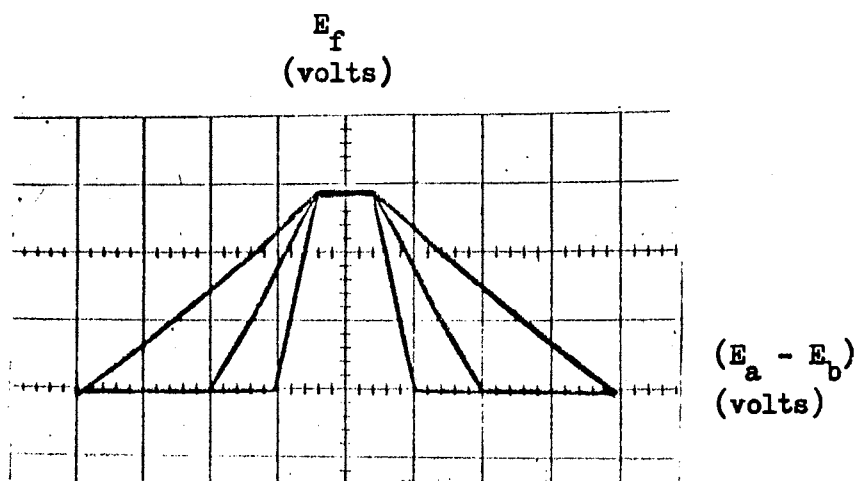
Figure 23. Desired Performance of First WAD Section

Horizontal scale: 5.0 volts/cm.

Vertical scale: 5.0 volts/cm.

Figure 24. Measured Performance of First WAD
Section

base-to-emitter voltage drops occurring in the active pair of the four transistors Q1, Q2, Q3, and Q4 in the absolute-difference amplifier. It is noted that the R6 resistance used for measuring the circuit performance is actually 2 kilohms for Case I, 4 kilohms for Case II, and 8.6 kilohms for Case III.

Figure 25 shows some waveforms measured at the nodes E_a, E_b, e_1, and E_f in the first WAD section. For these measurements, a 20-volt (peak-to-peak) 40 kHz sinusoidal signal was used as the input E_a and the other input was grounded. The voltage e_1 appears as a full-wave rectified voltage. The waveforms indicated by E_{f1}, E_{f2}, and E_{f3} are the voltages measured at the node E_f for Cases I, II, and III, respectively. The effects of adjusting the R6 resistance and zero-clamping by the D1 diode are observed in the E_{f1}, E_{f2}, and E_{f3} waveforms. It is seen that the circuit of the first WAD section can work effectively from DC to the frequency of 40 kHz, which is sufficient for the purposes of the studies reported here.

The second WAD section shown in Figure 22 has two functions of (1) clamping the voltage e_3 to the zero-volt potential of the common ground-bus G during the horizontal and vertical blanking-pulse intervals, and (2) developing by a single (5.2 K - 47 μF) low-pass filter a 244-msec running-time average of the residual-free signal voltage available at the Q7 emitter. The clamping is achieved by use of a simple analog-gating technique which utilizes the 12-volt horizontal blanking-pulse input H^+ and the 12-volt vertical blanking-pulse input V^+.

Since the residual signal at the Q6 emitter is actually a 14.3-volt, 12-microsecond signal during the horizontal blanking interval, and is a 14.3-volt, 1.0-msec signal during the vertical blanking
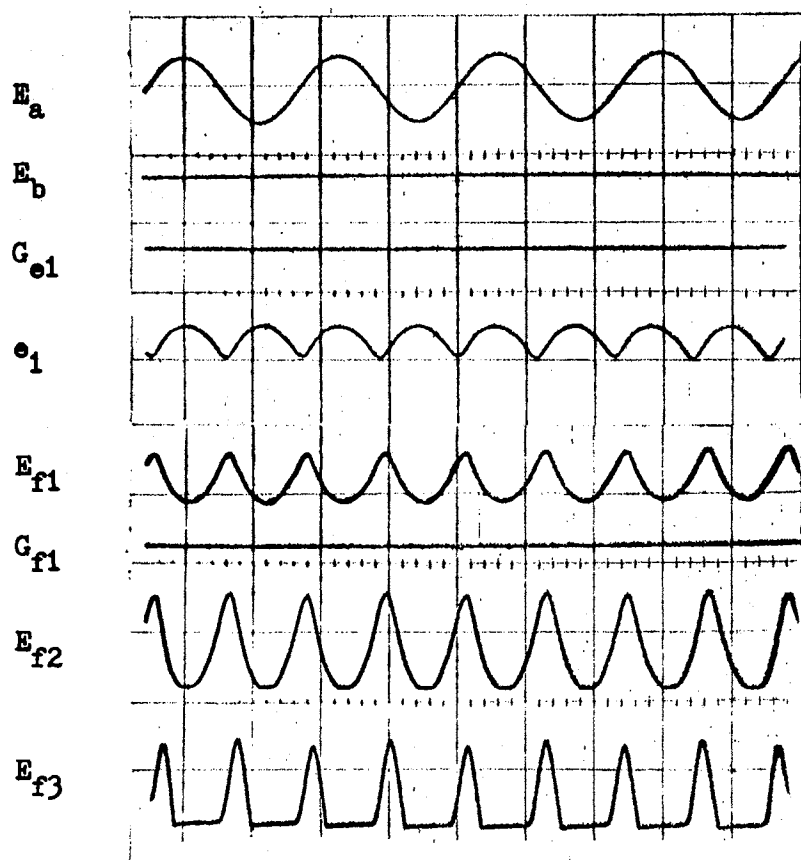
Vertical: E_a: Sine-Wave Input to WAD, 20 V/cm

E_b: Zero-Volt Input to WAD

G_{e1}: Zero-Volt Ground Reference for e_1

e_1: Absolute-Difference Voltage, 10 V/cm

E_{f1}: Output of First WAD Section (Case I), 10 V/cm

G_{f1}: Zero-Volt Ground Reference for E_{f1}

E_{f2}: Output of First WAD Section (Case II), 10 V/cm

E_{f3}: Output of First WAD Section (Case III), 10 V/cm

Horizontal: Time, 10 μsec/cm

Figure 25. Waveforms of First WAD Section

interval, the clamping function of the second WAD section is equivalent
to a linear subtraction of a constant voltage of 3.5 volts which is
determined by the numerical calculation of

$$(15.0 - 0.3) \left[1.0 - \left(\frac{63.5 - 12.0}{63.5} \right) \left(\frac{16.7 - 1.0}{16.7} \right) \right] \quad ,$$

where the number 15.0 corresponds to E_{cc}, 0.3 to the base-to-emitter
voltage drop of the Q6 resistor, 63.5 to the horizontal period, 12.0 to
the horizontal blanking-pulse interval, 16.7 to the vertical period,
and the subtrahend 1.0 to the vertical blanking-pulse interval. The
244-msec running-time average voltage available in the second WAD sec-
tion becomes the WAD output voltage E_w when measured at the emitter of
the Q8 transistor through the R13 resistance. It has been found that
if the (5.2 K - 47 µF) low-pass filter is directly connected to the R9
resistance without using the Q7 emitter follower, the corresponding
running-time average thus obtained is incorrect, due to the modulation
of the RC time constant by the blanking-pulses.

Subsystem Characteristics and Performance

Validity testing of the waveform-agreement-detection subsystem was
performed with the aid of two input horizontal-step images. The hori-
zontal-step image is an image of either "white" in the left half and
"black" in the right half, or "white" in the right half and "black" in
the left half. Some of the results obtained during the testing pro-
cedure are shown in Figures 26 and 27.

Figure 26 shows the development of the WAD output voltage E_w for a
pair of identical horizontal-step input images of "white" in the left
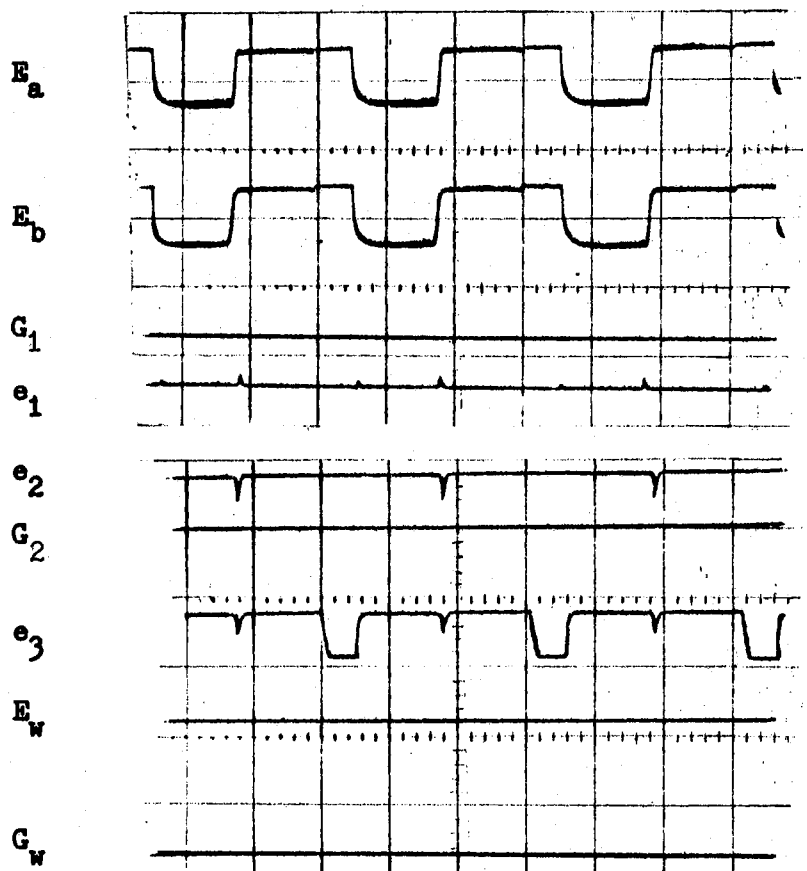
Vertical: E_a: Camera Input to WAD, 20 V/cm

E_b: Recorder Input to WAD, 20 V/cm

G_1: Zero-Volt Ground Reference for e_1

e_1: Absolute-Difference Voltage, 20 V/cm

e_2: Q6 Emitter Voltage, 20 V/cm

G_2: Zero-Volt Ground Reference for e_2

e_3: Clamped Voltage, 20 V/cm

E_w: WAD Output, ,5 V/cm

G_w: Zero-Volt Ground Reference for E_w

Horizontal: Time, 20 μsec/cm

Figure 26. Genesis of WAD Output Voltage, for
Identical Horizontal-Step Images
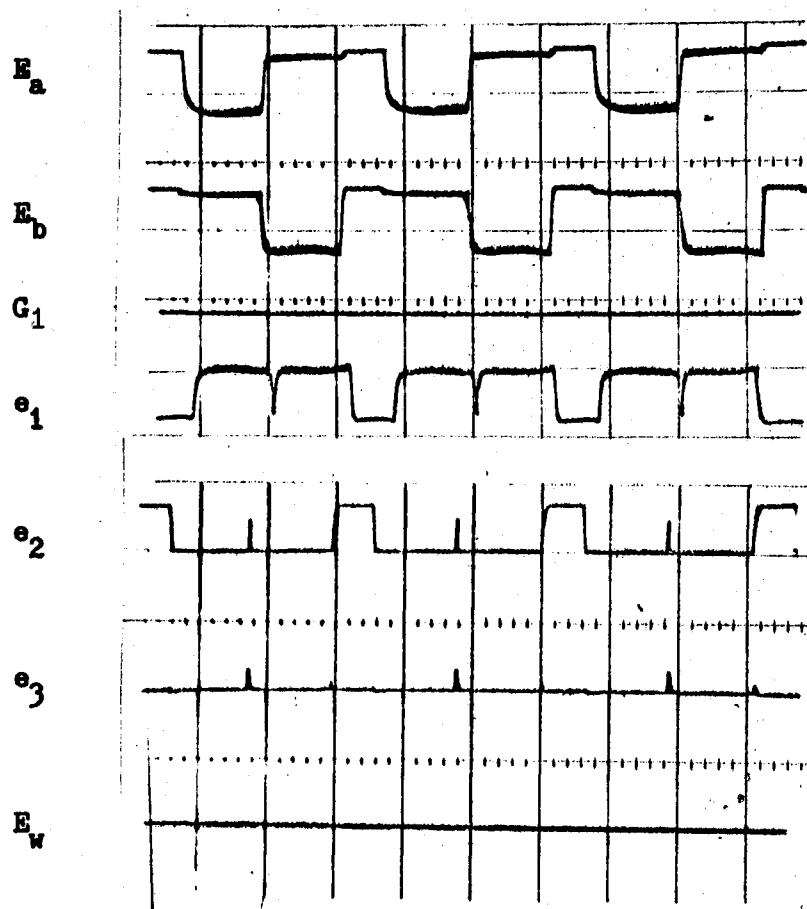
Vertical: E_a: Camera Input to WAD, 20 V/cm

E_b: Recorder Input to WAD, 20 V/cm

G_1 : Zero-Volt Ground Reference for e_1

e_1 : Absolute-Difference Voltage, 10 V/cm

e_2 : Q6 Emitter Voltage, 20 V/cm

e_3 : Clamped Voltage (Zero Volt)

E_w : WAD Output Voltage (Zero Volt)

Horizontal: Time, 20 μsec/cm

Figure 27. Genesis of WAD Output Voltage, for
Opposite Horizontal-Step Images

half and "black" in the right half. There are nine traces shown in Figure 26, each of which is identified by its corresponding alphabetical and numerical notation. The upper two traces show the sync-free amplified signal inputs E_a and E_b to the WAD circuit. The traces indicated by G_1, G_2, and G_w are, respectively, the zero-volt common ground references for the traces e_1, e_2, and E_w which are measured correspondingly at the nodes e_1, e_2, and E_w as shown in Figures 21 and 22. As the input images are in full agreement with each other, no current will flow through the resistance R4. Therefore, the voltage e_1 is equal to $-E_{cc}$ (= -15 volts), and voltage e_2 to E_{cc} - 0.3, or 14.7 volts, as shown by the e_1 and e_2 traces in Figure 26. The e_3 trace shows the effectiveness of the zero-clamping circuit of the second WAD section. For this pair of identical images, the output voltage E_w is found to be 10 volts when the R6 resistance is 4 kilohms as used in Case II of Figure 24.

Figure 27 similarly shows the development of the WAD output voltage E_w for the similar horizontal-step images as used in the case shown in Figure 26, except that the input image retrieved by the recorder is rotated by 180 degrees to put the black-half of the image on the left. The traces shown in Figure 27 are indicated in the same manner as those shown in Figure 26. Since the input images are in complete disagreement, there is a maximum current flowing through the R4 resistance. Thus, as shown by the e_1 and e_2 traces of Figure 27, there exist the maximum voltage drops across the R4 and R6 resistances all of the time except during the blanking-pulse time intervals, when the voltage e_1 is -15 volts and the voltage e_2 is 14.7 volts. The spikes appearing in the e_2 trace are believed due to the imperfection of images used and to the difficulty of image centering. The e_3 trace similarly shows the

effectiveness of the zero-clamping circuit of the second WAD section. For this pair of opposite horizontal-step images, the output voltage E_w is found to be zero when the R6 resistance used is 4 kilohms (Case II).

The overall performance of the waveform-agreement-detection video subsystem is illustrated in the 8-by-8 matrix arrangement of Figures 28, 29, and 30. In each diagram, the top row shows the various images retrieved by the recorder, and the left column shows the various images sensed by the camera. Each numerical entry in the 64-element matrix is the corresponding output voltage E_w of the waveform-agreement-detector circuit.

The matrices shown in Figures 28, 29, and 30 correspond, respectively, to Cases I, II, and III which have the performance curves shown in Figure 24. For the optical images illustrated in the three figures, the degree of waveform agreement between any pair of images is determined by the corresponding numerical entry, i.e., for a pair of image the larger the value of a numerical entry the higher the degree of waveform agreement. For all three matrices, the maximum value is found to be 10.2 (for a pair of identical images), and the minimum value is zero (for a pair of opposite images). Each numerical entry on the main diagonal of the three matrices is the voltage obtained for each pair of identical images. As the same group of eight patterns is used for both the images retrieved from the recorder and the images sensed by the camera, the corresponding waveform-agreement-detection matrices should be theoretically a symmetric matrix. The experimental results illustrated in Figures 28, 29, and 30 show that the values of the off-diagonal elements of the three matrices do not differ appreciably from those of the ideal symmetric matrix. The greatest error is found in

IMAGE RETRIEVED FROM RECORDER

INPUT IMAGE FROM CAMERA

10.2	8.2	6.2	4.1	2.5	4.2	6.1	8.2
8.2	10.2	8.2	6.2	4.3	2.7	4.3	6.2
6.2	8.2	10.0	8.1	6.2	4.3	2.4	3.8
4.0	6.1	8.0	9.9	8.3	6.4	4.5	2.7
2.6	4.5	6.4	8.4	10.0	8.2	6.5	4.2
4.9	2.5	4.4	6.6	8.2	10.2	8.4	6.3
6.5	4.6	2.4	4.6	6.6	8.4	10.2	8.2
8.5	6.5	4.5	2.7	4.3	6.3	8.4	10.2

Figure 28. Waveform-Agreement-Detector Output Voltages for Various Camera and Recorder Images (Case I)

IMAGE RETRIEVED FROM RECORDER

INPUT IMAGE FROM CAMERA

10.0	7.5	5.0	2.2	0	2.2	4.8	7.6
7.5	9.8	7.5	4.8	2.2	0	2.2	5.1
4.9	7.1	9.5	7.4	4.9	2.1	0	2.4
2.3	4.9	7.3	9.8	7.5	4.9	2.4	0
0	2.4	4.9	7.5	9.7	7.1	4.8	2.3
2.2	0	2.2	4.9	7.5	9.8	7.3	4.4
5.1	2.4	0	2.5	5.0	7.4	9.7	7.2
7.6	5.1	2.5	0	2.2	4.7	7.2	9.7

Figure 29. Waveform-Agreement-Detector Output Voltages for Various Camera and Recorder Images (Case II)

IMAGE RETRIEVED FROM RECORDER

	9.6	7.0	4.5	1.6	0	1.7	3.9	7.1
	7.1	9.3	6.9	4.0	1.6	0	1.6	4.6
	4.6	7.0	9.0	6.6	3.9	1.4	0	1.9
	1.9	4.4	6.6	9.3	6.9	4.2	2.0	0
	0	2.0	4.5	7.0	9.2	6.8	4.3	1.9
	1.6	0	1.5	4.2	6.8	9.0	6.6	3.1
	4.6	2.1	0	2.1	4.5	6.8	9.1	6.8
	7.2	4.7	2.1	0	1.7	4.2	6.5	9.1

INPUT IMAGE FROM CAMERA

Figure 30. Waveform-Agreement-Detector Output Voltages for
Various Camera and Recorder Images (Case III)

the difference of 1.1 between the entry of 3.1 in row 6 and column 8 of the matrix for Case III and the entry of 4.2 in row 8 and column 6 of the same matrix. It is seen that the accuracies of the various waveform-agreement-detection voltages are compatible with the normal tolerances of the electronic components used.

