VIDEO RESOLUTION, FRAME RATE AND GRAYSCALE TRADEOFFS UNDER LIMITED BANDWIDTH FOR UNDERSEA TELEOPERATION

by

VIVÉK RANADIVÉ

S.B. in Elec. Eng. & Computer Science, M.I.T.
1979

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1979
(i.e. February 1980)

Signature of Author.....
Department of Mechanical Engineering, Sept. 15, 1979

Certified by.....
Thesis Supervisor

Accepted by.....
Chairman, Departmental Committee of Graduate Students

VIDEO RESOLUTION, FRAME RATE AND GRAYSCALE
TRADEOFFS UNDER LIMITED BANDWIDTH FOR UNDERSEA TELEOPERATION

by

VIVÉK RANADIVÉ

Submitted to the Department of Mechanical Engineering
on September 26, 1979, in partial fulfillment of the
requirements for the degree of Master of Science.

ABSTRACT

The high costs associated with a diver working at depths approaching 1000m make the use of remotely controlled visual inspection and manipulation look very attractive. When the undersea teleoperator is untethered, and acoustic communication is used, there are severe restrictions on the bandwidth of the image transmission channel. Since the product of the Frame Rate (F), Resolution (R) and Grayscale (G) is directly proportional to the transmission baud rate, for a fixed channel capacity there are tradeoffs between F, R and G in the actual sampling of the picture. A manipulator was used in the MASTER/SLAVE mode to study these tradeoffs.

A mathematical model for image reconstruction was used to design a system which would allow the operator to set independently the F, R and G of the picture. This design was implemented using a commercially available charge coupled camera with its accompanying digitizer, and an electronic unit was constructed to interface with the digitizer and camera.

When subjects first saw the picture with which they had to perform tasks using the manipulator, they refused to believe that they could actually work under those circumstances. Much to their surprise, however, not only were they able to work with the digitized and itermittant picture but were also able to perform with a considerably degraded picture. It was found that frame rate, resolution and grayscale could be independently reduced without preventing the operator from accomplishing his/her task. Threshold points were found beyond which degradation would prevent any successful performance. In the range of operation* it was observed that frame rate and grayscale could be degraded considerably more than resolution, before operation became impossible.

---

*F = 28 f/s, R = 128x128 pixels, g = 4 bits grayscale.

Isoperformance curves (curves of constant performance) were found for two subjects for various combinations of frame rate, resolution, and grayscale. These results were represented along with isotransmission curves (curves along which the transmission rate is the same) and the two considered in light of each other.

It was the conclusion of the the author that a well trained operator could perform tasks familiar to him/her with a considerably degraded picture. Furthermore, for a given task and transmission rate, some FRG combinations are more effective than others. For the particular tasks experimented with, 50,000 bits/sec was found to be a satisfactory transmission rate.

Thesis Supervisor: Thomas B. Sheridan
Title: Professor of Engineering and Applied Psychology

ACKNOWLEDGEMENTS

Professor Sheridan, thank you for your help, guidance and encouragement. Al Woodin, your help with the hardware was invaluable and Professor Moray, your time and advice was greatly appreciated. Leslie Regan - thank you for typing this manuscript with such speed and accuracy.

Finally, I must thank the office of Naval Research for making this study possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION TO TELEOPERATED VEHICLES

Undersea operations are not only increasing in complexity but also the depths are getting to the point where diver safety is being compromised. Operations including exploration, inspection, construction, maintenance, salvage and rescue [1] are being carried out at depths greater than 300m, where the risks and dollar costs associated with using divers makes it necessary to find an alternative method. Thus, the primary reasons for research in the use of undersea teleoperators are economic and related.

A number of teleoperators have been built by government and commercial organizations. In fact, in view of the numerous design option available, these are many control/communication alternatives for an unmanned submersible. Figure 1.1 is an illustration of these various alternatives. The communication alternatives for unmanned vehicles will be important in determining the trade-off between human and computer control. The particular configuration will, of course, depend on the task to be accomplished, operating depth, size, speed, power source, duration, etc. A robotic teleoperator alone with no remote control link would be the most sophisticated mode of communication. This would, however, require significant progress in such fields as artificial intelligence before becoming feasible —— a level of technology that is not currently available.

Sonic link control is a more realistic alternative it falls between the state-of-the-art and technology in the near future. Present sonic links give channel capacities as high at 10,000 bits per second: enough to give a video picture at slow speeds. The main advantage of using a sonic link is that it eliminates the need for a cumbersome tether.

Thus, remotely controlled underwater vehicles [2] offer significant advantages over manned submersibles for missions that require the submersible to perform complex manipulative tasks in close proximity to such entangling obstructions as cables, nets, or structures. These vehicles have almost unlimited operational time at the bottom (if powered over a tether) and offer no risk to their operators. Similar considerations apply when the operations involve ordnance or explosives whose accidental detonation could cause the loss of a submersible and its crew.

Work systems developed in the past have proven to be very reliable ocean engineering tools, and were of inestimable value to the Navy during the H-bomb recovery operation at Palomares, Spain, the ocean engineering operation in the Azores, and the Pisces III rescue off Ireland.

Continuing research is directed at improving the operational capabilities of the submersible while decreasing the operational costs and shipboard support.

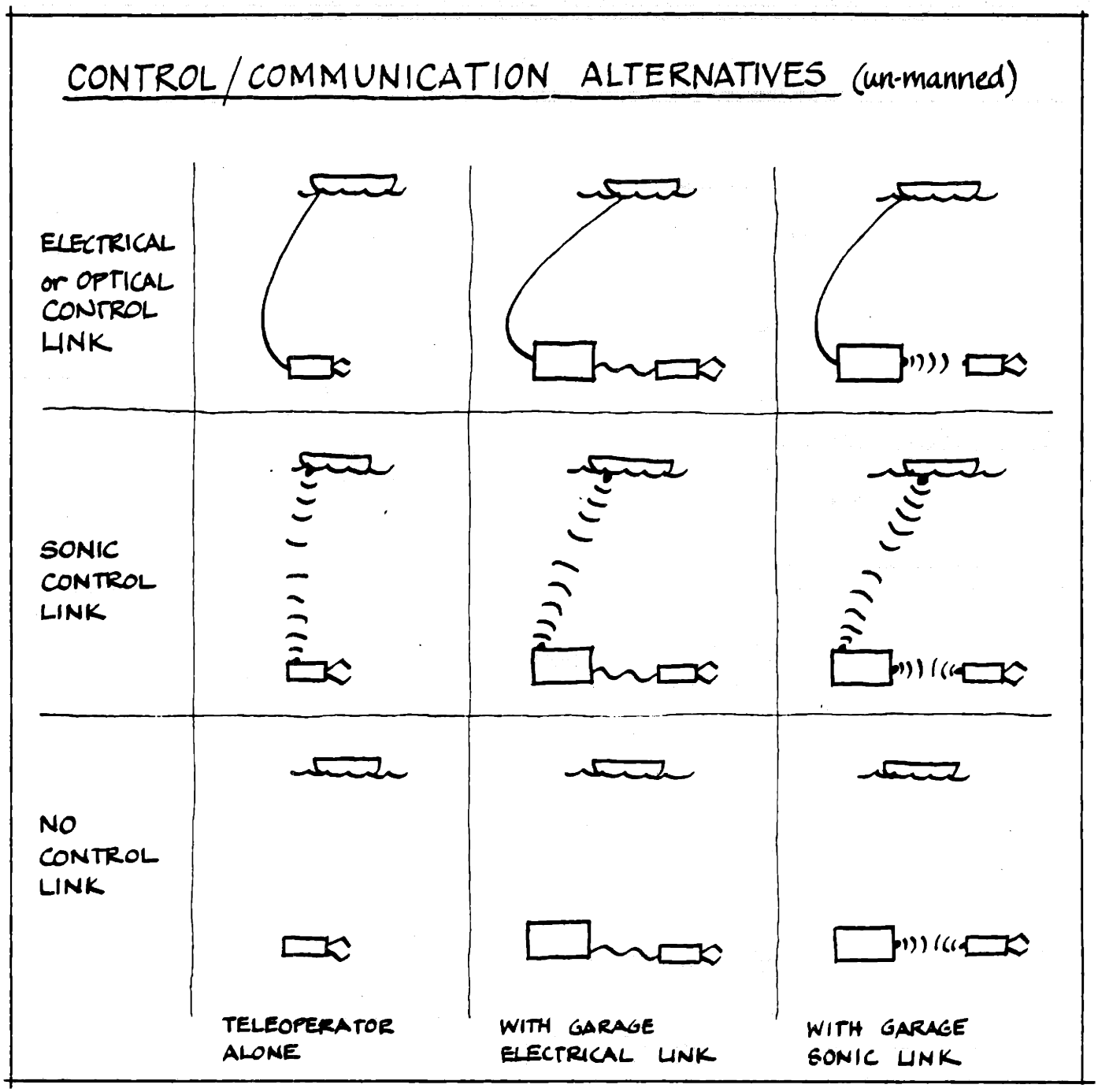


FIGURE 1.1: COMMUNICATION ALTERNATIVES for unmanned vehicles will be important in determining the trade-off between human and computer control. The particular configuration will, of course, depend on task to be accomplished, operating depth, size, speed, power source, duration, etc. The above matrix classifies alternative forms of communication: 1) with the surface ship (if any); 2) with an intermediary "garage" (if any).

CHAPTER II

IMAGE TRANSMISSION AND THE UNDERSEA TELEOPERATOR

There are various sensing modes available for undersea teleoperation. Unfortunately, formidable environmental problems [2] challenge progress in the underwater imaging field. Acoustic energy, capable of traveling great distances underwater, has been ill-suited for imaging purposes because of inherently low resolution. Light, the traditional image-forming medium, is severely attenuated in seawater. Water selectively absorbs light and the absorbed component no longer contributes to visual images. Light is also removed from the image-forming process by scattering. Scattering further reduces image contrast. When objects are illuminated by artificial sources, light is often reflected, or back scattered, creating an effect similar to that encountered when automobile headlights are used in a fog. In addition to the difficulty of transmitting light in the sea, the high pressures and corrosive action of seawater pose challenges to hardware design.

Inspite of all the drawbacks associated with optical imaging, it continues to be the most important sensing mode. Video is farily well developed and easily interpretable by vision, which is the most valuable sensory mode of the human organism.

The video information used by remotely operated data systems to perform tasks underwater has to be real time in nature. The video

signal must provide sufficient resolution, grayscale and framerate to allow operation of the system at its full potential. Conventional television fulfills this task, but places a severe demand on the communication link (cable, rf, or acoustic). Standard television requires approximately 5MHz of bandwidth using most of the available frequency spectrum in such systems. Let us now consider a sonic system.

Consider a teleoperator with a camera mounted on it. The bottleneck in this system is the information transmission capability from teleoperator to ship (or supervisor). Sonic communication has two sources of difficulty [1], time-delay and slow frame-rate. Round trip time-delay is the time for a command to travel to the vehicle and the first indication of response to travel back. This will at the very minimum be twice the distance divided by the speed of propagation. Limited channel capacity is the constraint that would cause a slow frame-rate.

A transmission channel of K bits/second could have

$\ell \times \ell$ pixels/frame

b bits of grayscale/frame

f frames/sec

Thus, information transmission rate is

$$(\ell \times \ell)\ \frac{\text{pix}}{\text{frame}} \times\ b\ \frac{\text{bits}}{\text{pixel}} \times\ f\ \frac{\text{frames}}{\text{sec}} = K\ \frac{\text{bits}}{\text{sec}}$$

Normal broadcast television has

$$\ell \times \ell \simeq 512 \times 512 \text{ pixels/frame}$$

assume $b = 4$ bits of gray

$f = 30$ frames/sec

This means that normal broadcast television requires upwards of 30,000,000 bits/second. Current acoustic technology makes it possible to transmit from 30,000 bits/second in the shorter ranges to 3,000 bits/second for longer distances. It is, therefore, imperative for operators to learn to perform with coarse, slow pictures. Given the restricted bandwidth of operation, compromises between the various features contributing to making up picture quality will exist.

At this point it is necessary to distinguish between image transmission and picture display. Image transmission is the actual transmission of bits (representing the picture) from teleoperator to the location of display. The variables in this process are the frame-rate, resolution and grayscale. The display process on the other hand, deals with the reconstruction of the image from the bits received through the transmission process. This reconstruction procedure can be manipulated so as to get a better picture. The enhancement techniques which result in the better picture have been known as digital picture processing methods.

Digital picture processing deals with the reconstruction of images from quantized samples [5]. Image processing methods can also be used for optimum sampling.

Pictures are continuous functions [5] since the gray level is a continuous function of position. Transmission or computer processing requires the quantization of the image. Thus quantizations consists of sampling the gray level in the picture at an MxN array of points. The gray level needs to be digitized and this is done by dividing the range of gray level into L intervals such that the gray level at any particular point can take on only one value. To get a good picture from this quantization, M, N and L have to be large. Ordinarily, the finer the sampling and quantization, the better the reproduced picture. Nothing is gained, however, by increasing M, N and L beyond the optical and gray scale resolution capabilities of either the receiver or the human observer.

The camera in the system could typically be a charge coupled device which outputs an electric charge proportional to the light intensity. The function of this solid state device, then, is to translate an optimal image [3] into a precise electrical signal. The camera has a microsensor which consists of an array of pixels (picture elements) each consisting of a pair of capacitors arranged to "share" charge through a p-coupled region. Within each pixel, depletion regions are caused by the application of row and column bias voltages to the capacitor plates, and are used to trap the positive charges (holes) generated by light striking the array.

An automation interface [4] takes the output of the charge coupled device and digitizes the analog information. This information is then transmitted via some channel to a computer - say on board a ship as in Figure 2.1.