The effect of the sensitivity (R6) adjustment of the waveform-agreement-detection video subsystem can be observed by comparing the three matrices shown in Figures 28, 29, and 30. The matrix of Figure 28 corresponds to Case I, i.e., the case of operating with less sensitivity. In this matrix, the row-column elements (5,1), (6,2), (7,3), (8,4), (1,5), (2,6), (3,7), and (4,8), corresponding to the pairs of opposite images, have the values of 2.6, 2.5, 2.4, 2.7, 2.5, 2.7, 2.4, and 2.7, respectively. According to the measured subsystem performance (Figure 24), these elements should be zero, since each corresponding voltage difference $(E_a - E_b)$ is either 20 volts or -20 volts. The non-zero results imply that a part of the voltage difference $(E_a - E_b)$ is neither 20 volts nor -20 volts. The imperfection of the voltage difference is possibly due to the non-uniform light intensity and the incorrect image alignment. This discrepancy may be compensated by adjusting the performance sensitivity, or in other words, by properly adjusting the R6 resistance in the WAD circuit shown in Figure 22.

The performance improvement by the sensitivity adjustment is verified by the zero-elements of the matrices in Figures 29 and 30. Further inspection of the matrices for Cases II and III shows that the values of the diagonal elements of the two matrices are somewhat smaller than those of the corresponding elements of the matrix for Case I. Theoretically, the diagonal elements of the three matrices should have the

same value. The voltage drop in the diagonal elements of the matrices for Cases II and III is possibly due to the finite base current required by the Q6 emitter follower shown in Figure 22. The drop is 0.5 volt for Case II and 1.2 volts for Case III. However, as far as the performance indicated in the matrix for Case II is concerned, an increase of the sensitivity significantly improves the performance of the video waveform-agreement-detection subsystem.

Thus, the characteristics of the video waveform-agreement-detection subsystem are seen to be compatible with the requirement that it be capable of generating a series of voltages which indicate the degrees of waveform agreement between an image sensed by the television camera and a series of reference images retrieved by the television tape recorder. Consequently, the video waveform-agreement-detection subsystem can be used to convert each optical image into an equivalent feature vector.

CHAPTER V

AN IMAGE RECONSTRUCTION ALGORITHM

It is noted that both the video cross-correlating subsystem and
the video waveform-agreement-detection subsystem, described in Chapters
III and IV, respectively, have the property of converting a single input
(test) image into a corresponding feature vector. The available feature
vector, however, will be a valid representation of an image only if it
consists of the significant information of the image. One method of
testing the validity of the feature-vector representation of an image
is to reconstruct the image directly from its corresponding feature
vector and then to compare the reconstructed image with the original
image. The purpose of this chapter is, therefore, to develop a mathe-
matical algorithm for recovering a test image from its equivalent fea-
ture vector.

Assumptions and Definitions

A "test image" or "input image", arbitrarily denoted by "B", re-
fers to the image which is sensed by the television camera when the
whole television apparatus operates in the so-called special playback
mode. Similarly, a "reference image" refers to the image pre-recorded
in the tape and retrieved from the recorder. A reference image will be
denoted by "R", and a series of k reference images are distinquished
from each other by numerical subscripts, i.e., $R_1, R_2, \ldots, R_k$.

The test and reference images used in the research reported here contain only the black and white information. It is assumed that the images are exposed in an environment of uniform light intensity. This assumption implies that the "black" means the "uniformly and purely black", and the "white" means the "uniformly and purely white". Therefore, the information content of an image can be represented in a binary-valued form. For example, the value +1 will be assigned to the white and -1 to the black.

Since the aspect ratio (the quotient of the width to the height of a rectangular display) of a standard television frame is 4/3, each image used in the experiments is confined within a fixed rectangular plane having the same aspect ratio of 4/3. In an image of m-by-n resolution, each of the many (m times n) rectangular or square planes is called a "fundamental plane". Figure 31 shows an original image of 4-by-3 resolution and the corresponding 12 fundamental planes indicated by (1,1), (1,2),(1,3),(2,1),...,(3,4). Any two fundamental planes are called "connected" if they share a common sideline, and are not connected if they share one or no common point. For example, as shown in Figure 31, the fundamental planes (1,1) and (1,2) are connected, but the fundamental planes (1,1), (2,2), and (3,4) are not connected.

The "reverse" of an original image is an image in which each fundamental plane is made by applying a complete set of black and white reversing operations to the original image. Figure 31 also shows the reversed image of an original image.

A "recovered image" is an image which is established by a certain reconstruction process from its corresponding feature vector. The recovered image of an image B is denoted by B^*.
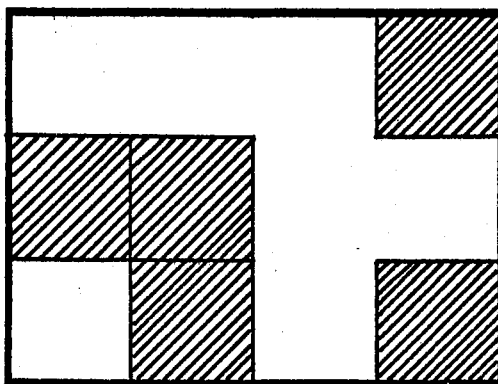
(1,1)	(1,2)	(1,3)	(1,4)
(2,1)	(2,2)	(2,3)	(2,4)
(3,1)	(3,2)	(3,3)	(3,4)
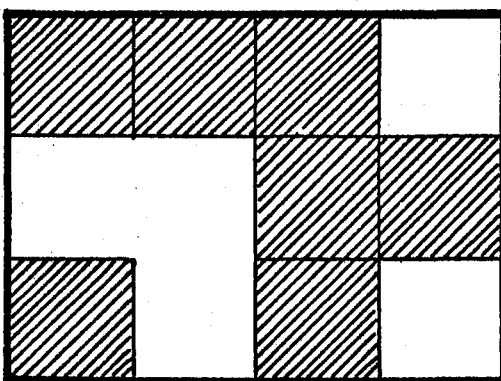
Figure 31. Original Image (Top), Fundamental Planes (Middle), and Reversed Image (Bottom)

The "information matrix" of an image is a matrix in which each entry represents the information content of an assigned and distinct fundamental plane of that image. Each entry of an information matrix of a reference or test image is assigned the value +1 when the corresponding fundamental plane is white and the value -1 when the corresponding fundamental plane is black. One exception is the case of the primary information matrix which will be described in the next section (each entry of a primary information matrix can be any real number). Accordingly, the information matrix of the original image shown in Figure 31 is

$$
\begin{bmatrix}
1 & 1 & 1 & -1 \\
-1 & -1 & 1 & 1 \\
1 & -1 & 1 & -1
\end{bmatrix}.
$$

The "mean" of an image is defined to be the quotient of the differential area (total white area subtracted by total black area) to the total image area. According to this definition, the image mean can be obtained by dividing the algebraic sum of the elements (+1 or -1) of the corresponding m-by-n information matrix by the product of m times n. For example, if the information matrix is

$$
\begin{bmatrix}
d_{11} & d_{12} & d_{13} \\
d_{21} & d_{22} & d_{23}
\end{bmatrix},
$$

the mean of the corresponding image will be $(d_{11}+d_{12}+d_{13}+d_{21}+d_{22}+d_{23})/6$.
When the total white and the total black share equally one-half of the
image area, the image is called a "zero-mean" image.

The "image product" of two images, G and H, of m-by-n resolution,
is denoted by G*H, and is defined to be an image of the same resolution
whose information matrix is established in such a way that if the in-
formation matrices of the two original images are

$$
[G] = \begin{bmatrix}
g_{11} & g_{12} & \cdots & g_{1n} \\
g_{21} & g_{22} & \cdots & g_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
g_{m1} & g_{m2} & \cdots & g_{mn}
\end{bmatrix}
\tag{5.1}
$$

and

$$
[H] = \begin{bmatrix}
h_{11} & h_{12} & \cdots & h_{1n} \\
h_{21} & h_{22} & \cdots & h_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
h_{m1} & h_{m2} & \cdots & h_{mn}
\end{bmatrix}
\tag{5.2}
$$

the information matrix of the image product G*H is

$$[G*H] = \begin{bmatrix} g_{11}h_{11} & g_{12}h_{12} & \cdots & g_{1n}h_{1n} \\ g_{21}h_{21} & g_{22}h_{22} & \cdots & g_{2n}h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1}h_{m1} & g_{m2}h_{m2} & \cdots & g_{mn}h_{mn} \end{bmatrix} \qquad (5.3)$$

Image Reconstruction Process

The image reconstruction process employed in the studies reported here is actually a computing process which can be used to reconstruct a test image directly from its equivalent feature vector. The image reconstruction process can be expressed by the equation

$$[B^*] = TF\{a_1 u_1 |u_1|^{(p_1-1)} [R_1] + a_2 u_2 |u_2|^{(p_2-1)} [R_2] + \dots$$

$$+ a_k u_k |u_k|^{(p_k-1)} [R_k]\}$$

$$= TF\{\sum_{i=1}^{i=k} a_i u_i |u_i|^{(p_i-1)} [R_i]\} \qquad (5.4)$$

where $[B^*]$ is the information matrix of the recovered image B^*; TF denotes a threshold operator; $a_1, a_2, \dots, a_k$ are a set of pre-assigned scaling factors; $u_1, u_2, \dots, u_k$ are the elements of the corresponding feature vector; $p_1, p_2, \dots, p_k$ are a set of pre-assigned non-zero positive numbers; and $[R_1], [R_2], \dots, [R_k]$ are the information matrices of the reference images. The validity of this reconstruction algorithm is demonstrated theoretically for a special case in Appendix B.

Equation (5.4) indicates that the first task in the reconstruction

process is to weight each reference matrix by the corresponding element of the feature vector, with pre-assigned modification, and by the pre-assigned scaling factor. Then, the weighted matrices are added together to form a matrix sum called here the "primary information matrix", i.e.,

$$\sum_{i=1}^{i=k} a_i u_i |u_i|^{(p_i-1)} [R_i] \quad .$$

In Equation (5.4), the product $u_i |u_i|^{(p_i-1)}$ is arranged so that the positive or negative sign of u_i is reserved regardless of whether the exponential power p_i is an odd or even number. The exponential powers $p_1, p_2, \ldots, p_k$ are used to provide a flexible modification on the feature vectors. For instance, the result may (or may not) be improved if some, or all, of the exponential powers are assigned different numbers. Also, the scaling factors $a_1, a_2, \ldots, a_k$ are used to give the flexibility of weighting the reference matrices. For example, when a certain reference image is considered much more significant than the rest of reference images, the corresponding scaling factor is then assigned a value greater than the others.

The threshold operator TF of Equation (5.4) is designed to perform the functions of (a) inspecting the primary information matrix, i.e., to find the largest and smallest entries in the matrix; (b) setting a proper threshold level for information decision; and (c) establishing the information matrix of the recovered image. It is seen that the threshold operation plays a significant role in the final decision of the information content of the recovered image. However, no restrictive rules are defined to the setting of the threshold level. The setting of a proper threshold level appears best done through the use of

cut-and-try techniques. A tentative threshold may be set at the mid-point between the largest and the smallest entries in the primary information matrix of the recovered image.

Tentative Reference Images

The tentative reference images employed in the investigation of image reconstructions can be categorized into three kinds of images: the window images; the belt images; and the orthogonal images.

A "window image" of m-by-n resolution is established by assigning to only one of the many (m times n) fundamental planes of an image the +1 information content of the white, and to all of the rest of the fundamental planes of the same image the -1 information content of the black. Figure 32 shows the complete set of window images of 4-by-4 resolution.

A "belt image" of m-by-n resolution is formed by assigning to only one of the m horizontal belts, or to only one of the n vertical belts, the +1 information content of the white, and to the rest of the fundamental planes of the same image the -1 information content of the black. It is noted that a horizontal belt is composed of n horizontally connected fundamental planes, and that a vertical belt is composed of m vertically connected fundamental planes. Figure 33 illustrates the complete set of belt images of 4-by-4 resolution.

Two images are called "orthogonal" if the corresponding cross-correlation coefficient is zero. Figure 34 shows a set of orthogonal images. The corresponding information matrices are listed in Table III. The images shown in Figure 34 have the specific features of (a) each image being of 4-by-4 resolution; (b) each image being a zero-mean

(32.01)   (32.02)   (32.03)   (32.04)

(32.05)   (32.06)   (32.07)   (32.08)

(32.09)   (32.10)   (32.11)   (32.12)

(32.13)   (32.14)   (32.15)   (32.16)
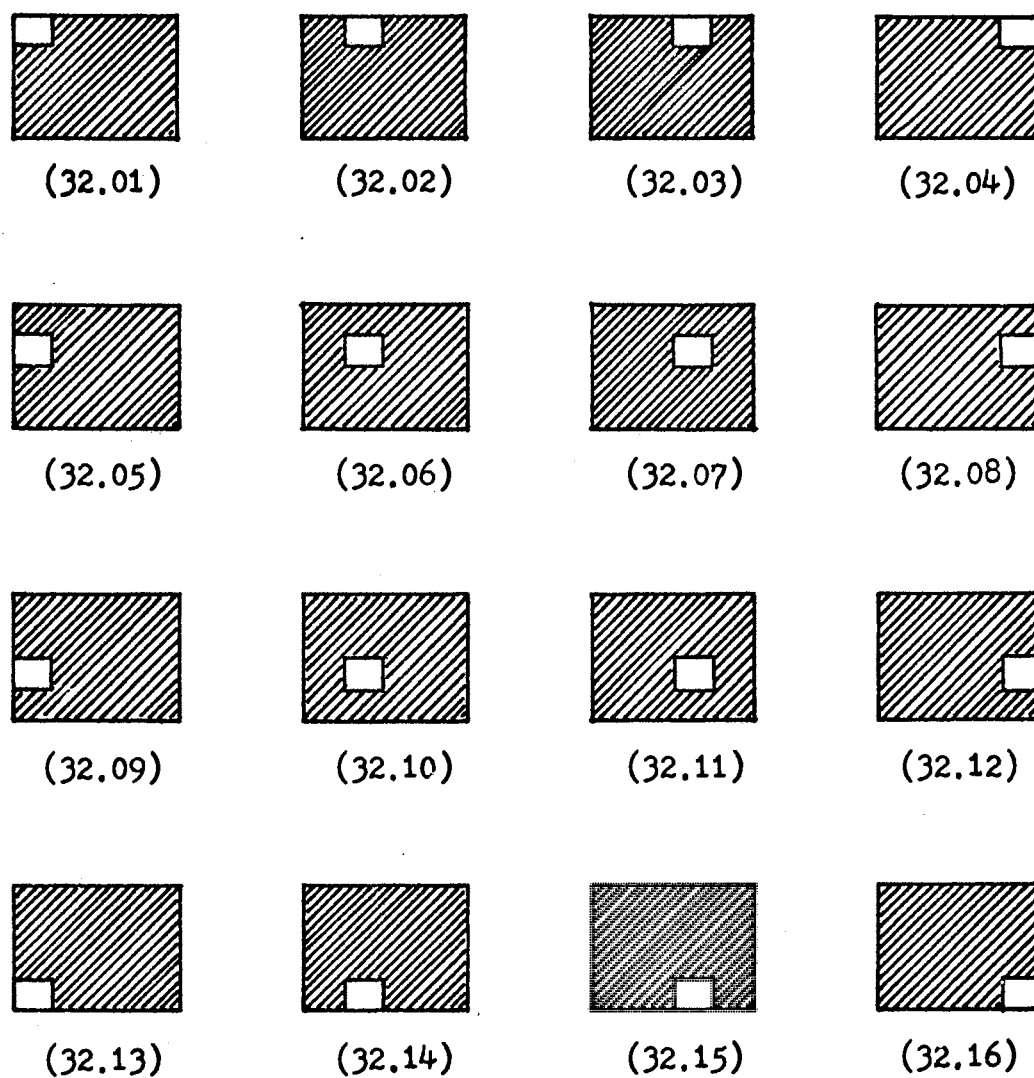
Figure 32.   A Set of Window Images of 4-by-4 Resolution
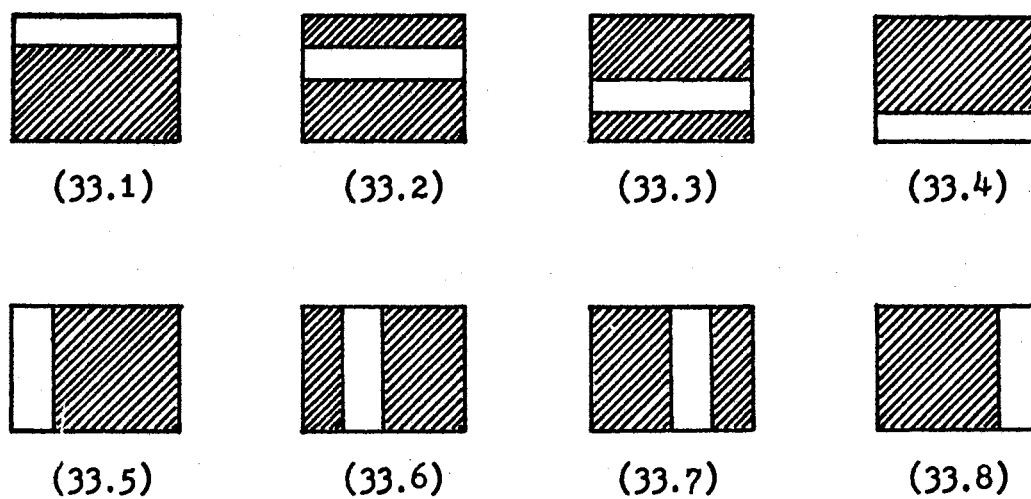
Figure 33. A Set of Belt Images of 4-by-4 Resolution

TABLE III

INFORMATION MATRICES OF ORTHOGONAL IMAGES

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(T3.01)

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$$

(T3.02)

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

(T3.03)

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

(T3.04)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{bmatrix}$$

(T3.05)

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$$

(T3.06)

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$$

(T3.07)

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

(T3.08)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(T3.09)

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$$

(T3.10)

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

(T3.11)

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

(T3.12)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix}$$

(T3.13)

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$$

(T3.14)

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$$

(T3.15)

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

(T3.16)

(34.01)   (34.02)   (34.03)   (34.04)

(34.05)   (34.06)   (34.07)   (34.08)

(34.09)   (34.10)   (34.11)   (34.12)

(34.13)   (34.14)   (34.15)   (34.16)
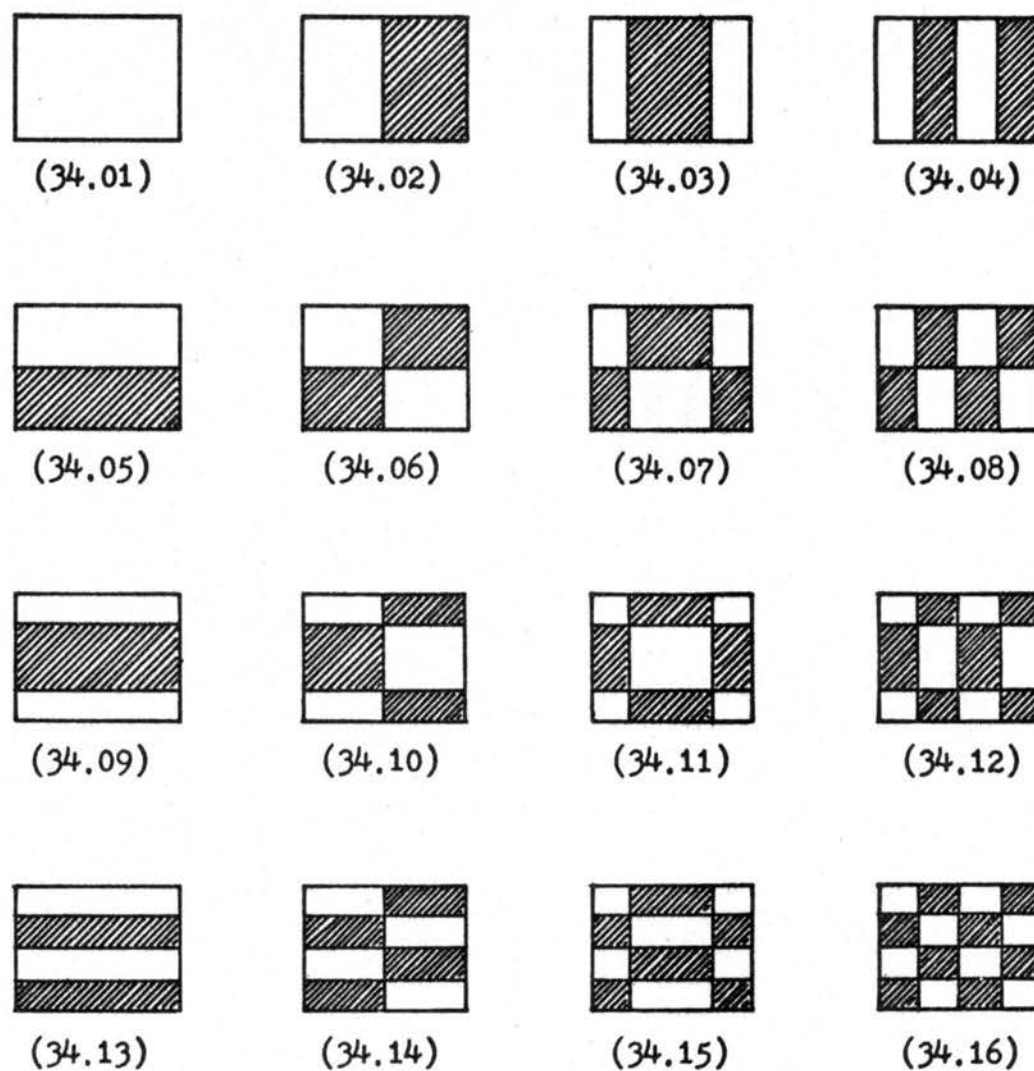
Figure 34. A Set of Orthogonal Images of 4-by-4 Resolution

image, except the image (34.01); and (c) the theoretical cross-correlation coefficient between any two images being zero. It is seen that some, or all, of the orthogonal images shown in Figure 34 can be replaced by their reversed images without changing the property of orthogonality.