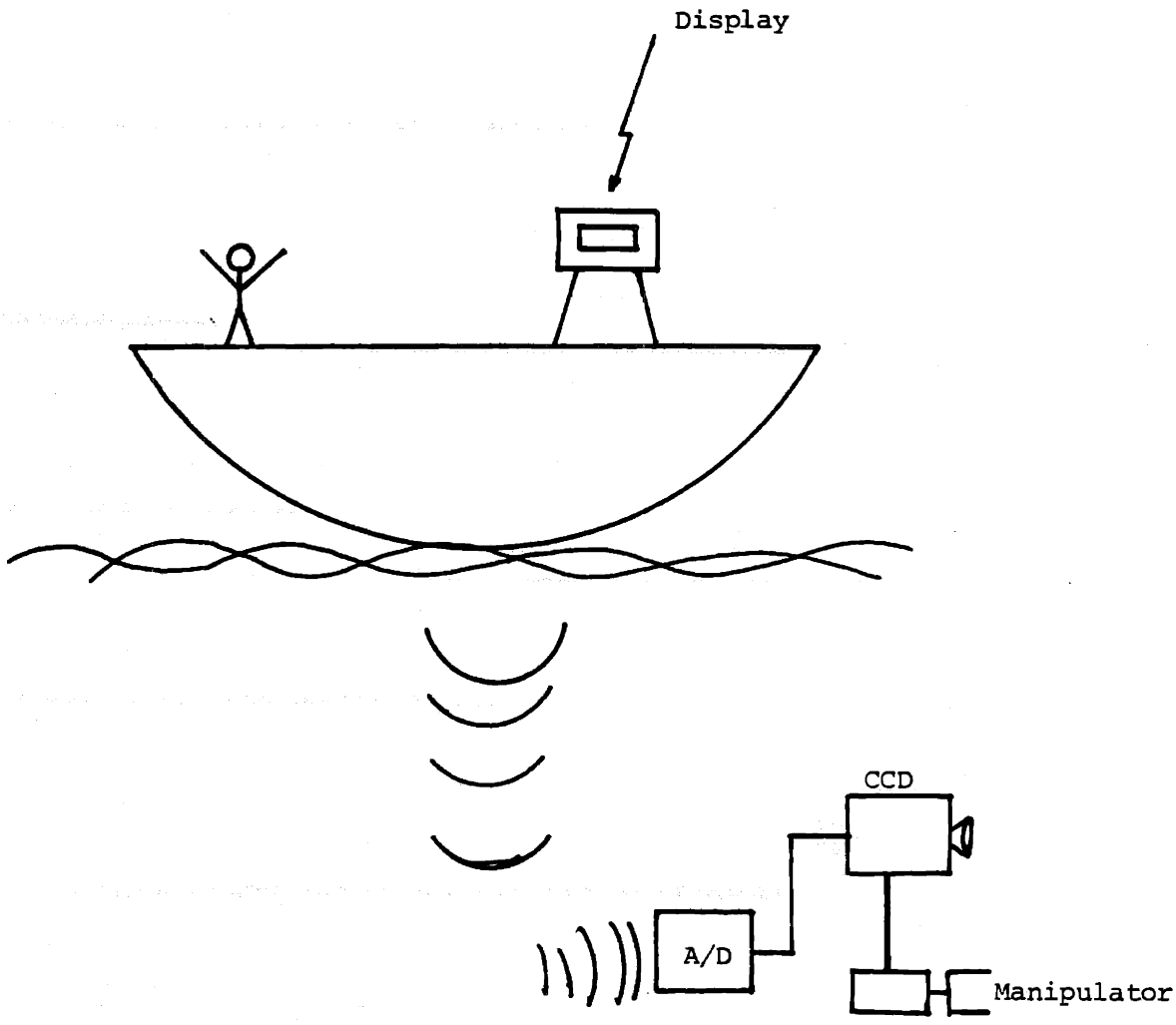


FIGURE 2.1: TELEOPERATOR SYSTEM

CHAPTER III

IMAGE DIGITIZATION

As mentioned in Chapter II, picture quantization is the process of sampling the gray level or a picture at an MxN array of points.

In a more general sense [4], the aim of sampling is to represent a continuous picture by a finite string or array of numbers called samples. The primary consideration on this data is that it should be possible to reconstruct a picture from it.

There are numerous methods of picture sampling. This chapter deals with two specific methods:

i) sampling using an array of points

ii) sampling using orthonormal functions

3.1 Sampling Using An Array of Points

In this technique the samples are the gray levels of the picture at an array of points. The method will first be established [4] for one-dimensional functions:

3.1.1 One-Dimensional Case

Consider a one-dimensional function f(t) is represented by samples f(kT), where k takes on integer values from $-\infty$ to $+\infty$ and T is the sampling period. Then, following Peterson & Middleton [5] use

an interpolation function g(t) such that:

$$f(t) = \sum_{k=-\infty}^{\infty} f(kT)\, g(t - kT) \tag{1}$$

In other words, the contribution of f(kT) to the reconstruction at time t is weighted by the factor g(t - kT). Assume that f and g are Fourier transformable.

Using the fact that

$$\iint_{-\infty}^{\infty} g(x,y)\,\delta(x - \alpha, y-\beta)\,dxdy = g(\alpha,\beta) \tag{1.1}$$

$$f(kT)g(t - kT) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)\delta(\tau - kT)d \tag{2}$$

substitute in (1)

$$f(t) = \sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} f(\tau)g(t - \tau)\ (\tau - kT)d \tag{3}$$

Inverting the order of the summation and integration

$$f(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)\left(\sum \delta(\tau - kT)\right)d\tau \tag{4}$$

Consider $\sum_{k=-\infty}^{\infty} \delta(\tau - kT)$

this is periodic and can therefore be represented by a Fourier series:

$$\sum_{k=-\infty}^{\infty} \delta(\tau - kT) = \sum_{n=-\infty}^{\infty} a_n\, e^{\left(j\,\frac{2\pi n}{T}\,\tau\right)} \tag{5}$$

where coefficients a_n are

$$a_n = \frac{1}{T} \int_{-\pi 2}^{T/2} \left(\sum_{k=-\infty}^{\infty} \delta(\tau - kT) \quad e^{(-j \frac{2\pi n}{T} \tau} \right) d\tau$$

$$= \frac{1}{T} \int_{-\pi/2}^{\pi/2} \delta(\tau) e^{(-j \frac{2\pi n}{T} \tau)} \, d\tau$$

Since only the k = 0 term of the sum is non-zero in the range of integration then using 1.1

$$a_n = 1/T \text{ for all } n \tag{6}$$

Substitute (5) and (6) in (4)

$$f(\tau) = \int_{-\infty}^{\infty} f(\tau)\, g(t - \tau) \left(\sum_{n=-\infty}^{\infty} \frac{1}{T} \quad e^{j \frac{2\pi n}{T} \tau} \right)$$

$$= \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \left[f(t)\, e^{j \frac{2\pi n}{T} \tau} \right] \quad \left(\frac{g(t - \tau)}{T} \right) d\tau \tag{7}$$

Realize now that each term in the above summation is a convolution of the functions

$$f(t)\, e^{(j \frac{2\pi nt}{T})} \text{ and } \frac{g(t)}{T} \tag{8}$$

Consider the convolution property of Fourier Transforms:

$$F\left\{\int_{-\infty}^{\infty}\int_{\infty}^{\infty} f_1(\alpha,\beta) f_2(x-\alpha, y-\beta)\, d\alpha\, d\beta\right\} = F\left\{f_1(x,y)\right\} \left\{f_2(x,y)\right\}$$

Using this property

$$F(\omega) = \frac{q(\omega)}{T} \sum_{h=-\infty}^{\infty} F\left(\omega - \frac{2\pi n}{T}\right) \tag{9}$$

Since steps (1) through (7) are reversible, (9) is a necessary and significant reconstructability of f(t) from its samples f(kT) using (1). Sufficient conditions for 9 and 7 will now be formulated for reconstructability to hold.