For a set of zero-mean orthogonal images, the test of orthogonality between any two images can be made with aid of the corresponding information matrices. Since the image representation is +1 for the white and -1 for the black, the algebraic sum of the many (m times n) elements of the information matrix of the image product of the two original images should be zero if the two images are orthogonal. As an example, the information matrices of two zero-mean images of 2-by-4 resolution may be represented as

$$\begin{bmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \end{bmatrix}$$

and

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \end{bmatrix}.$$

If the sum of products,

$$\sum_{i=1}^{i=2} \sum_{j=1}^{j=4} g_{ij} h_{ij}$$

is zero, the two zero-mean images are orthogonal.

Formation of Orthogonal Images

In the process of forming a set of zero-mean orthogonal images, it is assumed that for the images of m-by-n resolution, the numbers m and n satisfy the equations

$$m = 2^x, \text{ where } x = 1,2,3,\ldots \qquad (5.5)$$

and

$$n = 2^y, \text{ where } y = 1,2,3,\ldots \qquad . \qquad (5.6)$$

This assumption does not imply that no zero-mean orthogonal images can be formed if the numbers m and n do not satisfy Equations (5.5) and (5.6), but it is used to simplify the description of developing an algorithm for the formation of the zero-mean orthogonal images.

The first step of forming a set of zero-mean orthogonal images is to construct the vertical "stripe" images in which the white and black vertical stripes of equal horizontal width are interlaced. Figure 35 shows the vertical stripe images for the cases of n = 2, 4, 8, and 16. The second step is to construct the subset of the "basic" images by use of the obtained vertical stripe images. Figure 36 shows the development of the basic images of 8-by-8 resolution. In the case of n = 2, there is only one basic image (36.1). If n = 4, there are three basic images (36.2), and if n = 8, there are seven basic images (36.3).

Further inspection of the images shown in Figure 36 indicates that
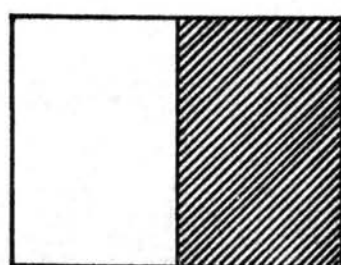
(35.1)
n = 2
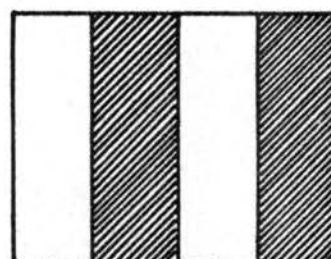
(35.2)
n = 4
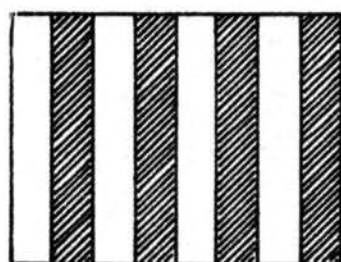
(35.3)
n = 8
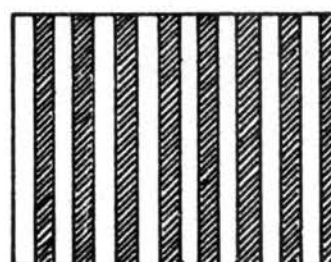
(35.4)
n = 16

Figure 35. Vertical Stripe Images
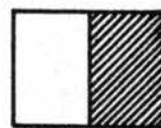
(36.1) , n = 2

(36.2) , n = 4

(36.3) , n = 8

Figure 36. Development of Basic Zero-Mean Orthogonal Images

the top-row image of (36.2) is identical to the basic image of (36.1), and that the top-row three images of (36.3) are identical to the three basic images of (36.2). It is also seen that in the images of (36.2), the bottom-right image is the image product of the top-row image and the bottom-left image, and that in the images of (36.3), the three images in the bottom-right are each the image products of the image in the bottom-left and the three images in the top row. Therefore, it is obvious that the image-product operation described may be used for the construction of more complex basic images.

In the last step of the orthogonal-image construction process, the subset of basic images is used to generate the whole set of zero-mean orthogonal images. The basic images are put in the top row, as illustrated by the images (34.02), (34.03), and (34.04) in Figure 34. The 90-degree rotated images of the basic images are put in the left-most column as shown by the images (34.05), (34.09), and (34.13) in Figure 34. The image-products of each image in the left-most column and the basic images in the top row become the corresponding row of newly-formed orthogonal images. The whole set of possible zero-mean orthogonal images comprises the subset of basic images, the subset of the 90-degree rotated images of the corresponding basic images, and the subset of newly-formed images.

The orthogonal images shown in Figure 34 correspond to the case of $m = 4$ and $n = 4$. However, the images in the top two rows correspond to the case of $m = 2$ and $n = 4$, and the images in the left two columns to the case of $m = 4$ and $n = 2$. The rule described here is also applicable to the formation of zero-mean orthogonal images of higher resolution.

Selection of Reference Images

In the process of image reconstruction, the particular selection of reference images may affect the validity of the results. A part of the significant information content of a recovered image may be inadvertently neglected or, conversely, unnecessarily emphasized if the reference images used in the reconstruction process are not properly selected.

The selection of reference images employed in the reconstruction process depends on the particular video subsystem used to generate the corresponding feature vector. If the feature vector is obtained by use of the video waveform-agreement-detection subsystem, the selection of reference images remains undefined. However, if the video cross-correlating subsystem is used to generate the feature vector, the use of the window and belt images is limited because the ratio of the white area to the black area in a window or belt image of high resolution is either too large or too small to be applicable in the operation of the video cross-correlating subsystem described in Chapter III. The inconvenience of excessively large or small white-to-black ratios is easily understood by recalling the formula

$$E_0 = 5 \; \frac{\overline{\alpha\beta}}{\sqrt{\overline{\alpha^2} \cdot \overline{\beta^2}}} \qquad\qquad (5.7)$$

which was used to describe the video cross-correlator output voltage E_0. When the white-to-black ratio of an image approaches a large number ($>>1$) or a small number ($<<1$), the numerator term $\overline{\alpha\beta}$, and one of the denominator terms $\overline{\alpha^2}$ or $\overline{\beta^2}$ approach zero with the result that the

formula for E_0 will contain a "zero/zero" ambiguity.

Thus, if the video cross-correlating subsystem is used to generate the feature vector, the window and belt images can be used only for the cases of low-resolution images. The superiority of the zero-mean orthogonal images over the window and belt images is evidently that the white-to-black ratio of any one of the zero-mean orthogonal images is always 1.0 regardless of the image resolution. The property of orthogonality in itself is also naturally desirable. Consequently, the set of zero-mean orthogonal images is considered here to be a proper set of reference images for the purpose of image reconstruction.

CHAPTER VI

EXPERIMENTS IN IMAGE RECONSTRUCTION

In order to verify the effectiveness of the image-reconstruction algorithm described in the previous chapter, further experiments were conducted with both the video cross-correlating subsystem and the video waveform-agreement-detection subsystem. This chapter is devoted to the detailed description of the illustrative experiments in image reconstruction. The pre-recorded reference images and the camera-sensed images employed in the experiments are of 4-by-4 resolution, and are of the same size of 3" high and 4" wide. The set of zero-mean orthogonal images which were shown in Figure 34 is used as the set of reference images. A set of ten arbitrarily made images, which are shown in Figure 37, and which are indicated as the images (37.01),(37.02),...,(37.10), is used as the set of test images. The image reconstruction process is simulated by a computer FORTRAN program.

Simulation of Reconstruction Process

The image reconstruction process previously indicated by Equation (5.4) can be simulated by a computer FORTRAN II program which is listed in Table IV. In general, the dimension statement in the program should be read as "DIMENSION IR(K,L), A(K), P(K), U(K), T(L), IS(L)" in which K is the total number of reference images used, and L is the total number of elements in a corresponding information matrix. The matrix

80
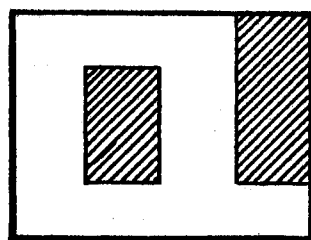
(37.01)
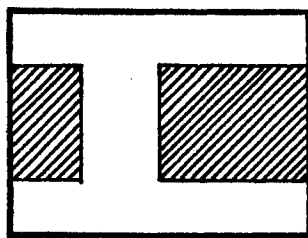
(37.02)
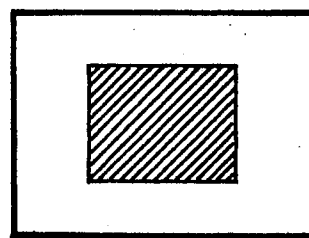
(37.03)
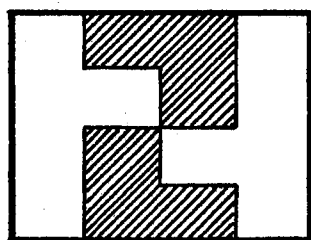
(37.04)
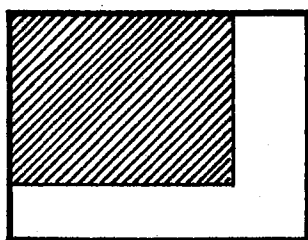
(37.05)
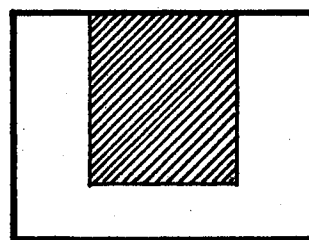
(37.06)
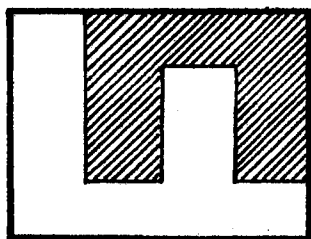
(37.07)
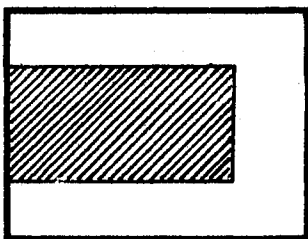
(37.08)
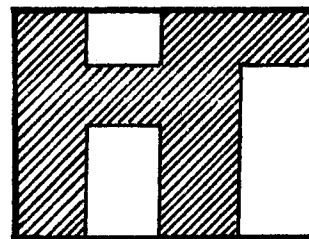
(37.09)
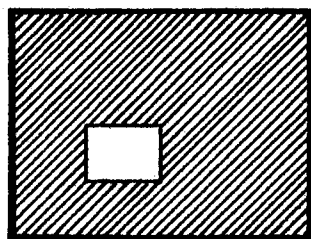
(37.10)

Figure 37. Test Images of 4-by-4 Resolution

TABLE IV

FORTRAN PROGRAM FOR IMAGE RECONSTRUCTION

```
ZZJOB
ZZFORX
C       FORTRAN PROGRAM FOR IMAGE RECONSTRUCTION
        DIMENSION IR(15,16), A(15), P(15), U(15), T(16), IS(16)
        K = 15
        L = 16
  10    FORMAT (5X, 16I4)
  11    FORMAT (5X, 8F7.2)
  12    FORMAT (23H THE RECOVERED IMAGE IS)
  13    FORMAT (F5.2)
  14    FORMAT (14HSENSE SWITCH 1)
  15    FORMAT (14HSENSE SWITCH 2)
  16    FORMAT (8HSCALE IS)
  17    FORMAT (34H THE PRIMARY INFORMATION MATRIX IS)
  18    FORMAT (3X, 4F6.1)
  19    FORMAT (3X, 4I4)
        DO 20 I=1,K
  20    READ 10, (IR(I,J), J=1,L)
        READ 11, (A(I), I=1,K)
        READ 11, (P(I), I=1,K)
  25    READ 11, (U(I), I=1,K)
        DO 28  I=1,K
        W=U(I)
        W1=ABSF(W)
        IF (W) 26,27,27
  26    Y=-1.0
        GO TO 28
  27    Y=1.0
  28    U(I)=Y*(W1**P(I))
        DO 40  J=1,L
        SUM=0.
        DO 30  I=1,K
        S=IR(I,J)
  30    SUM=SUM+S*A(I)*U(I)
  40    T(J)=SUM
        PUNCH 17
        PUNCH 18, (T(J), J=1,L)
  45    FMAX = T(1)
        DO 51  J=2,L
        IF (FMAX-T(J)) 50,51,51
```

TABLE IV (Continued)

```
50   FMAX = T(J)
51   CONTINUE
52   FMIN = T(1)
     DO 61 J=2,L
     IF (T(J)-FMIN) 60,61,61
60   FMIN = T(J)
61   CONTINUE
62   TYPE 16
     ACCEPT 13, SCALE
     FLAVL = FMIN + SCALE*(FMAX - FMIN)
     DO 72 J=1,L
     IF (T(J) - FLAVL)  70,71,71
70   IS(J) = -1
     GO TO 72
71   IS(J) = 1
72   CONTINUE
     TYPE 10, (IS(J), J=1,L)
     TYPE 14
     PAUSE
C    TURN SENSE SWITCH 1 ON FOR CHANGING 'SCALE'
     IF (SENSE SWITCH 1)  62,81
81   PUNCH 12
     PUNCH 19, (IS(J), J=1,L)
     TYPE 15
     PAUSE
     CALL RETYPE
C    TURN SENSE SWITCH 2 ON TO READ NEXT FEATURE VECTOR
     IF (SENSE SWITCH 2)  25,85
85   CALL EXIT
     END
```

IR(K,L) represents the complete set of reference images, i.e., each row corresponds to one of the reference images. The notation A(K) corresponds to the vector of preassigned scaling factors; P(K) to the vector of exponential powers; U(K) to the obtained feature vector; T(L) to the primary information matrix when converted to a vector form; and, IS(L) to the recovered image.

During the computing process, the primary information matrix is first obtained (statement number 40), and following which, the largest entry (FMAX) and the smallest entry (FMIN) are determined. The threshold (FLAVL) will be set when assigned an adjusting factor (SCALE). The cut-and-try technique is realized by the adjustment of the value of SCALE. Once a proper value of SCALE is assigned, the computer output will be the information content of the corresponding recovered image.

Measured Correlation Coefficients of Orthogonal Images

As previously noted, the coefficient of cross-correlation between any two zero-mean orthogonal images is theoretically zero, and the self-correlation coefficient of each image is 1. Accordingly, the correlation-coefficient matrix of a set of zero-mean orthogonal images should be ideally an identity matrix.

In order to test the accuracies of the correlation coefficients of the orthogonal images, a measurement was conducted with the set of orthogonal images shown in Figure 34. Table V shows the measured correlation-coefficient matrix which does not differ appreciably from an identity matrix.

TABLE V

MEASURED CORRELATION COEFFICIENTS OF ORTHOGONAL IMAGES

Recorder Image No.	Camera Image Number														
	(34.02)	(34.03)	(34.04)	(34.05)	(34.06)	(34.07)	(34.08)	(34.09)	(34.10)	(34.11)	(34.12)	(34.13)	(34.14)	(34.15)	(34.16)
(34.02)	.96	.14	.14	.14	.08	.10	.12	.08	.08	.16	.12	.08	.10	.12	.14
(34.03)	.14	.98	.12	.12	.14	.08	.12	.10	.12	.12	.12	.12	.10	.10	.12
(34.04)	.14	.10	.99	.12	.10	.12	.10	.06	.12	.14	.10	.14	.10	.12	.12
(34.05)	.08	.06	.08	.92	.08	.02	.08	.08	.10	.08	.10	.10	.08	.08	.10
(34.06)	.08	.12	.12	.16	.96	.10	.14	.06	.10	.12	.10	.08	.10	.12	.14
(34.07)	.12	.04	.12	.08	.12	.94	.10	.10	.10	.12	.14	.12	.12	.12	.14
(34.08)	.12	.12	.08	.18	.12	.16	.94	.06	.10	.14	.08	.12	.12	.12	.16
(34.09)	.10	.10	.10	-.02	.06	.10	.12	.94	.14	.04	.14	.02	.10	.12	.10
(34.10)	.08	.10	.10	.12	-.02	.08	.08	.08	.90	.10	.12	.16	.10	.08	.12
(34.11)	.12	.06	.10	.12	.10	.02	.10	.02	.14	.92	.14	.14	.10	.12	.14
(34.12)	.12	.12	.12	.12	.12	.14	.10	.12	.14	.18	.96	.16	.10	.14	.12
(34.13)	.10	.10	.10	.12	.10	.10	.12	.04	.10	.14	.12	1.00	.10	.06	.18
(34.14)	.06	.10	.08	.10	.12	.10	.10	.08	.04	.10	.12	.16	.90	.14	.12
(34.15)	.12	.10	.14	.10	.12	.12	.12	.08	.10	.08	.10	.08	.12	.96	.18
(34.16)	.12	.10	.06	.14	.10	.10	.14	.06	.08	.14	.12	.18	.12	.16	.96

Experiment One

In the first experiment in image reconstruction, the video cross-correlating subsystem is used to generate the feature vectors. The images shown in Figure 37 are used as the input (test) images. The set of reference images used consists of the set of orthogonal images shown in Figure 34, except the all-white image (34.01) which is not applicable to the operation of the cross-correlator. The corresponding correlator output voltages are listed in Table VI, in which each column is considered to be the feature vector of the input image identified by the corresponding test image number.