Consider Figures 3.1a,b,c. If f(t) is bandlimited such that

$$F(\omega) = 0 \quad \text{for } |\omega| \geq 2\pi f_g$$

then the shifted copies of F(ω) are far enough apart that no two of them can have non-zero values at any point, as in Figure 3.1b. For such can F, take

$$G(\omega) = \begin{cases} T & |\omega| < 2\pi f_c \\ 0 & \text{otherwise} \end{cases}$$

then (9) is satisfied since on the right-hand side the unshifted copy of F(ω) has weight 1, while all the other copies have weight 0.

Now $g(t) = \frac{1}{2\pi} \int_{-2_h f_c}^{2_h f_{cc}} T e^{j\omega t}\, d\omega = \frac{\sin 2\pi f_c t}{\pi t/T}$

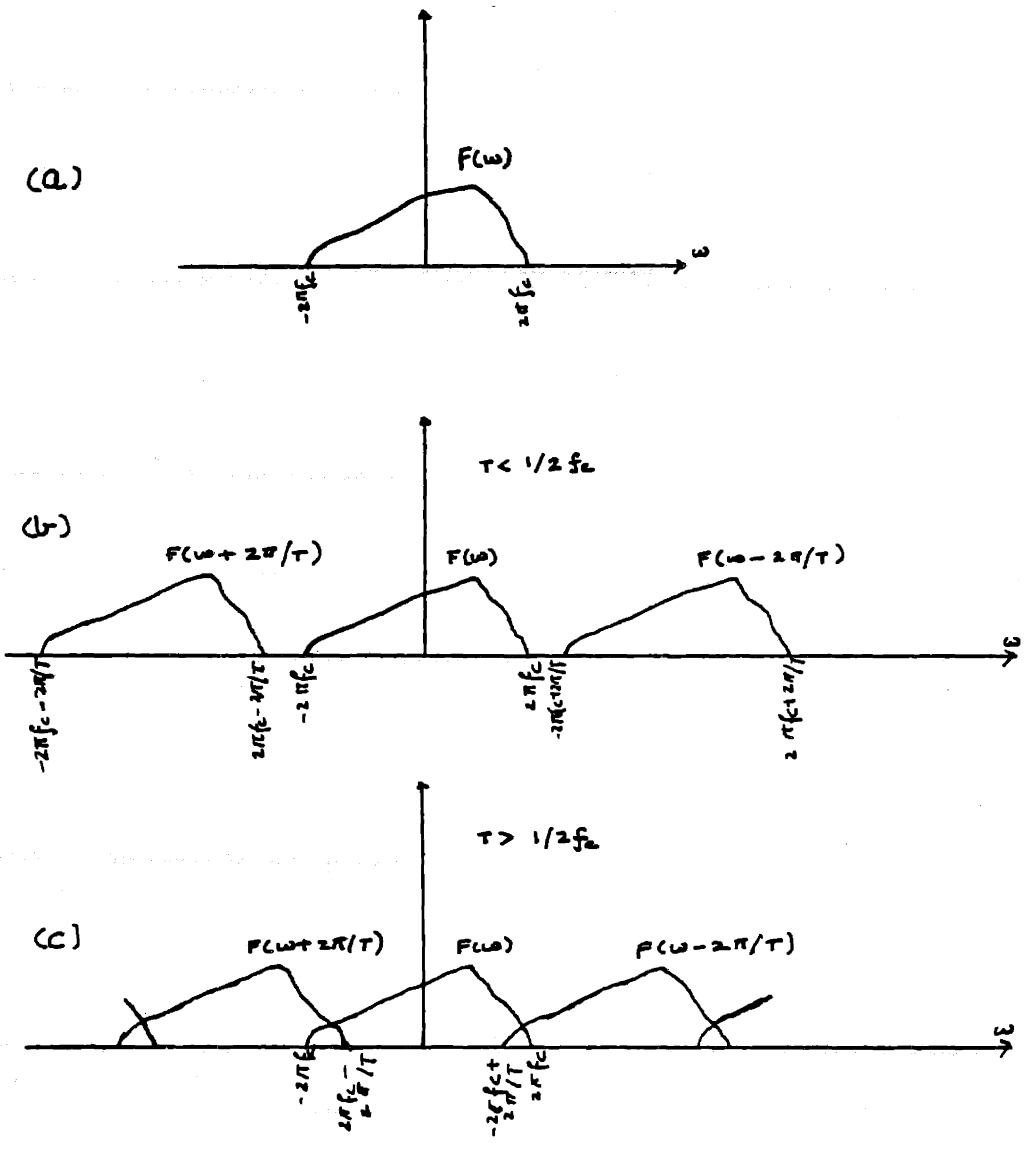


FIGURE 3.1: FOURIER SAMPLES

3.1

Thus, if $T = 1/2\ f_c$,

$$g(t) = \sin (2\pi f_c t)$$

If overlap of the copies is allowed + > $1/2\ p_c$ as in Figure 3.1c, and f is still permitted to be any function limited in frequency by $2\pi f_c$, no g(ω) will satisfy for all possible F. However, if $T < 1/2\ f_c$, we have some flexibility in defining 9.

This discussion [4] is summarized as the Whitaker-Kotelnikov-shannon theorem.

THEOREM:

If the Fourier transform of a function f(t) vanishes for $|\omega| \geq 2\pi f_c$, the f(t) can be exactly reconstructed from samples of its values taken $1/2f_c$ apart or closer.

3.1.2 Extension to Two-Dimensions

The theorem could be extended to the two-dimensional case by using sampling latices and reciprocal lattices [4]. The following theorem is stated here without proof but applying lattice methods to the previous subsection would yield the same.

THEOREM:

A function $f(\vec{r})$ whose Fourier transform $F(\vec{\omega})$ vanishes over all but a bounded region of spatial frequency space can be everywhere reproduced from its values taken over a lattice of points $(m\vec{r}, + n\vec{r}_2)$, m,n = D,±1, ±2..., provided vectors $\vec{r}_1$ and $\vec{r}_2$ are small enough to ensure nonover-

lapping of the spectrum $F(\bar{\omega})$ with its images on a periodic lattice of points $(p\bar{\omega}_1 + q\bar{\omega}_2)$, $p,q = 0, \pm 1, \pm 2, \ldots$, with $\bar{r}_i \cdot \bar{\omega}_i = 0$ for $i \neq j$, and $\bar{r}_i \cdot \bar{W}_j = 1$ for $i,j = 1$ and 2.

3.1.3 Generalization to Random Fields

So far it has been shown that given a bandlimited picture, one can determine the sampling strategy in the picture plane and the interpolation function such that the picture can be reconstructed without error from the samples. Rosenfield and Kak [4] formulate the same results for random fields.

In order to achieve perfect image reconstruction in a imaging system, it is necessary [5] to bandlimit the image to be sampled, spatially sample the image at the Nyquist or higher rate, and properly interpolate the image samples.

3.2 Sampling Using Orthonormal Functions

Sampling is used [4] to represent a picture by a finite string or array of numbers. If picture reconstruction is possible from these numbers it does not matter what specifically the numbers represent. Thus, the numbers need not represent grayscale and this section shows that if a picture function is expanded in terms of a set of orthonormal functions, the coefficients of the expansion can be taken as the picture samples.