The computer program listed in Table IV is used to reconstruct the test images from the obtained feature vectors. Table VII shows an example of input data to be read into the IBM 1620 computer. The first 15 rows correspond to the set of 15 reference images. The next two rows indicated by "A" are referred to the pre-assigned scaling factors, and the following two rows indicated by "P" to the exponential powers. It is seen that the value of 1.0 is assigned to each element of the rows A and P. The last 20 rows correspond to the feature vectors of the ten test images to be reconstructed. In each row, except the row indicated by A or P, the left-most number indicates the corresponding image number.

Table VIII shows the computer output data which correspond to the primary information matrices of the recovered images. It is seen that in each matrix the largest element is found to be coincidently the entry in row 1 and column 1. This implies that the (1,1) fundamental plane of each recovered image will contain the "white", regardless of whether the corresponding (1,1) fundamental plane of the original image

TABLE VI

FEATURE VECTORS USED IN EXPERIMENT ONE

Reference Image No.	Test Image Number									
	(37.01)	(37.02)	(37.03)	(37.04)	(37.05)	(37.06)	(37.07)	(37.08)	(37.09)	(37.10)
(34.02)	1.5	1.9	1.1	1.0	-.8	.7	1.5	-.4	.9	2.5
(34.03)	.6	-.2	3.0	3.6	2.0	3.4	1.4	1.8	1.1	.5
(34.04)	3.2	-.5	1.0	1.0	-.9	.8	3.2	-.5	-2.6	.2
(34.05)	.1	.7	.7	.7	-1.0	-.5	-1.2	.7	-.4	-.1
(34.06)	1.1	.8	.7	1.8	.2	.7	1.3	.7	.8	.1
(34.07)	.4	.9	.8	.7	1.2	1.4	1.3	.7	.2	2.1
(34.08)	1.2	.9	.8	-.1	.3	.7	1.2	.8	2.1	2.0
(34.09)	2.0	3.5	2.7	-.5	2.5	1.8	1.5	3.9	.9	0
(34.10)	1.4	-.5	1.0	.7	1.2	.7	1.0	1.8	1.9	0
(34.11)	.4	1.4	-1.5	1.7	.1	-.6	1.1	-.5	.2	2.0
(34.12)	-.7	1.9	.8	.8	1.4	.7	-.7	2.0	1.1	1.9
(34.13)	.4	.6	.7	1.0	-1.0	-.6	-.9	.6	1.1	2.0
(34.14)	1.2	.7	.8	-.1	.1	.3	.9	.7	1.9	2.2
(34.15)	.5	.9	.8	.7	1.2	1.7	1.5	.8	.2	.5
(34.16)	1.3	.6	.8	1.6	.1	.5	1.0	.6	1.1	.5

TABLE VII

COMPUTER INPUT DATA USED IN EXPERIMENT ONE

```
34.02   1    1   -1   -1    1    1   -1   -1    1    1   -1   -1    1    1   -1   -1
34.03   1   -1   -1    1    1   -1   -1    1    1   -1   -1    1    1   -1   -1    1
34.04   1   -1    1   -1    1   -1    1   -1    1   -1    1   -1    1   -1    1   -1
34.05   1    1    1    1    1    1    1    1   -1   -1   -1   -1   -1   -1   -1   -1
34.06   1    1   -1   -1    1    1   -1   -1   -1   -1    1    1   -1   -1    1    1
34.07   1   -1   -1    1    1   -1   -1    1   -1    1    1   -1   -1    1    1   -1
34.08   1   -1    1   -1    1   -1    1   -1   -1    1   -1    1   -1    1   -1    1
34.09   1    1    1    1   -1   -1   -1   -1   -1   -1   -1   -1    1    1    1    1
34.10   1    1   -1   -1   -1   -1    1    1   -1   -1    1    1    1    1   -1   -1
34.11   1   -1   -1    1   -1    1    1   -1   -1    1    1   -1    1   -1   -1    1
34.12   1   -1    1   -1   -1    1   -1    1   -1    1   -1    1    1   -1    1   -1
34.13   1    1    1    1   -1   -1   -1   -1    1    1    1    1   -1   -1   -1   -1
34.14   1    1   -1   -1   -1   -1    1    1    1    1   -1   -1   -1   -1    1    1
34.15   1   -1   -1    1   -1    1    1   -1    1   -1   -1    1   -1    1    1   -1
34.16   1   -1    1   -1   -1    1   -1    1    1   -1    1   -1   -1    1   -1    1


A      1.   1.   1.   1.   1.   1.   1.   1.
A      1.   1.   1.   1.   1.   1.   1.


P      1.   1.   1.   1.   1.   1.   1.   1.
P      1.   1.   1.   1.   1.   1.   1.


37.01  1.5   0.6   3.2   0.1   1.1   0.4   1.2   2.0
37.01  1.4   0.4  -0.7   0.4   1.2   0.5   1.3
37.02  1.9  -0.2  -0.5   0.7   0.8   0.9   0.9   3.5
37.02 -0.5   1.4   1.9   0.6   0.7   0.9   0.6
37.03  1.1   3.0   1.0   0.7   0.7   0.8   0.8   2.7
37.03  1.0  -1.5   0.8   0.7   0.8   0.8   0.8
37.04  1.0   3.6   1.0   0.7   1.8   0.7  -0.1  -0.5
37.04  0.7   1.7   0.8   1.0  -0.1   0.7   1.6
37.05 -0.8   2.0  -0.9  -1.0   0.2   1.2   0.3   2.5
37.05  1.2   0.1   1.4  -1.0   0.1   1.2   0.1
37.06  0.7   3.4   0.8  -0.5   0.7   1.4   0.7   1.8
37.06  0.7  -0.6   0.7  -0.6   0.3   1.7   0.5
37.07  1.5   1.4   3.2  -1.2   1.3   1.3   1.2   1.5
37.07  1.0   1.1  -0.7  -0.9   0.9   1.5   1.0
37.08 -0.4   1.8  -0.5   0.7   0.7   0.7   0.8   3.9
37.08  1.8  -0.5   2.0   0.6   0.7   0.8   0.6
37.09  0.9   1.1  -2.6  -0.4   0.8   0.2   2.1   0.9
37.09  1.9   0.2   1.1   1.1   1.9   0.2   1.1
37.10  2.5   0.5   0.2  -0.1   0.1   2.1   2.0   0.0
37.10  0.0   2.0   1.9   2.0   2.2   0.5   0.5
```

TABLE VIII

PRIMARY INFORMATION MATRICES OBTAINED IN EXPERIMENT ONE

$$\begin{bmatrix} 14.5 & .7 & .4 & -5.7 \\ 1.6 & -6.1 & 1.4 & -5.9 \\ 2.8 & -5.7 & 1.8 & -5.5 \\ 2.2 & 1.9 & .8 & .9 \end{bmatrix}$$

(T8.01)

$$\begin{bmatrix} 13.5 & 1.7 & 1.7 & 1.9 \\ -4.5 & 2.8 & -6.3 & -5.3 \\ -5.5 & 3.0 & -6.9 & -4.7 \\ 1.3 & 2.7 & 2.7 & 1.7 \end{bmatrix}$$

(T8.02)

$$\begin{bmatrix} 14.2 & 1.2 & .8 & .1 \\ 1.9 & -7.3 & -6.9 & 1.5 \\ 2.1 & -7.1 & -7.1 & 1.3 \\ 1.9 & 1.8 & 1.0 & .3 \end{bmatrix}$$

(T8.03)

$$\begin{bmatrix} 14.5 & -5.3 & -5.5 & 1.1 \\ 2.7 & 1.9 & -5.3 & 1.3 \\ 2.9 & -4.5 & 2.3 & 2.3 \\ 1.9 & -6.3 & -5.7 & 1.3 \end{bmatrix}$$

(T8.04)

$$\begin{bmatrix} 6.5 & -4.1 & -3.7 & 3.3 \\ -4.5 & -4.1 & -4.5 & 3.3 \\ -5.1 & -3.9 & -4.7 & 3.9 \\ 4.3 & 4.7 & 4.7 & 4.0 \end{bmatrix}$$

(T8.05)

$$\begin{bmatrix} 11.6 & -5.4 & -4.8 & 1.4 \\ 2.6 & -5.2 & -5.4 & 1.2 \\ 1.8 & -6.4 & -5.8 & 2.8 \\ 3.2 & 3.2 & 3.0 & 1.8 \end{bmatrix}$$

(T8.06)

$$\begin{bmatrix} 14.0 & -5.8 & -5.8 & -4.7 \\ 3.2 & -5.0 & 1.3 & -6.6 \\ 3.1 & -5.3 & 1.9 & -4.5 \\ 3.8 & 3.8 & 3.8 & 2.6 \end{bmatrix}$$

(T8.07)

$$\begin{bmatrix} 13.6 & 2.2 & 2.4 & 2.2 \\ -6.0 & -5.8 & -6.0 & 2.9 \\ -6.4 & -5.8 & -6.8 & 3.3 \\ 2.4 & 2.6 & 2.8 & 2.2 \end{bmatrix}$$

(T8.08)

$$\begin{bmatrix} 10.4 & 3.7 & -3.9 & -3.8 \\ -6.2 & -2.6 & -3.8 & 3.2 \\ -3.1 & 4.4 & -5.0 & 6.0 \\ -3.4 & 4.1 & -5.4 & 5.6 \end{bmatrix}$$

(T8.09)

$$\begin{bmatrix} 16.3 & -2.9 & -3.3 & -2.3 \\ -1.7 & -1.5 & -2.7 & -2.1 \\ .4 & 12.9 & -2.5 & -2.3 \\ -2.1 & -1.1 & -2.3 & -1.7 \end{bmatrix}$$

(T8.10)

being black or white. It is, therefore, predicted that the error may be found in the (1,1) fundamental plane of the recovered image.

The fact that the (1,1) fundamental plane of a recovered image tends to be the white may be related to the fact that all of the (1,1) fundamental planes of the orthogonal images shown in Figure 34 contain identically the white. Because of this undesirable performance, the (1,1) entry will not be taken into account in the process of searching the largest entry among the entires in the primary information matrix. Therefore, in this experiment, the statement 45 in the program listed in Table IV is changed to be read as FMAX = T(2) instead of FMAX = T(1), in order to assure proper operation of the computing process.

Table IX shows the computer output data which indicate the information matrices of the recovered images. The recovered images are shown in Figure 38, in which those fundamental planes with an asterisk * are wrongly reconstructed. The value of 0.5 is found to be the proper one to be assigned to the threshold adjusting factor which was denoted as SCALE in Table IV. The proper threshold levels are found to be -1.7, -2.0, -2.7, -1.7, -0.2, -1.6, -1.4, -1.8, -0.1, and 4.8, corresponding to the primary information matrices (T8.01),(T8.02),...,(T8.10), respectively.

In Figure 38, it is seen that 7 out of 10 test images are correctly recovered, and that in each of the remaining three images (38.05), (38.09), and (38.10), only the (1,1) fundamental plane is wrongly recovered. The error shown in the information content of the (1,1) fundamental planes of the three wrongly reconstructed images is in agreement with that which was predicted.

TABLE IX

INFORMATION MATRICES OF·RECOVERED IMAGES

$$\begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(T9.01)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(T9.02)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(T9.03)

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

(T9.04)

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(T9.05)

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(T9.06)

$$\begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(T9.07)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

(T9.08)

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$

(T9.09)

$$\begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{bmatrix}$$

(T9.10)

(38.01)   (38.02)   (38.03)

(38.04)   (38.05)   (38.06)

(38.07)   (38.08)   (38.09)

(38.10)
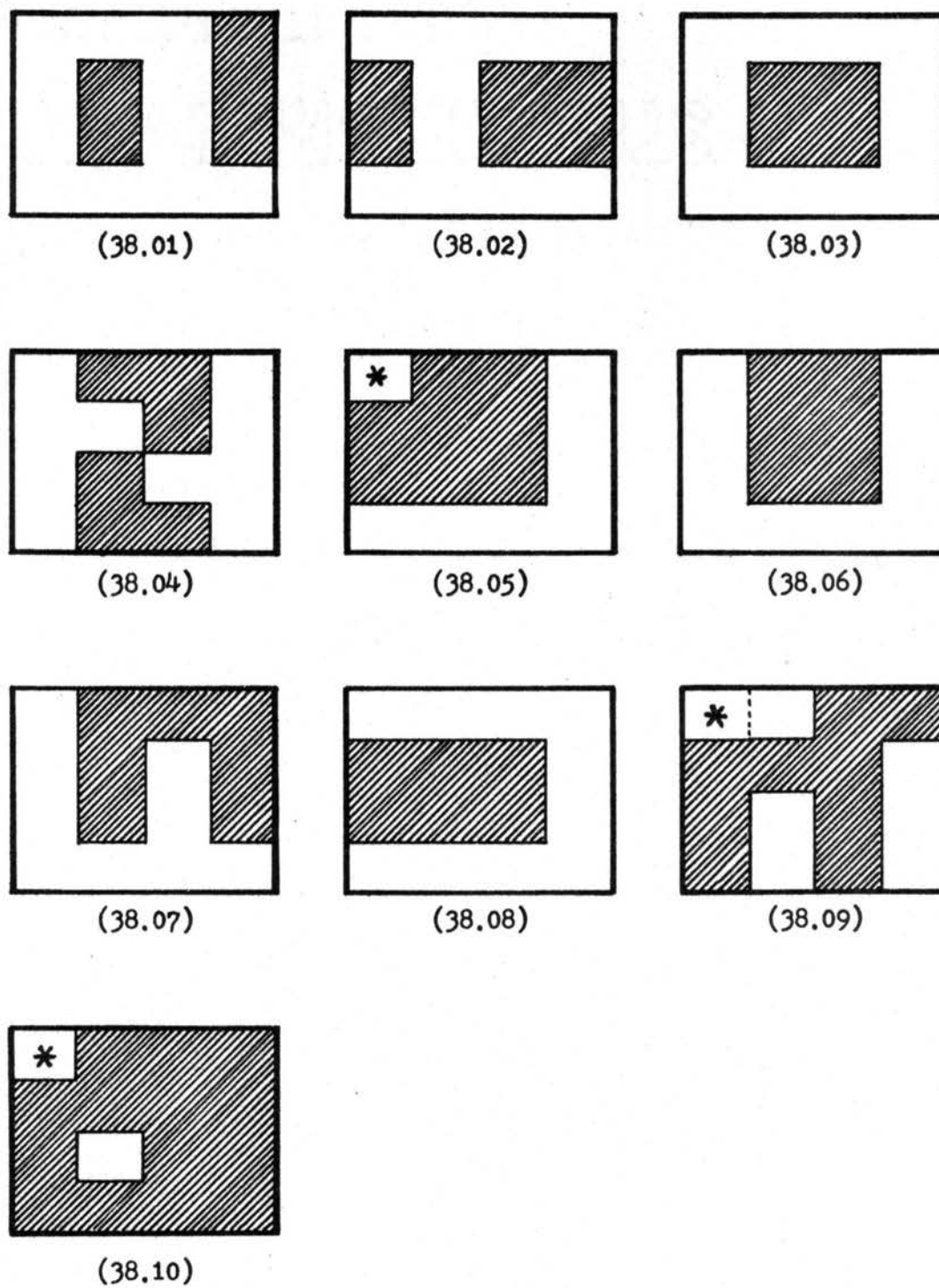
Figure 38.   Recovered Images

Experiment Two

The second experiment is similar to the first experiment, except that in this case the feature vectors are generated by the video waveform-agreement-detection subsystem. The corresponding output voltages of the waveform agreement detector are shown in Table X, in which each column represents the feature vector of the input image identified by its test image number.

The computer input data used in this experiment are listed in Table XI, in which the first 19 rows are exactly the same as those shown in Table VII, and in which the last 20 rows correspond to a different set of feature vectors for the same set of test images used in the first experiment.

Table XII indicates the computer output data which correspond to the primary information matrices of the recovered images. Since the same set of reference images is used in this experiment, and since all of the elements of the feature vectors generated by the waveform agreement detector are non-negative, the (1,1) element in each primary information matrix will always be the largest entry. Therefore, the error may be found in the (1,1) fundamental plane of a recovered image. For this reason, the threshold is set at the midpoint between the secondly largest entry and the minimum entry of each primary information matrix. The threshold levels are found to be -6.2, -5.9, -5.9, -5.6, -4.7, -5.3, -5.5, -6.1, -4.7, and -3.1, corresponding to the primary information matrices (T12.01),(T12.02),...,(T12.10), respectively.

The information matrices of the recovered images obtained in the second experiment is found to be exactly the same as those shown in Table IX. Therefore, the corresponding recovered images will be the

TABLE X

FEATURE VECTORS USED IN EXPERIMENT TWO

Reference Image No.	Test Image Number									
	(37.01)	(37.02)	(37.03)	(37.04)	(37.05)	(37.06)	(37.07)	(37.08)	(37.09)	(37.10)
(34.02)	5.2	5.8	4.5	4.8	3.3	4.6	5.7	3.9	5.3	6.1
(34.03)	4.5	4.2	7.1	8.0	6.6	7.9	5.4	6.3	5.5	5.2
(34.04)	6.9	3.8	4.5	4.8	3.3	4.5	7.3	3.9	2.1	4.9
(34.05)	4.6	5.2	5.2	5.0	3.4	3.7	3.2	5.3	4.0	4.8
(34.06)	5.3	5.1	4.6	5.9	4.6	4.9	5.8	5.1	5.3	5.0
(34.07)	4.7	5.1	5.2	5.2	5.5	5.7	5.5	5.1	5.1	5.6
(34.08)	5.2	5.0	4.8	4.0	4.5	4.7	5.4	5.2	6.2	5.6
(34.09)	6.5	8.5	7.1	3.9	6.8	6.0	5.8	8.3	5.0	4.6
(34.10)	5.6	3.8	5.0	4.8	5.4	4.8	5.2	6.0	6.3	4.9
(34.11)	4.6	5.6	3.1	5.9	4.4	3.8	5.3	4.1	4.5	5.6
(34.12)	3.8	6.1	5.0	4.9	5.5	4.8	3.6	6.4	5.2	5.6
(34.13)	4.8	5.2	5.0	5.1	3.6	3.8	3.3	5.2	5.1	5.9
(34.14)	5.6	4.9	4.8	4.0	4.3	4.4	5.1	5.0	6.2	5.9
(34.15)	4.6	5.2	4.9	4.9	5.5	5.9	5.6	5.1	4.4	5.1
(34.16)	5.4	4.8	4.8	5.6	4.3	4.4	5.0	5.0	5.2	5.0

TABLE XI

COMPUTER INPUT DATA USED IN EXPERIMENT TWO

34.02	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1
34.03	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1
34.04	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1
34.05	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1
34.06	1	1	-1	-1	1	1	-1	-1	-1	-1	1	1	-1	-1	1	1
34.07	1	-1	-1	1	1	-1	-1	1	-1	1	1	-1	-1	1	1	-1
34.08	1	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1
34.09	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1
34.10	1	1	-1	-1	-1	-1	1	1	-1	-1	1	1	1	1	-1	-1
34.11	1	-1	-1	1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	1
34.12	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	1	-1	1	-1
34.13	1	1	1	1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1
34.14	1	1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1
34.15	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1
34.16	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1

A	1.	1.	1.	1.	1.	1.	1.	1.
A	1.	1.	1.	1.	1.	1.	1.	

P	1.	1.	1.	1.	1.	1.	1.	1.
P	1.	1.	1.	1.	1.	1.	1.	

37.01	5.2	4.5	6.9	4.6	5.3	4.7	5.2	6.5
37.01	5.6	4.6	3.8	4.8	5.6	4.6	5.4	
37.02	5.8	4.2	3.8	5.2	5.1	5.1	5.0	8.5
37.02	3.8	5.6	6.1	5.2	4.9	5.2	4.8	
37.03	4.5	7.1	4.5	5.2	4.6	5.2	4.8	7.1
37.03	5.0	3.1	5.0	5.0	4.8	4.9	4.8	
37.04	4.8	8.0	4.8	5.0	5.9	5.2	4.0	3.9
37.04	4.8	5.9	4.9	5.1	4.0	4.9	5.6	
37.05	3.3	6.6	3.3	3.4	4.6	5.5	4.5	6.8
37.05	5.4	4.4	5.5	3.6	4.3	5.5	4.3	
37.06	4.6	7.9	4.5	3.7	4.9	5.7	4.7	6.0
37.06	4.8	3.8	4.8	3.8	4.4	5.9	4.4	
37.07	5.7	5.4	7.3	3.2	5.8	5.5	5.4	5.8
37.07	5.2	5.3	3.6	3.3	5.1	5.6	5.0	
37.08	3.9	6.3	3.9	5.3	5.1	5.1	5.2	8.3
37.08	6.0	4.1	6.4	5.2	5.0	5.1	5.0	
37.09	5.3	5.5	2.1	4.0	5.3	5.1	6.2	5.0
37.09	6.3	4.5	5.2	5.1	6.2	4.4	5.2	
37.10	6.1	5.2	4.9	4.8	5.0	5.6	5.6	4.6
37.10	4.9	5.6	5.6	5.9	5.9	5.1	5.0	

TABLE XII

PRIMARY INFORMATION MATRICES OBTAINED IN EXPERIMENT TWO

$$\begin{bmatrix} 77.2 & -2.1 & -2.8 & -8.6 \\ -4.4 & -10.3 & -3.0 & -8.9 \\ -3.2 & -9.5 & -2.6 & -9.7 \\ -3.0 & -2.9 & -2.4 & -3.1 \end{bmatrix}$$

(T12.01)

$$\begin{bmatrix} 78.2 & -1.2 & -1.1 & -.3 \\ -9.8 & -2.7 & -11.3 & -10.1 \\ -10.4 & -2.8 & -11.5 & -9.1 \\ -2.6 & -1.9 & -.9 & -2.1 \end{bmatrix}$$

(T12.02)

$$\begin{bmatrix} 75.5 & -3.1 & -2.7 & -.4 \\ -3.8 & -11.3 & -10.9 & -1.3 \\ -4.4 & -10.8 & -11.1 & -2.7 \\ -3.0 & -3.0 & -3.3 & -3.0 \end{bmatrix}$$

(T12.03)

$$\begin{bmatrix} 76.7 & -9.7 & -10.2 & -.7 \\ -1.3 & -2.7 & -10.0 & -1.7 \\ -2.4 & -8.9 & -2.2 & -1.5 \\ -2.5 & -10.3 & -9.5 & -2.2 \end{bmatrix}$$

(T12.04)

$$\begin{bmatrix} 70.9 & -8.1 & -8.2 & .6 \\ -8.5 & -8.9 & -9.4 & -1.0 \\ -9.1 & -8.8 & -8.8 & .3 \\ -.3 & -.4 & 0.0 & 0.0 \end{bmatrix}$$

(T12.05)

$$\begin{bmatrix} 73.8 & -9.4 & -10.1 & -.3 \\ -1.8 & -9.6 & -10.3 & -2.4 \\ -2.8 & -10.2 & -10.1 & -.2 \\ -1.0 & -1.7 & -1.5 & -1.6 \end{bmatrix}$$

(T12.06)

$$\begin{bmatrix} 77.1 & -8.9 & -10.0 & -9.0 \\ -.5 & -8.7 & -3.0 & -11.1 \\ -2.3 & -9.3 & -2.3 & -8.6 \\ -.5 & -.8 & .1 & -1.5 \end{bmatrix}$$

(T12.07)

$$\begin{bmatrix} 79.8 & -2.2 & -1.2 & -1.0 \\ -10.3 & -10.1 & -10.6 & -1.6 \\ -11.1 & -10.1 & -11.0 & -1.3 \\ -2.1 & -2.7 & -2.1 & -1.8 \end{bmatrix}$$

(T12.08)

$$\begin{bmatrix} 75.3 & -.9 & -9.7 & -8.1 \\ -8.4 & -7.6 & -8.0 & -.4 \\ -7.8 & -.2 & -8.2 & .5 \\ -7.5 & -.3 & -8.8 & .3 \end{bmatrix}$$

(T12.09)

$$\begin{bmatrix} 79.7 & -5.3 & -7.0 & -6.1 \\ -5.3 & -5.4 & -6.2 & -5.8 \\ -3.5 & .7 & -6.0 & -5.2 \\ -5.9 & -6.0 & -6.4 & -6.0 \end{bmatrix}$$

(T12.10)

same images shown in Figure 38, in which 7 out of 10 test images are
correctly recovered, and in which the error is found to be in those
(1,1) fundamental planes with an asterisk *, exactly being in agreement
with that which was predicted.

Comparison of Experimental Results

It is seen that with the same sets of test and reference images,
the set of recovered images obtained in the first experiment is the
same as the one obtained in the second experiment. This means that the
feature vectors used in the first experiment and the corresponding
feature vectors used in the second experiment give similar information.

The identical experimental results may be due to the fact that the
set of zero-mean orthogonal images is applicable to the operation of
both the cross-correlating subsystem and the waveform-agreement-
detection subsystem. As previously mentioned in Chapter V, the applica-
bility of the set of zero-mean orthogonal images in the operation of
the cross-correlating subsystem is based on the fact that the coeffi-
cient of cross-correlation between any two orthogonal images is zero.

However, further inspection on the set of zero-mean orthogonal
images shown in Figure 34 shows that for any two zero-mean orthogonal
images, the subtotal area of waveform agreement and the subtotal area
of waveform disagreement share equally one-half of the total image area.
This implies that when the set of zero-mean orthogonal images is used
simultaneously as the sets of reference and test images, the desired
output voltage of the waveform-agreement-detector will have the largest
value (e.g. 15 volts) for the cases of identical images, and will be
one-half of the largest value (e.g. 7.5 volts) for the cases of

different images. The corresponding matrix will be a symmetric matrix in which the diagonal elements each have a unique value of 15, and the off-diagonal elements each have another unique value of 7.5, according to the desired performance of the waveform-agreement-detector described in Chapter IV. If each element of the mentioned matrix is first subtracted by a constant of 7.5 and then divided by a factor of 7.5, the matrix thus obtained will be equivalent to the corresponding correlation-coefficient matrix which is an identity matrix. It is this property that makes the set of zero-mean orthogonal images also applicable to the operation of the waveform-agreement-detection subsystem in the task of image reconstruction.

Figure 39 shows the similarity between the feature vectors generated by the cross-correlating subsystem and the feature vectors generated by the waveform-agreement-detection subsystem. In each plot of Figure 39, the upper trace indicates the feature vector generated by the waveform agreement detector, and the lower trace indicates the feature vector generated by the cross-correlator. The feature vectors shown in Figure 39 were previously listed in Tables VI and X. The plots (39.01),(39.02),...,(39.10) correspond to the test images (37.01), (37.02),...,(37.10), respectively. It is seen that the variances of the upper and lower traces in each plot of Figure 39 are nearly equal or proportional. Thus, if the set of zero-mean orthogonal images is used as the set of reference images in the case of using the cross-correlating subsystem and in the case of using the waveform-agreement-detection subsystem, the results of image reconstruction obtained in both cases will be exactly or nearly identical.

(39.01) (39.02)

(39.03) (39.04)

Ampli-
tude
(volts)

(39.05) (39.06)

(39.07) (39.08)

(39.09) (39.10)
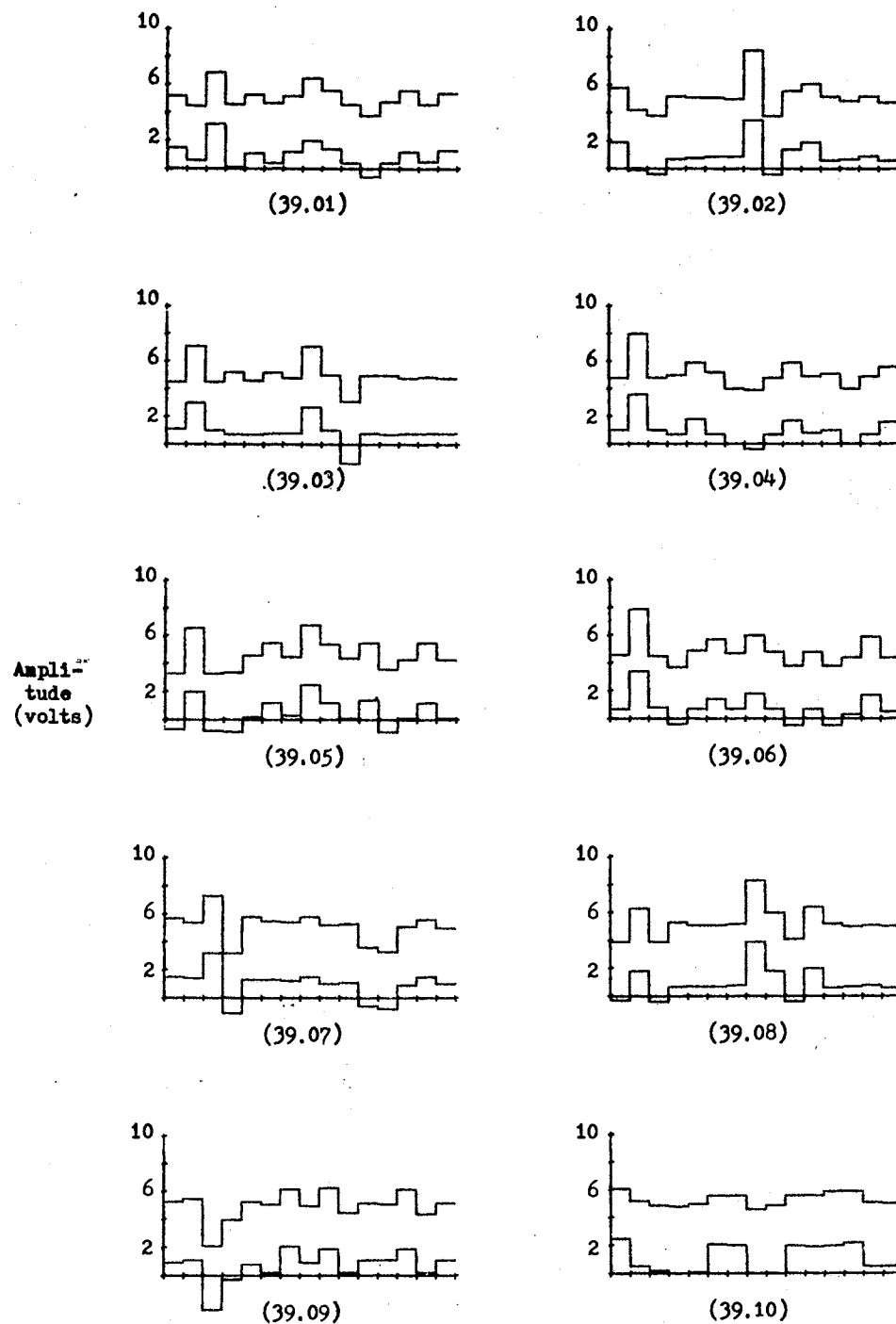
Upper Traces: Output Voltages of Waveform-Agreement-Detector
2.0 Volts/Div
Lower Traces: Output Voltages of Cross-Correlator
2.0 Volts/Div

Horizontal Scale: Each division corresponds to one
reference image.

Figure 39. Plots of Feature Vectors

CHAPTER VII

SUMMARY AND CONCLUSIONS

The video image reduction studies reported in this thesis can be summarized by the following items:

1. A literature review was made to examine previous works in the areas of artificial intelligence and pattern recognition.

2. The basic video camera and recording equipment used in the research was described. Special circuits were developed to ensure continuous synchronism between the camera and recorder scanning signals.

3. A specialized video cross-correlating subsystem was designed, constructed and tested in order to convert a camera-sensed input image into a corresponding feature vector.

4. As a parallel approach, a video waveform-agreement-detection subsystem was constructed and tested in order to convert a camera-sensed input image into a corresponding feature vector.

5. A mathematical algorithm for image reconstruction was developed and illustrated with equations. The selection of reference images was discussed, and an algorithm for the construction of orthogonal images was presented.

6. Experiments in image reconstruction were illustrated, using computer simulation techniques. The experimental results were analyzed and discussed.

It was found that both the video cross-correlating subsystem and

the video waveform-agreement-detection subsystem had the property of converting a camera-sensed input image into a corresponding feature vector. As verified by the image-reconstruction experiments, the feature vectors obtained by these two video subsystems were seen to be valid representations of two-dimensional optical images. Since there is only a limited set of significant features of interest contained in a given input image, the number of reference images needed for converting the input image into a meaningful feature vector should be much less than the number of the image resolution elements. Thus, by efficient selection of the reference images, the number of bits required to define the feature vector of a given input image can be made to be substantially less than the number of bits conventionally required to represent that input image. Therefore, it is believed that these two video subsystems can be used as image preprocessors to reduce the dimensionality of the input patterns without significantly sacrificing their information content. It was also found that the mathematical image-reconstruction algorithm was applicable to the task of recovering a test image from its corresponding feature vector, and that the set of orthogonal images had specific features which were advantageous in the processing of image reconstruction.

It is felt that the studies reported in this thesis have shown that the image-reduction techniques hold considerable promise for use in optical pattern preprocessing operations. On the basis of the studies reported herein, it is believed that the development of automatic image-coding techniques, as well as the design of video hybrid systems for automatic pattern recognition, would be fruitful.

BIBLIOGRAPHY

1. Rosenblatt, Frank. "Perceptron Simulation Experiments." Proceedings of the Institute of Radio Engineers, Vol. 48, No. 3 (March, 1960), 301-309.

2. Widrow, Bernard and Marcian E. Hoff. "Adaptive Switching Circuits." Wescon Convention Record, Part 4 (August, 1960), 96-104.

3. Kazmierczak, H. and K. Steinbuch. "Adaptive Systems in Pattern Recognition." IEEE Transactions on Electronic Computers, Vol. EC-12 (December, 1963), 822-835.

4. Bolie, Victor, W. "A Cognitive Learning Program Based on Alphabets of Linearly Independent Symbols." Proceedings of the National Aerospace Electronics Conference, (May, 1965), 381-390.

5. Bolie, Victor W. "Experiments in Machine Learning." College of Engineering, Oklahoma State University (1970).

6. Cutrona, L. J., E. N. Leith, C. J. Palermo and L. J. Porcello. "Optical Data Processing and Filtering Systems." IRE Transactions on Information Theory, Vol. IT-6 (June, 1960), 386-400.

7. Montgomery, W. D. and P. W. Broome. "Spatial Filtering." Journal of the Optical Society of America, Vol. 52, No. 11 (November, 1962), 1259-1275.

8. Vander Lugt, A. "Signal Detection by Complex Spatial Filtering." IEEE Transactions on Information Theory, Vol. IT-10 (April, 1964), 139-145.

9. Tretiak, Oleh J., Murray Eden, and William Simon. "Internal Structure From X-Ray Images." 8th ICMBE Pattern Recognition and Image Processing, Session 12-1, Palmer House, Chicago, Illinois (July 20 - 25, 1969).

10. Howell, P. N. "Optical Cross-Correlation Studies of Two-Dimensional Patterns." (unpub. Ph.D. thesis, Oklahoma State University, 1969).

11. Morris, D. E. "A Computer-Controlled Video System for Image Analysis and Pattern Recognition." (unpub. Ph.D. thesis, Oklahoma State University, 1970).

12. Nagy, George. "State of the Art in Pattern Recognition." Proceedings of the IEEE, Vol. 56, No. 5 (May, 1968), 836-862.

13. Levine, Martin D. "Feature Extraction: A Survey." Proceedings of the IEEE, Vol. 57, No. 8 (August, 1969), 1391-1407.

14. Kanal, Laveen N. Pattern Recognition. Proceedings of the IEEE Workshop on Pattern Recognition, Dorado, Puerto Rico. Washington, D.C.: Thompson Book Co., 1968.

15. Cheng, G. C., R. S. Ledley, D. K. Pollock, and A. Rosenfeld. Pictorial Pattern Recognition. Proceedings of Symposium on Automatic Photointerpretation sponsored by the Office of Naval Research, the University of Maryland, and the Pattern Recognition Society. Washington, D.C.: Thompson Book Co., 1968.

16. Service Manual for Sony Video Camera Model AVC-3200 and Viewfinder Model AVF-3200. Sony Corporation, 1970.

17. Service Manual for Sony Videocorder Model AV-3600. Sony Corporation, 1970.

18. Service Manual for Sony Video Monitor Model CVM-110U. Sony Corporation, 1970.

19. Luxenberg, H. R. and R. L. Kuehn. Display Systems Engineering. New York: McGraw-Hill, 1968.

20. Bolie, Victor W. and Wen-Shyong Su. "A Microelectronic CrossCorrelation Computer for Real-Time Video Pattern Abstraction." SWIEEECO Record of Technical Papers, Houston, Texas (April, 1971).

21. Bolie, Victor W. "Video Pattern Recognition System." (Patent Application in Process).

APPENDIX A

A MICROELECTRONIC CROSS-CORRELATION COMPUTER FOR REAL-TIME

VIDEO PATTERN ABSTRACTION

The manuscript which comprises this appendix is a detailed description of the video cross-correlating subsystem which was constructed as a part of this thesis. It was presented at the Southwestern Institute of Electrical and Electronics Engineers Conference held in Houston, Texas, during April 28 - 30, 1971.

A CROSS-CORRELATING VIDEO SYSTEM FOR
RAPID ABSTRACTION OF OPTICAL PATTERNS

by

Victor W. Bolie[1] and Wen-Shyong Su[2]

I. Introduction

It is by now well known that the complexities of implementing various "nearest neighbor" algorithms for automatic recognition of non-trivial video images are substantially reducible only through the use of preprocessors which can ideally minimize the dimensionality of the input patterns without sacrificing their information content. Recent studies (1-15) have shown that dimensionality reduction by conventional methods requires too much computing time to be economical for arbitrary video patterns.

One approach to the ideal preprocessor is through the use of a video cross-correlator and a set of stored reference images to convert each input image into a corresponding feature vector. Thus, if the sequence of reference images are linearly independent then each of the successive elements of the feature vector will convey a separate attribute of the input pattern, and the preprocessing operations will be economical if high-speed analog techniques can be used to generate each cross-correlation coefficient.