To illustrate this method of sampling, it is necessary to talk about ORTHONORMAL EXPANSIONS. Let $f(x,y)$ be a real function

defined over a region f of the xy-plane. Assume that f(x,y) is square integrable.

$$[f(x,y)]^2 \, dxdy < \infty \tag{10}$$

If we are given a set of square integrable functions $\phi_{min}(x,y)$, m = 0,1,2...; n = 0,1,2,...; defined over the same region ρ. The set of functions is carried orthonormal if $\iint \phi_{mn}(x,y) \, \phi^*_{pq}(x,y) dxdy = 0$ (11) for $m \neq p$ or $n \neq q$, m, n, p, q = 0,1,2...3. If, in addition, the following property holds

$$\iint_{\rho} |\phi_{mn}(x,y)|^2 \, dxdy = 1, \; m, n = 0,1,2 \tag{12}$$

then the set of functions is called ORTHONORMAL. The functions ϕ_{mn} may be either real or complex valued.

To approximate f(x,y) at all points within ρ by a sum of the form

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a_{mn} \phi_{mn}(x,y) \tag{13}$$

in such a way that mean square error

$$e^2_{mn} = \iint_{Y} \left| f(x,y) - \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a_{mn} \phi_{mn}(x,y) \right|^2 dxdy \tag{14}$$

is minimized, the following theorem is useful.

THEOREM:

The constants a_{mn} that minimize e^2_{mn} are given by

$$a_{mn} = \iint_{\rho} f(x,y) \, \phi^*_{mn}(x,y)\,dxdy \tag{15}$$

(stated without proof, refer to [4] for it).

The relevant issue now is whether the approximation stated by (13) becomes more accurate as the number of terms in (13) increases. The dependence of the mean square error e^2_{mn} on M and N depends on the nature of the orthonormal functions.

An orthonormal set of functions $\phi_{mn}(x,y)$ m = 0,1,2...; n = 0,1,2...; is called complete if for every square integrable function f we have $\lim_{M\to\infty, N\to\infty} e^2_{mn} = 0$ (16)

In otherwords, if the approximation mean square error approaches zero as the number of terms approaches infinity. A complete orthonormal set of functions is called an ORTHONORMAL BASIS.

It can be said that given an orthonormal basis $\phi_{mn}(x,y)$, m = 0,1,2...; n = 0,1,2...; defined over a region ρ of the xy-plane, then any function f(x,y), square integrate over ρ, can be expanded as

$$f(x,y) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} a_{mn}\,\phi_{mn}(x,y) \tag{17a}$$

FIGURE 3.2(a): PHOTOGRAPHIC PRINT OF A PICTURE BEFORE SAMPLING (Courtesy Perkin-Elmer Corp.)

Fig. 10a Photographic print of a picture before sampling.

FIGURE 3.2(b): PICTURE OF FIGURE 3.3(a) AFTER SAMPLING AND RECONSTRUCTION (Courtesy Perkin-Elmer Corp.)

Fig. 10b Picture of Fig. 10a after sampling and reconstruction. (Courtesy Perkin-Elmer Corp. These pictures appear in the book cited in [9].)

where $a_{mn} = \iint_{\rho} f(x,y)\phi^{*}_{mn}(x,y)\,dxdy$ (17b)

this expression being valid only over the region ρ of the xy plane.

If it is desired to trnsmit an image over a communication link, then (17) shows that we need transmit the coefficients a_{mn} only since by (17a) they can be used to reconstruct the picture at the receiving end.

There are various methods of sampling using orthonormal expansions:

1] Fourier Sampling

If the region in the x-y plane over which the pictures are to be sampled is a rectangle with sides A and B and the orthonormal basis of the function

$$\phi_{mn}(x,y) = \frac{1}{\sqrt{AB}} \exp[j2\pi\,(\frac{m}{A}x + \frac{n}{B}y)] \qquad (18)$$

the sampling in this way is called Fourier sampling.

The Fourier samples of a picture f(x,y) are obtained by substituting 18 in 17b:

$$a_{mn} = \frac{1}{\sqrt{AB}} \int_{-B/2}^{B/2}\int_{-A/2}^{A/2} f(x,y)\ \exp[-j2\pi(\frac{m}{A}x + \frac{n}{B})y)]dxdy \qquad (19)$$

where it has been assumed that the origin is at the center of the picture. Using a Fourier transform relationship, 19 can be written as

$$a_{mn} = \frac{1}{\sqrt{AB}} F\left(\frac{m}{A}, \frac{n}{B}\right) \tag{20}$$

where $F\left(\frac{m}{A}, \frac{n}{B}\right)$ is the value of the Fourier transform $F(u,r)$ of the picture at $u = M/A$ and $v = n/B$. Various methods of computing the Fourier samples a_{mn} of a picture are currently in use.

ii] Standard Sampling

In this technique each sample is obtained by averaging a small portion of the picture. Let the region in the xy-plane over which the pictures are defined be again a rectangle with sides A and B. Over this region the orthonormal basis functions are defined as

$$\phi_{mn}(x,y) = \frac{MN}{AB} \begin{cases} \frac{mA}{A} \leq x < \frac{(m+1)A}{M} \\ \\ \frac{nB}{N} \leq y < \frac{(n+1)B}{N} \end{cases}$$

$$= 0 \quad \text{elsewhere} \quad m = 0,1,2,\ldots,M-1$$
$$n = 0,1,2,\ldots,N-1 \tag{21}$$

where it has been assumed that the origin coincides with the lower left corner of the rectangle. In other words, the picture area is divided into MN rectangular regions, and each $\phi_{mn}(x,y)$ is constant

over one of these regions and zero elsewhere. For example, if M = 8 and N = 6, the set of orthonormal functions generated by the previous definition would have 48 members. As an illustration the member $\phi_{3,2}(x,y)$ of this set is shown in Figure 3.3.

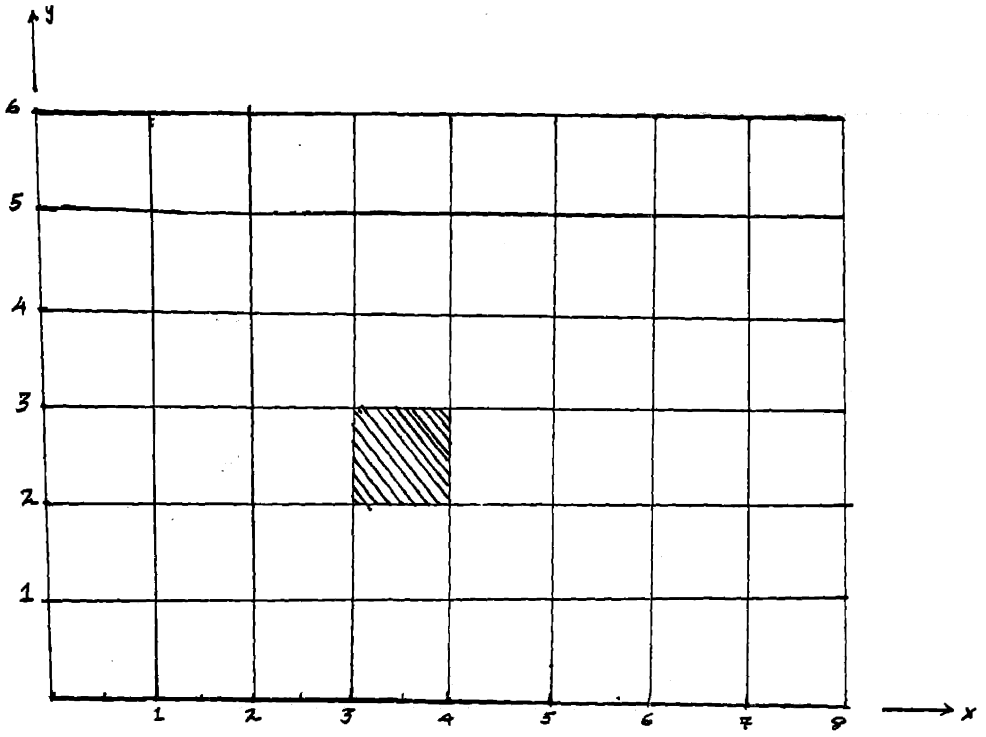


FIGURE 3.3: AN EXAMPLE OF AN ORTHONORMAL FUNCTION USED IN STANDARD SAMPLING THE FUNCTION $\phi_3(x,y)$ SHOWN HERE BELONGS TO THE ORTHONORMAL SET GENERATED BY (21) WITH M = 8 and N = 6

[iii] Optimal Sampling

Optimal sampling of a random field defined over a region ρ of the xy-plane is achieved by the set of orthonormal functions $\phi_{mn}(x,y)$ which for every M and N yields the minimum value of the sampling error.

Figure 3.4 compares these three methods.

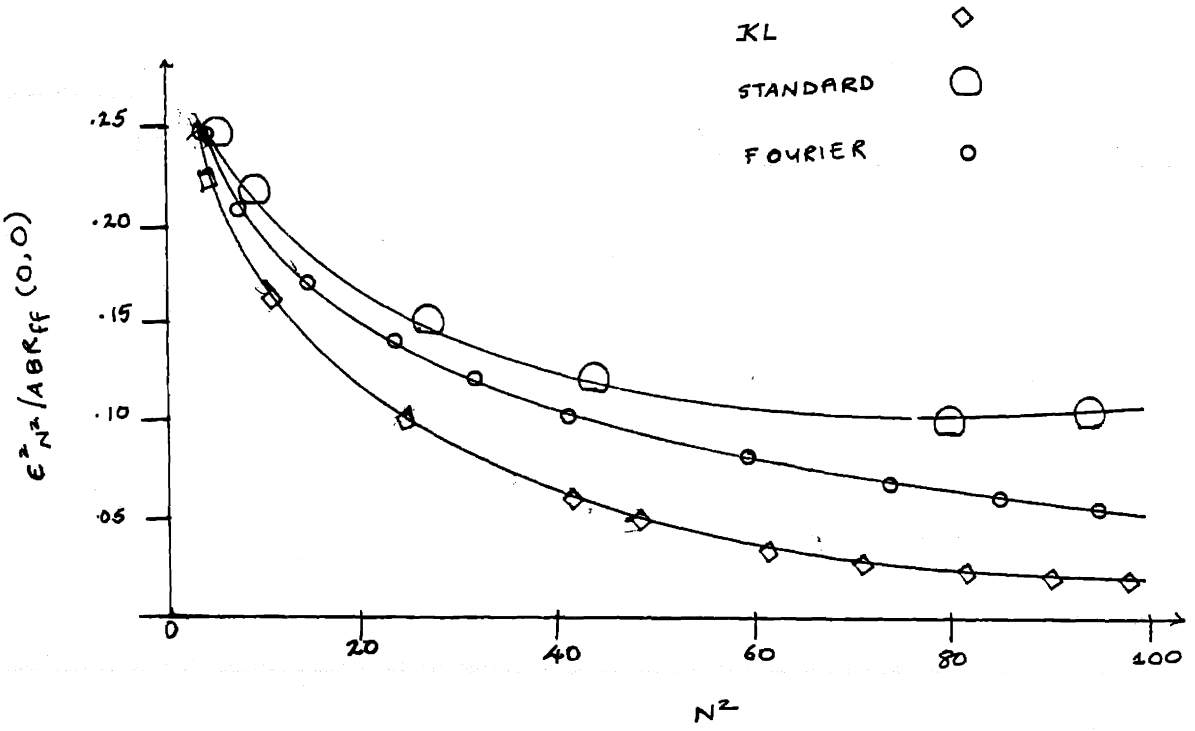


FIGURE 3.4: NORMALIZED SAMPLING ERROR FOR A RANDOM FIELD USING FOURIER , STANDARD , AND OPTIMUM SAMPLING TECHNIQUES AS A FUNCTION OF THE NUMBER OF SAMPLES N^2

3.3 Method Chosen: Sampling Using an Array of Points

Sampling using the method of which required the samples to be the grey levels at an array of point was chosen mainly due to practical considerations. The validity of this method was established in 3.1. The video components currently available use this technique. The change coupled device built by General Electric and used by the author uses an array of sensor to convert light to electrical signals. The magnitude of the electric signals is directly proportional to the light intensity and represents grayscale. An array of sensors consists of:

M rows

N columns

So as to make up MxN sensors, each of these sensors thus represents a single picture element. The electrical signal that emerges from the charge coupled device needs to be digitized before transmission. The digitizer in the system performs this conversion throught a high speed A/D converter. The number of digitized levels available for each signal from the CCD corresponds to the "grayscale" in the picture(6 bits).

Now each picture is represented by

MxNx6 bits/picture.

For a dynamic scene, the "framerate" or # of frames (pictures)/second is also of importance. If this is f frames/sec, the

total transmission rate is MNbf bits/second.

In summarizing, it has been shown that MxN picture elements can be used to represent a picture. The greater M and N are the better (closer to original) the reconstructed picture will be. Each of the picture elements has grayscale. The higher the number of levels of gray, the better the reconstructed image will be. Furthermore, a faster framerate will more accurately picture the scene.

CHAPTER IV

SYSTEM GREAF: A SYSTEM TO PROVIDE ADJUSTABLE GREYSCALE, RESOLUTION AND FRAME RATE

The main purpose of this study was to examine the trade-offs between frame rate, resolution and grayscale for a fixed bandwidth. Towards accomplishing this end, it was necessary to construct a system which would allow an operator to set these parameters while making sure that their product (the baud rate) was not so high as to be out of sight of current undersea transmission technology.

The system constructed was a combination of commercially available components and special purpose hardware which had to be built.

The commercially available components included a camera and its accompanying interface. To avoid interfacing problems it was decided to buy a camera from a company which would also sell the accompanying interface. This interface is essentially a high speed A/D converter and is commonly known as a "digitizer".

4.1 Camera & Digitizer

In purchasing a camera and digitizer, the following considerations were taken into account:

(i) Performance: The digitizer is a highspeed analog to digital converter. Different companies offer vastly different accessories with this converter and this results in a tremendous difference in costs. The camera would also be able to perform at different frequencies given the flexible nature of the experiments

were to be conducted.

Thus, performance requirements were that the camera and digitizer be functional under different conditions. That is, it was required the system performance not degrade so rapidly as to be inoperable at settings other than ideal. Naturally, speeds, gray levels and resolutions were performance factors to be considered while evaluating camera and digitizer.

(ii) Reliability: It was of vital importance that the product obtained performed as expected. Failure in the device could lead to delays in the carrying out of experiments. Naturally, a company with a good track record was preferable - even at some extra costs.

(iii) Interfacing Ease: Since a device would be designed to interface with the digitizer, it was of some significance that interfacing be fairly straightforward. Furthermore, a system that was easy to use would save time in the long run.