The present paper describes the design and operating characteristics of a video feature-abstraction system which generates a series of self-normalized coefficients of cross-correlation between an image sensed by

---

[1]Professor (Albrecht Naeter Chair), Oklahoma State University

[2]Doctoral Candidate in Electrical Engineering, Oklahoma State University

a television camera and a series of reference images retrieved from a television tape recorder. The complete equipment includes a subsystem which establishes and maintains synchronism between the camera scanning beam and the playback video head, as well as a subsystem which prevents the inherent horizontal and vertical sync and blanking pulses from entering the correlator.

II. The Overall System

The overall system of which the cross-correlation computer is a part is illustrated schematically in Figure 1. The video camera and video recorder each furnish through a 75-ohm coaxial cable a ±0.5 volt standard video signal. The recorder video output can be observed on the monitor, which contains internal circuitry to extract the horizontal and vertical sync pulses contained in the recorder video output. The special sync unit synchronizes the camera scan with the recorder video output by use of the sync signals extracted by the monitor. The recorder and camera video outputs are fed to a pair of sync-clipping video amplifiers, which remove the sync pulses and amplify the video to a scale of +15 volts for "black" and -5 volts for "white." The outputs of the two sync-clipping video amplifiers are used as the signal inputs to a pair of correlator pre-conditioners. Each correlator pre-conditioner subtracts out the DC-average value of its clipped video input, and clamps-to-zero any residual signal existing in either a horizontal or vertical blanking-pulse interval. The clamping control signals for each correlator pre-conditioner is obtained from a clamping pulse generator which is controlled by signals generated in the special sync unit.

The wideband output signals of the two pre-conditioners are analog voltages α and β which vary within the ranges of $-10 \leq \alpha \leq +10$ volts and $-10 \leq \beta \leq +10$ volts. The pre-conditioner output voltages α and β are the two signal inputs to the cross-correlator circuit. The cross-correlator circuit produces the output voltage E_0, which falls within the range $-5 \leq E_0 \leq +5$ volts, and which is given by the formula

$$E_0 = 5 \frac{\overline{\alpha\beta}}{\sqrt{\overline{\alpha^2} \cdot \overline{\beta^2}}}$$,

in which the superbar denotes the running time-averaging function of a first-order low-pass filter having a cut-off frequency not higher than 1.0 Hz.

In the overall video system illustrated in Figure 1, the camera, recorder, and monitor units actually used for the studies reported here consisted of the Sony AVC-3200 video camera, the Sony AV-3600 video recorder, and the Sony CVM-110U video monitor. The remaining units were constructed in the laboratory by use of discrete solid-state components.

III. Recorder-Camera Synchronism

The output voltage E_0 of the overall video system of Figure 1 will be a valid representation of the cross-correlation of the camera and recorder video outputs only if the camera scanning beam is synchronized with the playback video signal. The synchronization is achieved by making use of the horizontal and vertical sync pulses available internally in the monitor. The horizontal sync pulses (H) developed in the monitor comprise a 15,750 Hz train of 14-microsecond 15-volt positive

pulses superimposed on a -1.5 volt DC level. The vertical sync pulses

(V) developed in the monitor comprise a 60 Hz train of 0.5-millisecond

6-volt positive pulses superimposed on a +3.5 volt DC level.

The structure of the special sync unit is shown schematically in

Figure 2. The input horizontal sync pulses (H) are fed into a horizon-

tal sync delay circuit, which produces delayed horizontal-sync output

pulses (H´) as well as auxiliary square-wave output triggering pulses

(E_T). The input vertical pulses (V) are fed into a vertical sync in-

verter circuit, which produces inverted vertical-sync output pulses (V´);

the input vertical pulses (V) are also used as auxiliary triggering

pulses.

Figure 3a shows the detail of the vertical sync inverter circuit;

when connected for operation the waveform at the common node of the

15-microfarad capacitor and the 0.82-kilohm and 1.0-kilohm resistors is

a 60 Hz train of 4-volt negative pulses superimposed on the DC level of

E_{cc} = +15 volts. Figure 3b shows the detail of the horizontal sync

delay circuit, which consists of a pair of monostable multivibrators

and an input trailing-edge trigger. Upon arrival of the trailing edge

of each rectangular input pulse (H), the collector of the normally-on

transistor in the first mono swings from its 0.7-volt quiescent level

to the positive maximum level of +13.7 volts, where it remains for a

time lapse of 40 microseconds. This collector voltage furnishes the

auxiliary output signal E_T, which is an asymmetrical square wave having

a period of $(10^6)/(15,750)$ = 63.5 microseconds. By use of a 10K -

1000pF differentiator, the trailing edge of this asymmetrical square

wave is used to trigger the initiation of a 14-microsecond pulse formed

by the second mono. When connected for operation the waveform at the

common node of the 0.01-microfarad capacitor and the 1.9-kilohm and 3.6-kilohm resistors is a 15,750 Hz train of 7-volt positive pulses superimposed on a +5 volt DC level.

The reason for the 54-microsecond total delay of the two monos in the horizontal sync delay circuit of Figure 3b is that due to inherent delays in the camera circuitry the horizontal sync pulse in the camera video output does not appear until approximately 10 microseconds after the arrival of the camera input sync pulse. It was found experimentally that the delay of the vertical sync pulse by the camera circuitry was in the range of 100 to 200 microseconds, i.e., a time lapse equal to the duration of several horizontal sweeps. However, in view of the difficulties of precisely centering the optical image sensed by the camera, and in view of the inaccuracies of the remaining parts of the system, further improvements in the alignment of the camera and recorder scanning-voltage waveforms were not attempted.

IV. Sync-Clipping Video Amplifiers

Figure 4a shows the circuit details for each of the sync-clipping video amplifiers. The input signal voltage E_i is a ± 0.5 volt composite video signal, as represented by the upper trace in Figure 4b. The circuit output voltage E_a varies within the range $-5 \leq E_a \leq +15$ volts, as represented by the lower trace in Figure 4b.

The function of the sync-clipping video amplifier is to clip off the horizontal and vertical sync pulses present in the composite-video input, and to amplify the resulting signal to a 20-volt peak-to-peak amplitude. The capacitance input coupling has a low-frequency cutoff of 0.84 Hz. The first amplifier stage has a gain of ten, and, with the

power supply voltages $\pm E_{cc}$ = ± 15 volts, the collector voltage of the transistor Q2 is 9 $\pm$ 5 volts.

The coupling capacitor C2 operates in conjunction with the transistor Q3 and the resistances R5 and R6 so as to bias the "ceiling" of the amplified-and-inverted video signal to a level which is about 0.7 volt above the Q3 base voltage. Thus, by adjustment of the resistance R6 the maximum positive excursion of the inverted video signal voltage applied to the base of transistor Q4 can be set at any level between zero and +15 volts.

Since the Q4 emitter is connected through the resistance R10 to the ground bus, no current can flow through the resistances R8 and R9 into the $-E_{cc}$ bus unless the instantaneous voltage applied to the Q4 base is negative with respect to ground. This arrangement makes it easy to clip off any desired portion of the positive-going sync pulses existing in the inverted video signal, by adjustment of the "clipping" potentiometer R6.

The values of the resistances R8, R9, and R10 are chosen so that by adjustment of the "gain" potentiometer R9 a signal gain of 2.0 is readily achieved when transistor Q4 is in the conducting state. Consequently, when the "clipping" potentiometer R6 is adjusted so as to barely eliminate the sync pulses, and the "gain" potentiometer R9 is adjusted for a gain of about 2.0, the signal voltage applied to the base of transistor Q5 will vary between -15 and -5 volts. The resistances R11 and R12 are chosen to give a signal gain of about 2.0, so that the output voltage E_a at the Q5 collector varies between -5 and +15 volts.

Thus, as illustrated by the lower trace in Figure 4b, the sync-clipping video amplifier can be adjusted so as to convert the ± 0.5 volt composite video signal into an amplified sync-free video-signal output voltage E_a which varies through the range $-5 \leq E_a \leq +15$ volts. The response speed and temperature stability of the complete amplifier are more than sufficient for the purposes of the studies reported here.

V. Correlator Pre-Conditioners

Figure 5 shows the circuit details for each of the correlator pre-conditioners. The input voltage E_a varies within the range of $-5 \leq E_a \leq +15$ volts, and the output voltage α varies within the range of $-10 \leq \alpha \leq +10$ volts. The average value $\bar{\alpha}$ of the output voltage α must be zero in order to assure proper operation of the cross-correlation circuit.

The first section of the correlator pre-conditioner, consisting of the first three transistors and their associated passive components, is designed so that the voltage E_K at the emitter of the third transistor is given by the formula

$$E_K = \frac{E_a - \bar{E}_a{}'}{2} \quad ,$$

in which the voltage $\bar{E}_a{}'$ is determined by the formula

$$E_{cc} - \bar{E}_a = \left(\frac{63.5 - 12}{63.5}\right)\left(\frac{16.7 - 1.0}{16.7}\right)\left(E_{cc} - \bar{E}_a{}'\right) \quad .$$

In this expression, the average value $\bar{E}_a$ of the input voltage E_a corresponds to the voltage at the base of the first transistor, as determined

by the 10K - 33μF low-pass filter. It should be noted that $\bar{E}_a$ is the 330-millisecond running average of the input (clipped video) signal voltage E_a, which has the value of E_a = +15 volts during the horizontal and vertical blanking-pulse intervals. Conversely, the related voltage $\bar{E}_a{}'$ is the mean value of the input signal voltage E_a, as averaged over all the time except the blanking intervals in the 330-msec time span. When simplified, the formula for $\bar{E}_a{}'$ becomes $\bar{E}_a{}' = 1.3\bar{E}_a - 0.3E_{cc}$, which if substituted into the E_K formula gives $E_K = 0.15E_{cc} - 0.65\bar{E}_a + 0.50E_a$. It will be seen that this latter expression is realized by the choice of the passive-component values in the first section of the pre-conditioner, and that $-10 \leq E_K \leq +10$ volts.

The remaining (second) section of the correlator pre-conditioner has the function of making the output voltage α equal to the internal voltage E_K at all times except during the horizontal and vertical blanking-pulse intervals, when α is clamped to the zero-volt potential of the common ground bus G. The clamping is achieved by use of a simple analog-gating technique which utilizes ±12 volt horizontal blanking-pulse inputs H^+ and H^-, as well as ±12 volt vertical blanking-pulse inputs V^+ and V^-.

VI. Clamping-Pulse Generator

Figure 6 shows the scheme of the clamping-pulse generator, which produces the horizontal and vertical blanking pulses H^+, H^-, V^+, and V^- required by the two correlator pre-conditioners. The complete clamping-pulse generator is comprised of two sections, i.e., a vertical clamping-pulse generator which produces the outputs V^+ and V^-, and a horizontal clamping-pulse generator which produces the outputs H^+ and H^-.

Figure 7a shows the circuit details for the vertical clamping-pulse generator, which is seen to consist of a 1.0-msec monostable multivibrator which is operated from a ±15 volt power supply and is triggered through an emitter follower by the input vertical-sync pulses (V). Figure 7b shows the circuit of the horizontal clamping-pulse generator. In this case two monostable multivibrators are used. The first mono is triggered by the trailing edge of each input pulse E_T, and has a delay of 22 μsec, which is sufficient to reach to the beginning of the subsequent blanking interval. The second mono has a switched-on duration of 12 μsec, which is equal to the width of the horizontal blanking pulse in the original video signal. The component values used in the horizontal and vertical clamping-pulse generator circuits are chosen to give ample current-drive in the outputs H^+, H^-, V^+, and V^-.

VII. <u>Cross-Correlator Circuit</u>

Figure 8 shows the details of the cross-correlator circuit. The complete circuit makes use of four analog multipliers M1, M2, M3, and M4, two square-root units Q1 and Q2, one analog-division module D1, and one operational amplifier SN72709 arranged as a 0.4 Hz low-pass filter. Each of the modules M1, M2, M3, M4, Q1, Q2, and D1 used in the studies reported here consisted of a Burr-Brown Type 4097 Differential-Input Multiplier/Divider. The frequency response of the 4097 module extends from DC to 25 kHz for a one-degree phase shift, and from DC to 200 kHz for full-power output.

Using the ±10 volt analog input signals α and β, the first three multipliers M1, M2, and M3 are arranged to produce the output voltages $0.1\alpha^2$, $0.1\beta^2$, and $0.1\alpha\beta$, respectively. A 160-msec running time-average of each of these three output voltages is developed by a simple

1K - 160µF low-pass filter, to produce the three corresponding voltages $0.1\overline{\alpha^2}$, $0.1\overline{\beta^2}$, and $0.1\overline{\alpha\beta}$.

The multiplier M4 and the two square-root units Q1 and Q2 are arranged so as to develop at the output terminal of M4 the voltage $0.1\sqrt{\overline{\alpha^2} \cdot \overline{\beta^2}}$, which is used as the denominator-input to the analog-division module D1. The numerator input to D1 is made to be the time-averaged product voltage $0.1\overline{\alpha\beta}$, thereby causing the output voltage of D1 to be $-10\overline{\alpha\beta}/\sqrt{\overline{\alpha^2} \cdot \overline{\beta^2}}$.

The "709" operational amplifier with its associated passive components is arranged to serve as a 0.4 Hz low-pass filter having a DC gain of -0.5 when connected to the output terminal of the analog-division module D1. As a result, the output voltage E_0 of the complete cross-correlator circuit is the attenuated and additionally-smoothed cross correlation voltage $5\overline{\alpha\beta}/\sqrt{\overline{\alpha^2} \cdot \overline{\beta^2}}$.

VIII. System Characteristics

Figure 9 shows a photograph of the over-all video system which was illustrated in schematic form in Figure 1. The camera and its input "horizontal-step" image are shown on the left, and the recorder and monitor are shown on the right. The breadboard of remaining components of the system are shown with an output voltmeter in the center. The correlator output signal voltage E_0, and its associated zero-volt ground reference, are obtained from the two conductors hanging downward from the breadboard. The various components mounted on the breadboard include the necessary input-output connectors and ±15 volt power supply.

Validity testing of the overall video system can be performed with the aid of input step-images similar to that being monitored by the

camera in Figure 9. Some of the results obtained during such a testing procedure are shown in Figures 10 and 11.

Figure 10 shows the genesis of the left and right correlator-inputs α and β, for both identical and opposite horizontal-step images. In each of the four photographs, the upper trace shows the input to the sync-clipping video amplifier, the center trace shows the input to the corresponding correlator pre-conditioner, and the lower trace shows the corresponding input to the cross-correlator circuit. In Figures 10a and 10b, the optical input images sensed by the camera and generated by the recorder were identical and essentially the same as those shown in Figure 9. The negative-going horizontal sync pulses contained in the ± 0.5 volt video outputs of the camera and recorder are seen to be nearly absent in the amplified-and-inverted video signals represented by the center traces in Figures 10a and 10b. The lower traces of Figures 10a and 10b show the effectiveness of the zero-clamping sections of the correlator pre-conditioners. The information shown in Figures 10c and 10d is similar, except that the optical input image sensed by the camera was rotated by 180 degrees to put the black-half of the image on the left. The correlator output voltage E_0 was found to be +4.7 volts when the horizontal-step images were identical (Figures 10a and 10b), and was found to be -4.6 volts when the camera image was reversed (Figures 10c and 10d).

Figure 11 similarly shows the genesis of the left and right correlator-inputs α and β, for the case of identical and opposite vertical-step images. In Figures 11a and 11b, the camera and playback images were identical and essentially the same as those ·shown in Figure 9, except that each image was rotated by 90 degrees to put the white-half

of the image at the top. The negative-going vertical sync pulses contained in the video outputs of the camera and recorder are similarly seen to be essentially absent in the amplified-and-inverted video signals represented by the center traces of Figures 11a and 11b. The lower traces of Figures 11a and 11b show further the effectiveness of the zero-clamping circuit. The information shown in Figures 11c and 11d is similar, except that the playback image generated by the recorder was rotated by 180 degrees to put the white-half of the image at the bottom. The correlator output voltage E_0' was found to be +4.9 volts when the vertical-step images were identical (Figures 11a and 11b), and was found to be -4.5 volts when the playback image was reversed.

IX. System Performance

The overall performance of the cross-correlating video system is illustrated in the 8-by-8 matrix arrangement of Figure 12. The various images sensed by the camera are shown in the top row of the diagram. Similarly, the various images retrieved by the recorder are shown in the left column. Each numerical entry in the 64-element matrix is the corresponding output voltage E_0 of the cross-correlator circuit.

The numerical entries on the main diagonal of the matrix of Figure 12 are the experimentally obtained self-correlation coefficients, multiplied by the common scale factor of 5. It is seen that the values of the off-diagonal elements do not differ appreciably from those of the ideal symmetric matrix. The greatest error is found in the non-zero entry of 0.9 volt in row 6 and column 8. In general, the accuracies of the various cross-correlation voltages are seen to be compatible with the normal tolerances of the electronic components used.

Thus, the characteristics of the overall video feature-extraction system described herein are seen to be compatible with the requirement that it be capable of generating a series of self-normalized coefficients of cross-correlation between an image sensed by the television camera and a series of reference images retrieved by the tape recorder. Each optical input image is thereby converted into an equivalent feature vector, which can serve as the input to a pattern recognition system. The successive elements of each feature vector naturally depend on the choice of images stored in the recorder tape. Consequently, the overall system provides great flexibility in optical pattern preprocessing operations.

Acknowledgement

References

1. Cutrona, L. J., Leith, E. N., Palermo, C. J., and Porcello, L. J. "Optical Data Processing and Filtering Systems", IRE Transactions on Information Theory, IT-6: 386-400, (1960).

2. Montgomery, W. D. and Broome, P. W. "Spatial Filtering", Journal of the Optical Society of America, 52(11): 1259-1275, (1962).

3. Vander Lugt, A. "Signal Detection By Complex Spatial Filtering", IEEE Transactions on Information Theory, IT-10: 139-145, (1964).

4. Nagy, George. "State of the Art in Pattern Recognition", Proceedings of the IEEE, 56(5): 836-862, (1968).

5. Kanal, Laveen N. Pattern Recognition, Proceedings of the IEEE
    Workshop on Pattern Recognition, held at Dorado, Puerto Rico,
    Washington, D. C., Thompson Book Co., 1968.

6. Cheng, G. C., Ledley, R. S., Pollock, D. K., and Rosenfeld A.
    Pictorial Pattern Recognition, Proceedings of Symposium on Auto-
    matic Photointerpretation sponsored by the Office of Naval Research,
    the University of Maryland, and the Pattern Recognition Society,
    Washington, D. C., Thompson Book Co., 1968.

7. Levine, Martin D. "Feature Extraction: A Survey", Proceedings of
    the IEEE, 57(8): 1391-1407, (1969).

8. Tretiak, Oleh J., Eden, Murray, and Simon, William. "Internal
    Structure From X-Ray Images", 8th ICMBE Pattern Recognition and
    Image Processing, Session 12-1, Palmer House, Chicago, Illinois
    (July 20-25, 1969).

9. Andrews, Harry C. and Pratt, William K. "Transform Image Coding",
    Proceedings of the Symposium on Computer Processing in Communica-
    tions (New York 1969), Polytechnic Press of the Polytechnic Insti-
    tute of Brooklyn, Brooklyn, New York, distributors: John Wiley &
    Sons, Inc., New York, 1970, pp. 63-84.

10. Spencer, David R. "Bit-Plane in Coding of Continuous Tone Pic-
    tures", ibid, pp. 101-120.

11. Lippel, Bernard. "Experiments with a New Message Format for Digi-
    tal Representation of Photographs", ibid, pp. 121-139.