(iv) Cost: No purchasing analysis is complete without price considerations. Each additional dollar put into a product gives decreasing benefits. It had to be decided at what point to draw the line.

(v) Delivery: The devices we wanted to buy had long waiting periods. This consideration proved to be the most important since we wanted a product which could be bought "off-the-shelf".

After considering various companies such as CVl & Solid State Sales it was decided to get a General Electric TN2200 automation camera and its accompanying PN210A Automation Interface.

The TN2200 is a totally solid state instrument [3] which translates an optical image into a precise electrical signal. The C.I.D. microsensor contains not only the 16,384 pixels of the 128x128 camera but all of the circuit logic necessary to perform a sequential raster scan and to generate synchronization signals. The CID arrays are fabricated as a silicon P-MOS device similar in many respects to some microprocessor and memory arrays. [For technical specifications see Appendix].

THEORY OF OPERATION OF TN2200

The TN2200 has three printed circuit boards housed in a cylindrical camera case as shown in Figure 4.1. The physical arrangement is also shown in this figure. The CID microsensor is the most important part of the camera. It is made up to an array of picture elements, each of which consists of a pair of capacitors arranged to "share" charge through a P-coupled region. Figure 4.2 represents a typical pixel construction and the arrangement of depletion regions. Array arrangement is such that horizontal electrodes connect to the "Row" capacitor in the pixel and vertical electrodes connect to the "Column" capacitor. The row electrodes are made of polycrystaline silicon and are translucent. The column electrodes are aluminum and thus, opaque.

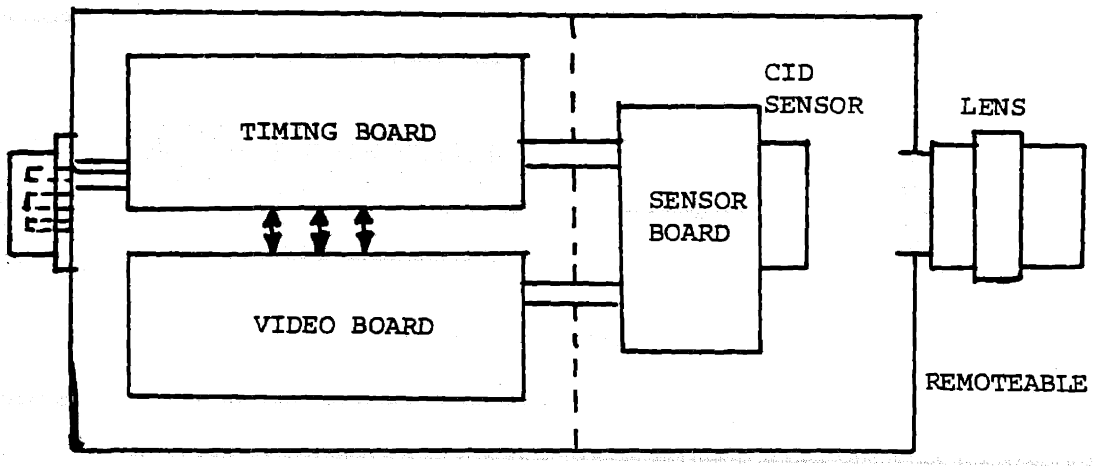


FIGURE 4.1: CAMERA BLOCK DIAGRAM

Within each pixel, depletion regions are caused by the application of row and column bias voltages to the capacitor plates, and are used to trap the positive carriers (holes) generated by light striking the array. The extraction of signal from a single pixel is performed as follows (see Figure 4.2).

I. INTEGRATION (Figure 4.2A)

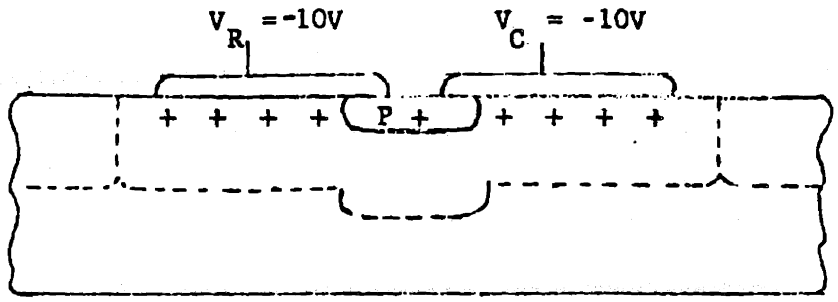
Negative bias is applied to both plates and photon-generated charge is collected. Periodic removal of bias from one plate or the other results in the transfer of charges to the plates which still has bias, but no charge is lost. This half-select condition occurs whenever bias is removed from just a row or column. Charge is lost, or injected, only at the intersection of a selected row and a selected column.

II. ROW SELECT (Figure 4-2B)

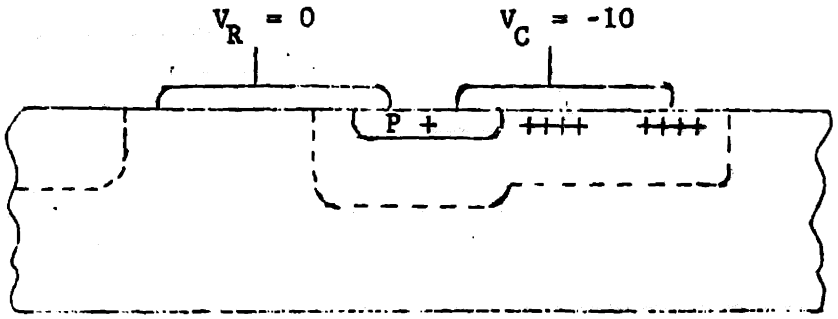
Row selection for read out results in the bias to the tow plate being returned to zero volts, which causes a "half sheet" condition because of the collapse of the energy well. Charge is transferred under the column plates for all elements in the row as described above.

III. CHARGE INJECTION (Figure 4.2C)

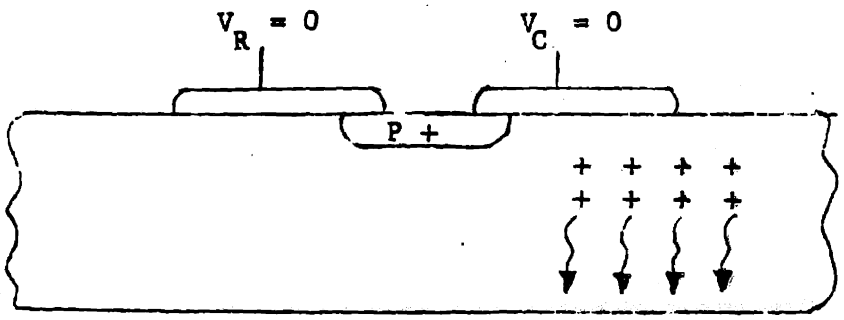
To read out the signal from a single element, the appropriate column electrode is subjected to a capacitively coupled positive pulse which brings the column to 0 volts. This results in half-select condition on the entire column, except at the row position previously



A) CHARGE INTEGRATION



B) ROW SELECTION



C) CHARGE INJECTION

FIGURE 4.2: SIGNAL EXTRACTION FROM A PIXEL

selected. At that site, the charge has no bias to retain it during the positive injection pulse and so it is injected out of the site and into the substrate.