12. Stell, Douglas E., Hess, Donald T., and Clarke, Kenneth K. "Digi-
    tal Storage, Measurement, Transformation, and Recovery of Video
    Information", ibid, pp. 141-150.

13. Deutsch, E. S. "On Form Preprocessing Techniques for Character
    Recognition", ibid, pp. 221-234.

14. Howell, P. N. "Optical-Cross-Correlation Studies of Two-Dimen-
    sional Patterns", Ph.D. Thesis, Oklahoma State University (May
    1969).

15. Morris, D. E. "Computer-Controlled Video System for Image
    Analysis and Pattern Recognition", Ph.D. Thesis, Oklahoma State
    University (July 1970).

LIST OF CAPTIONS FOR PHOTOGRAPHS

Figure   1.   Scheme of Video Cross-Correlating System

Figure   2.   Scheme of Special Sync Unit

Figure   3a.   Vertical Sync Inverter Circuit

Figure   3b.   Horizontal Sync Delay Circuit

Figure   4a.   Sync-Clipping Video Amplifier

Figure   4b.   Video Amplifier Input and Output

Figure   5.   Cross-Correlator Pre-conditioner

Figure   6.   Scheme of Clamping-Pulse Generator
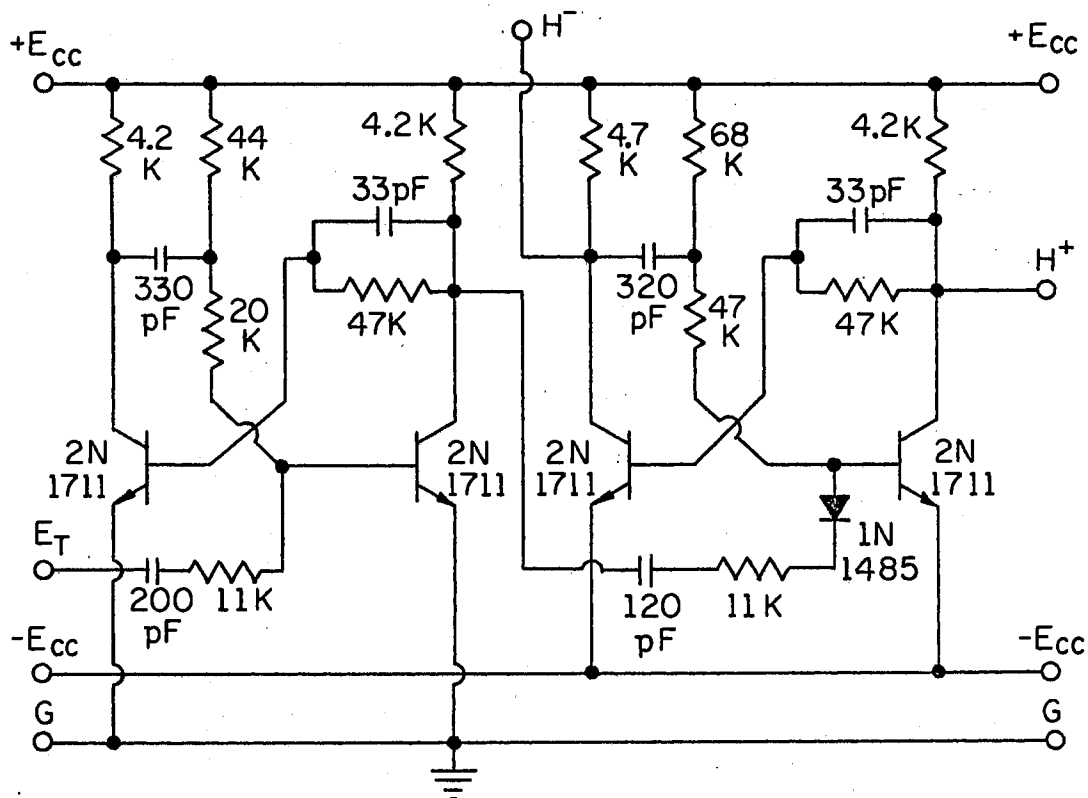
Figure   7a.   Vertical Clamping-Pulse Generator Circuit

Figure   7b.   Horizontal Clamping-Pulse Generator Circuit

Figure   8.   Video Cross-Correlator Circuit

Figure   9.   Photograph of Video Cross-Correlating System

Figure   10.   (See Attached Scheme)

Figure   11.   (See Attached Scheme)

Figure   12.   Correlator Output Voltages for Various Camera and
             Recorder Image Patterns
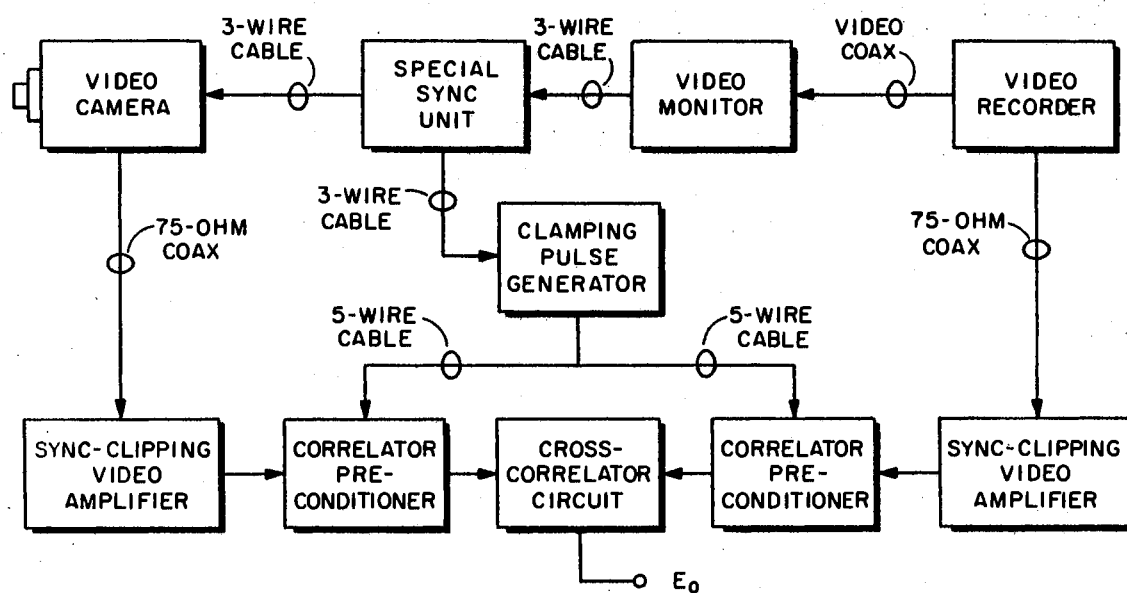
Figure 1.   Video Cross-Correlating System
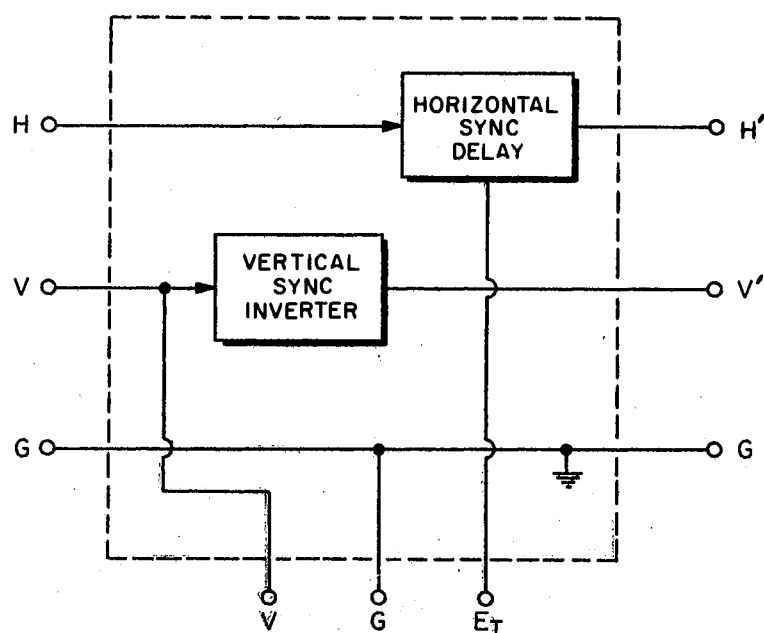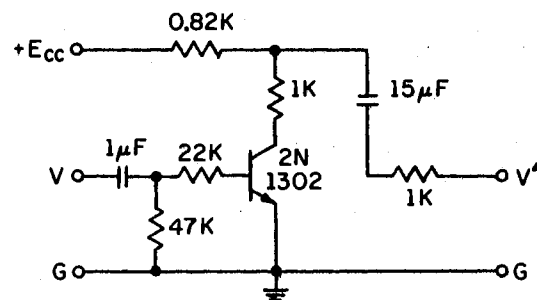
Figure 2.   Scheme of Special Sync Unit

Figure 3a. Vertical Sync Inverter Circuit
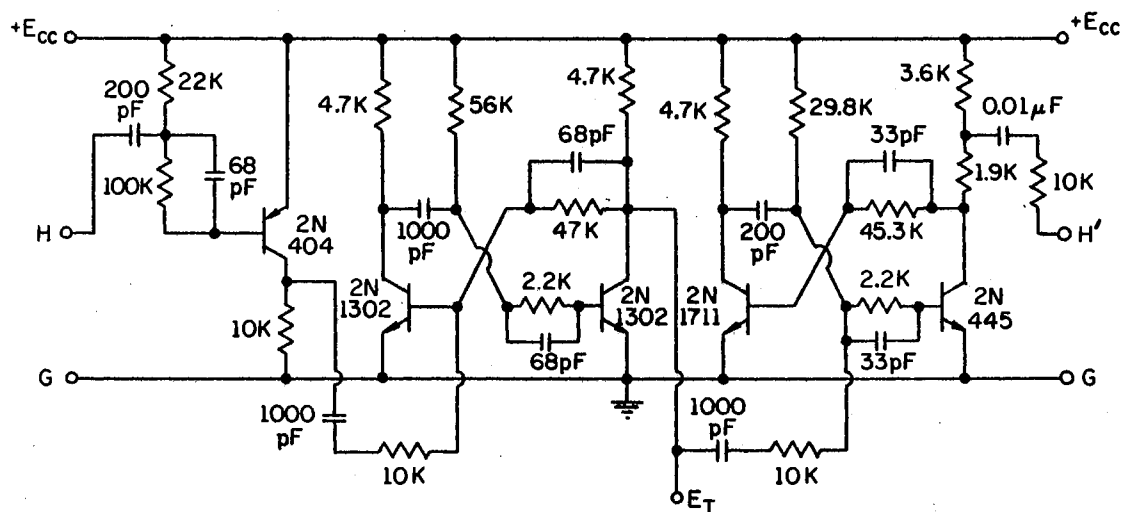
Figure 3b. Horizontal Sync Delay Circuit
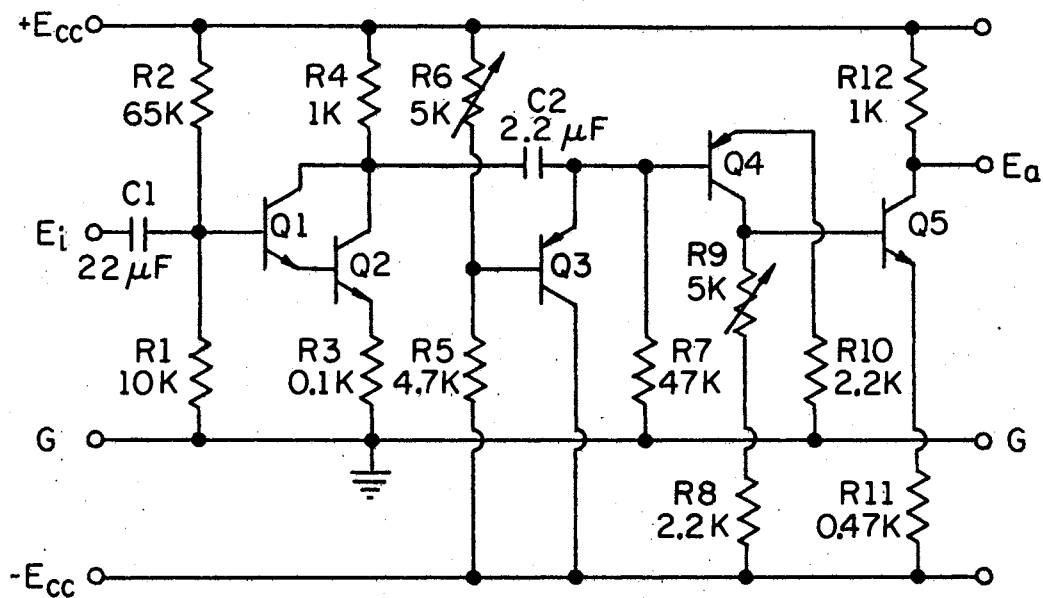
Figure 4a. Sync-Clipping Video Amplifier
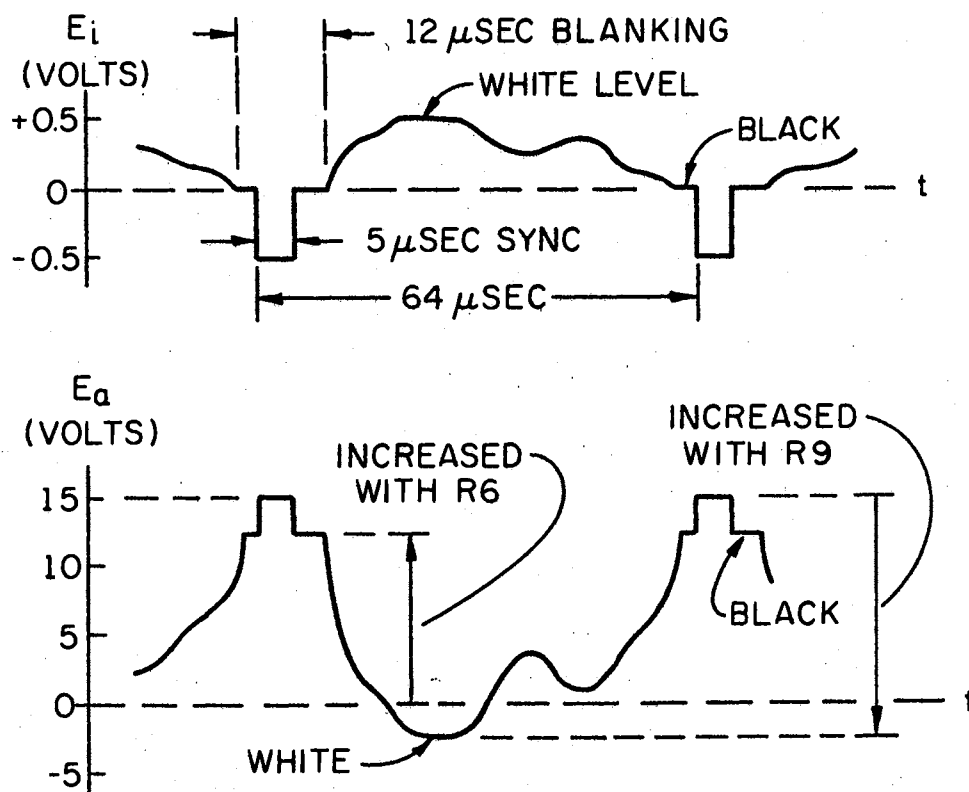
Figure 4b. Video Amplifier Input and Output
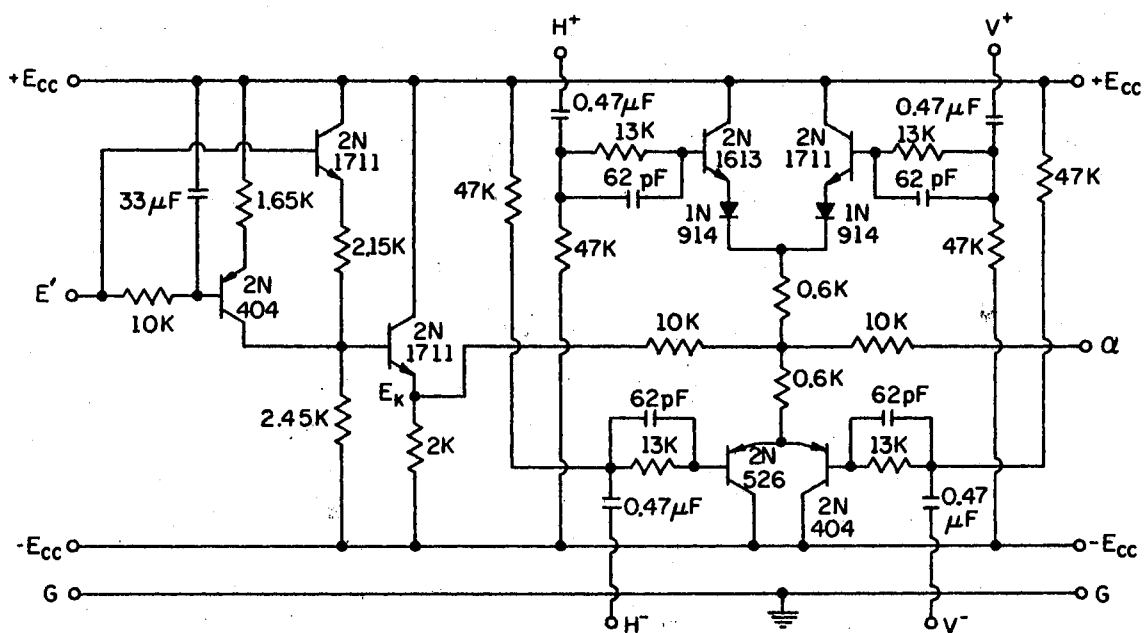
Figure 5.   Correlator Pre-conditioner
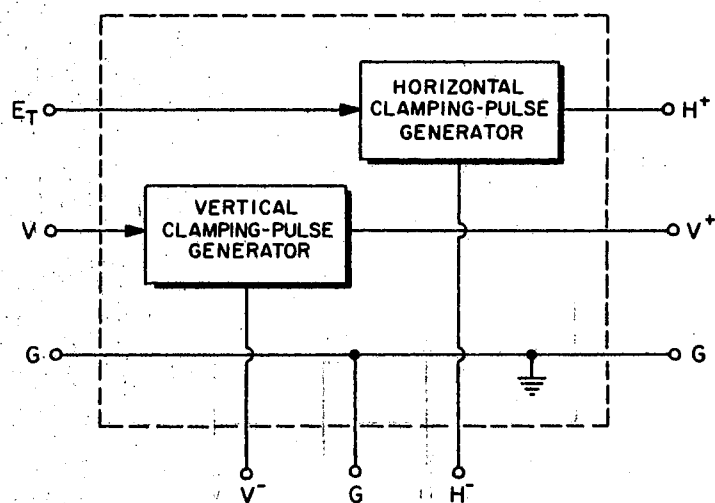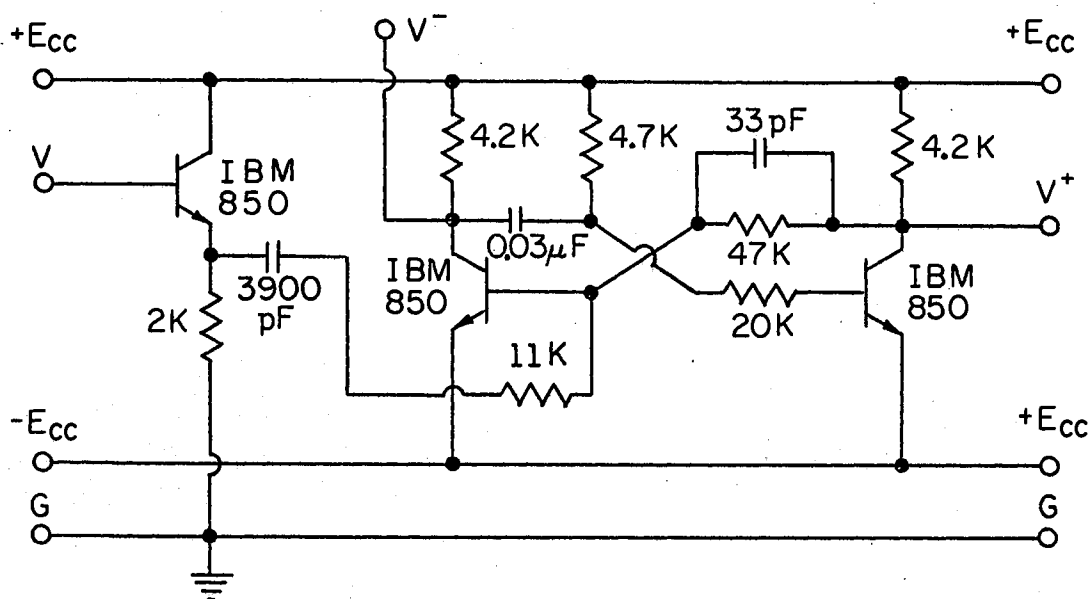
Figure 6.   Scheme of Clamping-Pulse Generator

Figure 7a. Vertical Clamping-Pulse Generator Circuit

Figure 7b. Horizontal Clamping-Pulse Generator Circuit
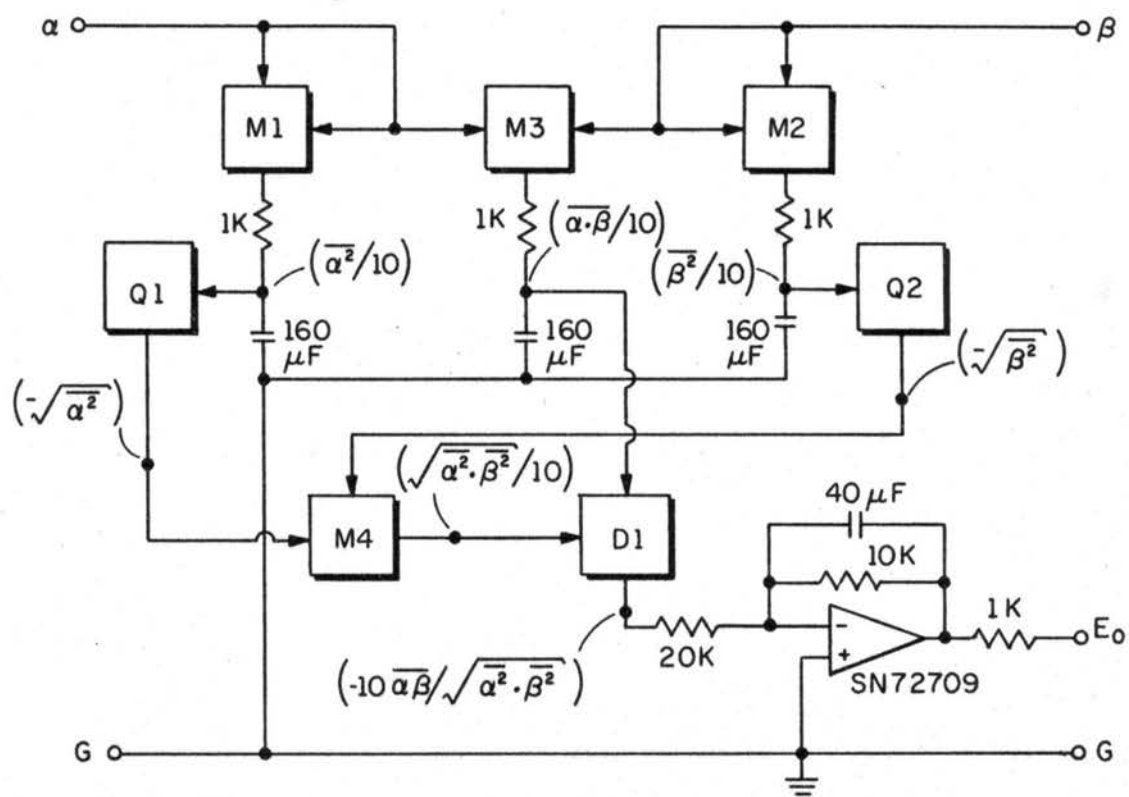
Figure 8. Cross-Correlator Circuit
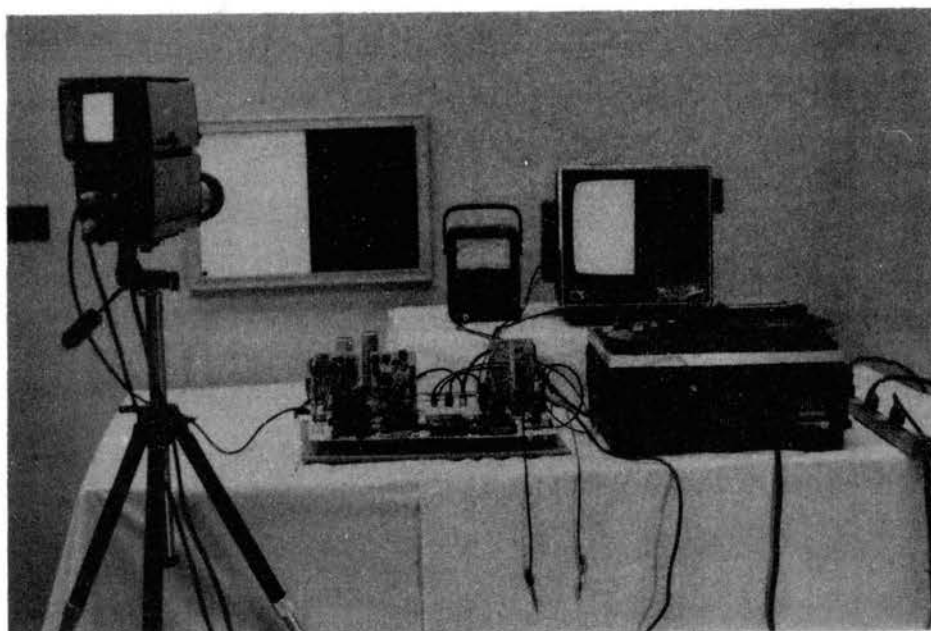
Figure 9. Photograph of Video Cross-Correlating System
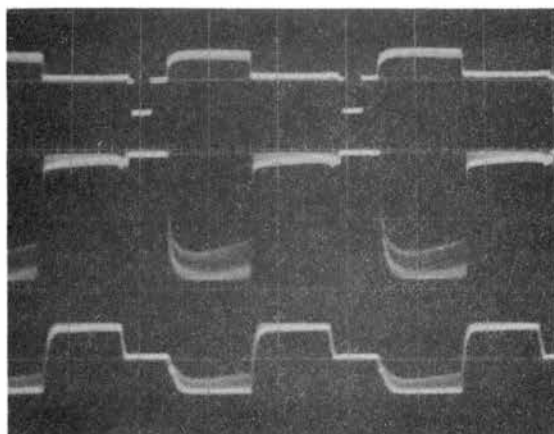
Figure 10a

Camera image is white left,
black right. Lower trace is α.
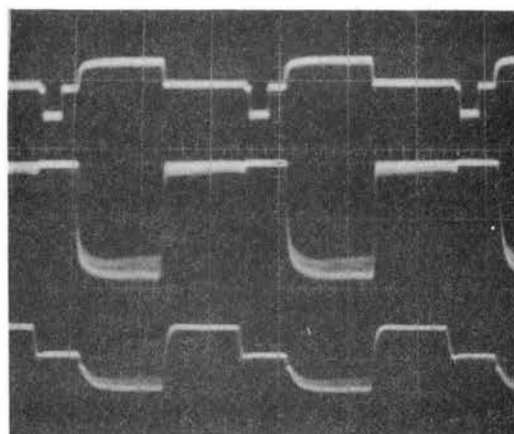
Figure 10b

Playback image is white left,
black right. Lower trace is β.
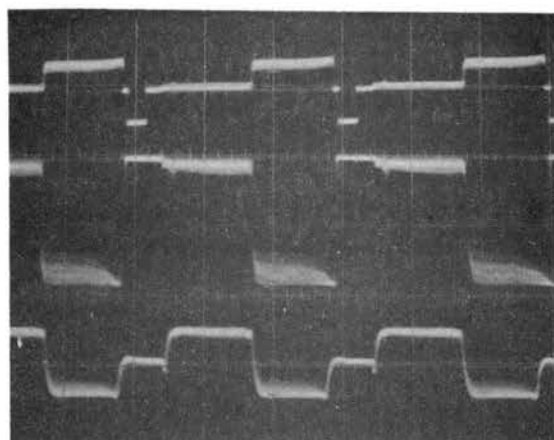
Figure 10c

Camera image is black left,
white right. Lower trace is α.
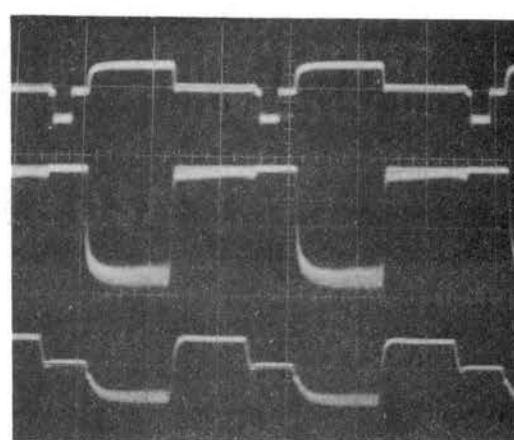
Figure 10d

Playback image is white left,
black right. Lower trace is β

Upper Traces: Input to clipping amplifier,   1.0 V/cm
Center Traces: Input to pre-conditioner   ,   10 V/cm
Lower Traces: Input to cross-correlator   ,   10 V/cm

Horizontal Scale: Time, 20 µsec/cm

Figure 10.   Genesis of Left and Right Inputs to Correlator,
for Identical and Opposite Horizontal-Step Images
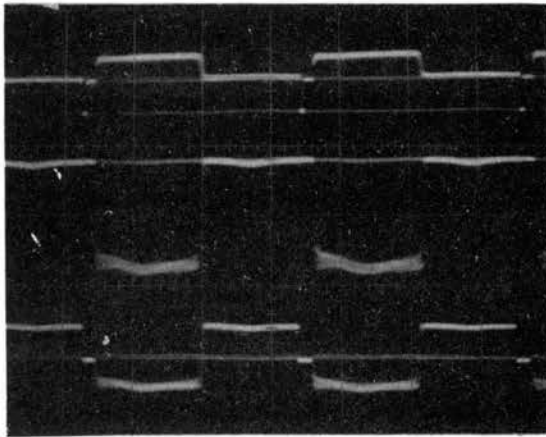
Figure 11a

Camera image is white above,
black below. Lower trace is α.
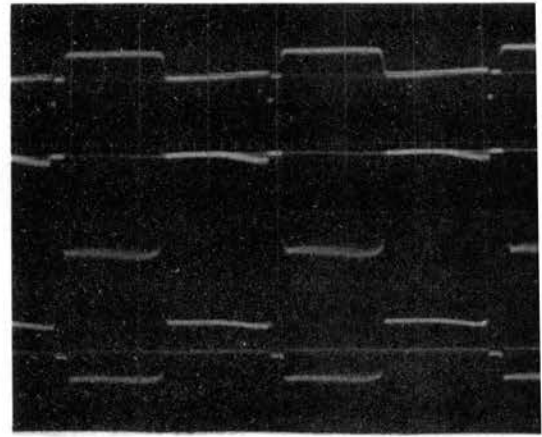
Figure 11b

Playback image is white above,
black below. Lower trace is β.
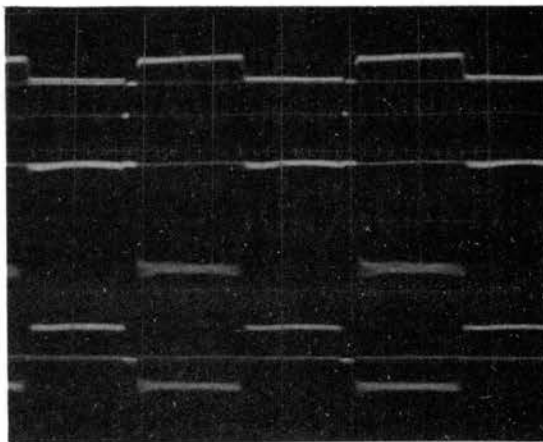
Figure 11c

Camera image is white above,
black below. Lower trace is α.
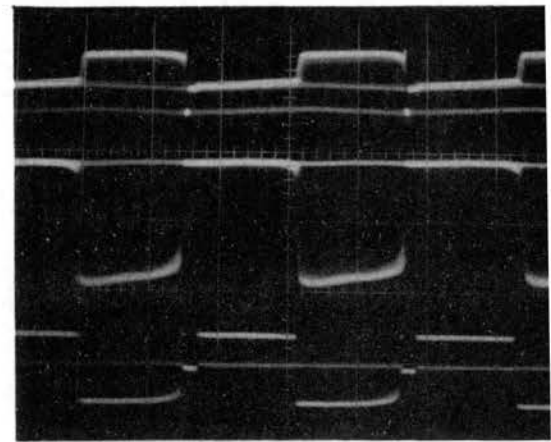
Figure 11d

Playback image is black above,
white below. Lower trace is β

Upper Traces: Input to clipping amplifier,   1.0 V/cm
Center Traces: Input to pre-conditioner   ,   10 V/cm
Lower Traces: Input to cross-correlator   ,   10 V/cm

Figure 11.   Genesis of Left and Right Inputs to Correlator,
for Identical and Opposite Vertical-Step Images

IMAGE INPUT FROM CAMERA

4.7	2.6	0.1	-2.6	-4.6	-2.1	0.4	2.9
2.5	4.9	2.9	0.3	-2.3	-4.6	-2.5	0.1
0.2	2.9	4.9	2.6	0.2	-2.5	-4.5	-2.3
-2.4	0.3	3.0	4.9	2.9	0.2	-2.4	-4.5
-4.6	-2.4	0.3	3.0	5.0	2.8	0.4	-2.2
-1.8	-4.5	-2.7	-0.1	2.6	5.0	3.3	0.9
0.4	-2.7	-4.6	-2.2	0.4	3.3	4.9	3.0
2.8	-0.2	-2.7	-4.5	-2.2	0.7	3.1	4.8

IMAGE RETRIEVED FROM RECORDER

Figure 12. Correlator Output Voltages for Various Camera and Recorder Image Patterns

APPENDIX B

VALIDITY OF THE IMAGE RECONSTRUCTION ALGORITHM

This appendix contains a theoretical demonstration of the validity of the image reconstruction algorithm which was described in Chapter V. Only a special case is considered.

Theoretical Proof for a Special Case

It is recalled that the image reconstruction process was expressed by Equation (5.4) which is repeatedly written as

$$[B^*] = TF\{ \sum_{i=1}^{k} a_i u_i |u_i|^{(p_i-1)} [R_i] \} \qquad . \qquad (B.1)$$

In particular, it is assumed that the images concerned are of 4-by-4 resolution, and that the set of zero-mean orthogonal images, including the all-white image, shown in Figure 34 of Chapter V is used as the set of reference images. It is also assumed that $a_i = 1$, and $p_i = 1$, for $i = 1, 2, \ldots, 16$. Therefore, Equation (B.1) is simplified to be

$$[B^*] = TF\{ \sum_{i=1}^{16} u_i [R_i] \} \qquad . \qquad (B.2)$$

If the information matrix (after subtracting out the DC average value) of a certain test image B is expressed by the equation

$$[B] = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 \\ b_5 & b_6 & b_7 & b_8 \\ b_9 & b_{10} & b_{11} & b_{12} \\ b_{13} & b_{14} & b_{15} & b_{16} \end{bmatrix} \qquad (B.3)$$

and if the information matrices of the reference images are expressed as

$$[R_i] = \begin{bmatrix} \gamma_{i,1} & \gamma_{i,2} & \gamma_{i,3} & \gamma_{i,4} \\ \gamma_{i,5} & \gamma_{i,6} & \gamma_{i,7} & \gamma_{i,8} \\ \gamma_{i,9} & \gamma_{i,10} & \gamma_{i,11} & \gamma_{i,12} \\ \gamma_{i,13} & \gamma_{i,14} & \gamma_{i,15} & \gamma_{i,16} \end{bmatrix} \qquad (B.4)$$

in which $i = 1,2,\ldots,16$, then the corresponding cross-correlation coefficients are obtained as expressed by

$$u_i = c_i \sum_{j=1}^{16} b_j \gamma_{i,j} \qquad (B.5)$$

where c_i is a constant determined by the mean of the corresponding reference image. Since the reference images are of zero-mean (except the all-white image), it is supposed that $c_i = c$, for $i = 1,2,\ldots,16$. Substituting Equation (B.5) into Equation (B.2), it is seen that

$$[B^*] = TF\{c \sum_{i=1}^{16} \sum_{j=1}^{16} b_j \gamma_{i,j} [R_i]\} \qquad . \qquad (B.6)$$

Accordingly, the elements of the matrix $[B^*]$ will be

$$b_m^* = TF\{c \sum_{j=1}^{16} b_j \sum_{i=1}^{16} \gamma_{i,m} \gamma_{i,j}\} \quad , \quad (B.7)$$

in which $m = 1,2,\ldots,16$. However, further inspection on the information matrices shown in Table III of Chapter V shows that

$$\sum_{i=1}^{16} \gamma_{i,m} \gamma_{i,j} = 16, \text{ if } j = m$$

$$= 0, \text{ if } j \neq m \quad . \quad (B.8)$$

Thus,

$$b_m^* = TF\{c \cdot 16 \cdot b_m\} \quad (B.9)$$

where $m = 1,2,\ldots,16$. Consequently, the matrix $[B^*]$ is linearly related to the matrix $[B]$.

VITA

Wen-Shyong Su

Candidate for the Degree of

Doctor of Philosophy

Thesis: COMPUTERIZED VIDEO IMAGE REDUCTION STUDIES

Major Field: Electrical Engineering

Biographical:

Personal Data: Born in Taichung, Taiwan, June 24, 1943, the son of Wei-Sir and Gay L, Su.

Education: Graduated from Taiwan Provincial Taichung First High School in June, 1959; received the Diploma in Electrical Engineering from Taiwan Provincial Taipei Institute of Technology in June, 1964; received the Master of Science degree with a major in Electrical Engineering from Wichita State University in June, 1968; completed the requirements for the Doctor of Philosophy degree at Oklahoma State University in July, 1971,

Professional Experience: Student Assistant for the Department of Electrical Engineering, Wichita State University, from September, 1967 to May, 1968; Graduate Assistant for the School of Electrical Engineering, Oklahoma State University from June, 1968 to August, 1969; Research Assistant for the Engineering Research, Oklahoma State University from September, 1969 to June, 1971.

Professional Organizations: Member of the Institute of Electrical and Electronics Engineers.