AUTOMATION INTERFACE (PN2110):

The PN2110 automation interface supports the TN2200 camera. The signals it provides [4] include Analog & Digitized Video, Thresholded Video, Power, Clock Signal, TTL signal level buffering and conversion, and analog sweeps for CRT display presentation as shown in Figure 4.3.

Video output signals are available in the form of analog video, 8-bit digitized video on thresholded binary video. This interface will be interfaced to a device which was designed to enable the operator to set the frame rate, gray scale and resolution. For further details refer to the appendix.

4.2 SPOX: The Special Purpose Box

SPOX was the "black box" designed and constructed to provide variable greyscare, resolution and framework. Its function is to take as input 8 bits corresponding to the intensity of the video signal and provide for adjustable parameters.

VARIABLE FRAME RATE:

The purpose of this adjustment is to enable the operator to set the sampling frequency of the video signal. There are two features in this procedure: (1) Sampling and (2) Display. The GE TN2200 can have various sampling periods. The range of variation for these is from

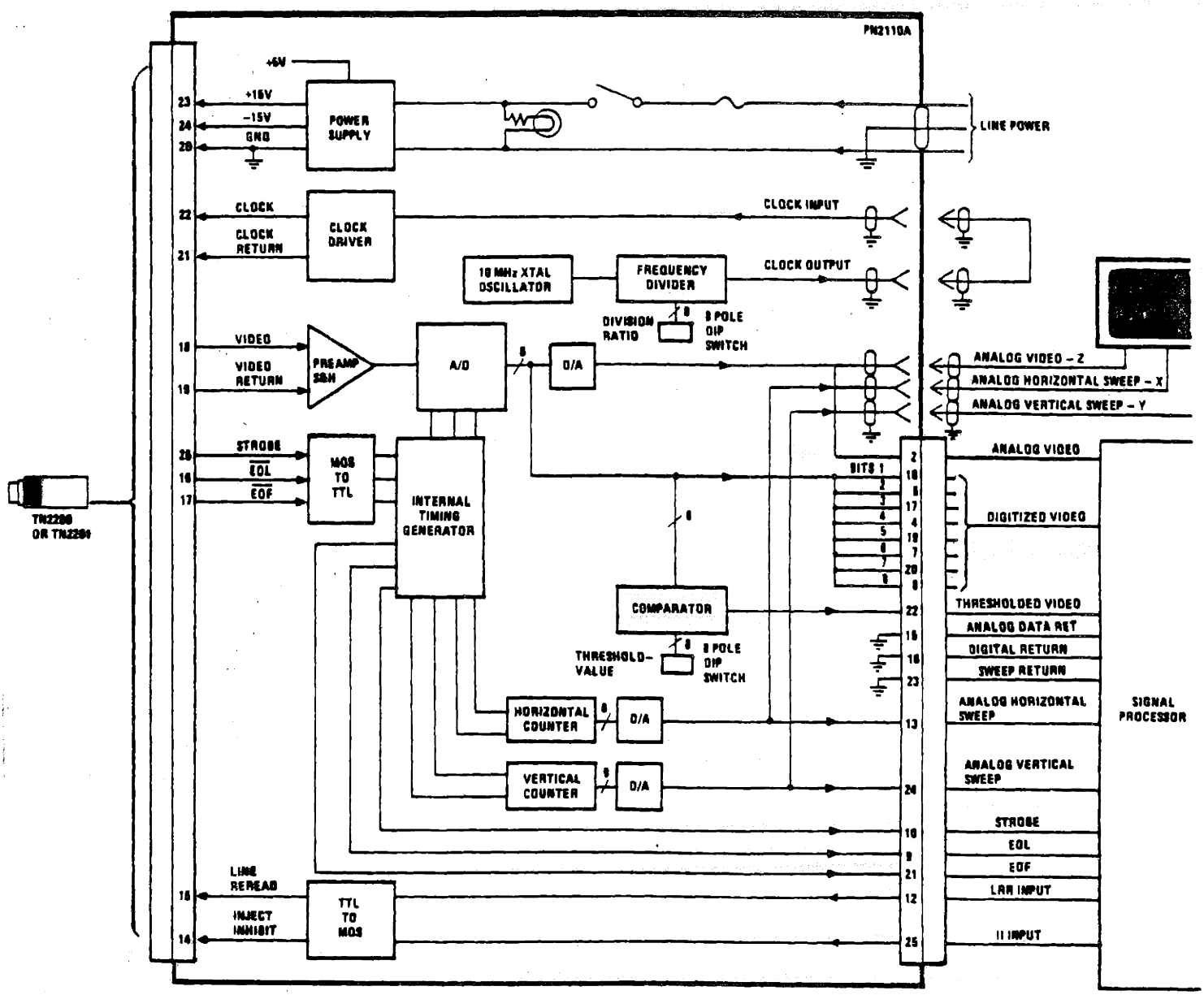


FIGURE 4.3: FUNCTIONAL BLOCK DIAGRAM OF AUTOMATION INTERFACE

2 psec/pixel to 50 μsec/pixel. Performance faster than 2 μsec/pixel is not possible since the Analog to Digital converter cannot function any quicker. Curiously enough, the system cannot work if run too slowly either since this causes integration of noise which saturates the picture. The frame rate adjustment is provided by directly using the FRAME RATE switches in the PN2110 automation interface. Table 4.1 lists the various frame rates corresponding to different switch settings.

The black box was designed so as to continuously display 28 frames/sec regardless of the actual sampling period. If the sampling speed is 28 frames/second then each frame can be displayed as it is sampled. For slow speeds however, it is necessary to store frames in RAM and display each frame more than once. Thus, for a frame rate of 14 frames/sec, each frame would be displayed twice, therefore resulting in a display frame rate of 28 frames/sec. In brief, the adjustable FRAME RATE switches on the box enable the operator to set the sampling frequency while the display frequency is kept constant. [See appendix for more details]. These switches control the strobe and thus the clock rate which determines the sampling frequency.

VARIABLE RESOLUTION:

The automation interface has two registers - x and y, corresponding to the horizontal and vertical scans. The x register. counts from 1 to 134 (the last few counts are for special signals) while the y register goes from 10 to 132.

0 = OFF

1 = ON

MOST SIGNIFICANT SWITCH							LEAST SIGNIFICANT SWITCH	
MSS						LSS		FRAME RATE
0	0	0	0	0	0	0	0	28.26 f/s
0	0	0	0	0	0	0	1	14.13 f/s
0	0	0	0	0	0	1	0	9.42 f/s
0	0	0	0	0	1	0	0	5.65 f/s
0	0	0	0	1	0	0	0	3.14 f/s
0	0	0	1	0	0	0	0	1.66 f/s
0	0	1	0	0	0	0	0	.856 f/s
0	1	0	0	0	0	0	0	.43 f/s [2-3s/f]
1	0	0	0	0	0	0	0	.219 f/s [4.56s/f]
1	1	0	0	0	0	0	0	.146
1	1	1	0	0	0	0	0	

TABLE 4.1: FRAMERATE ADJUSTMENT SWITCHES

For each (x,y) pair an intensity (z output) is produced illuminating each individual pixel. The purpose of the blackbox is to modify the counts from the x and y registers to result in the desired resolution. There are some digital signal processing issues involved since many pixels are now going to have a unique representation. A modification of the standard technique described in Chapter 3 is used for the same purpose. What is essentially being done is data compression. Consider the following transformation in Figure 4.4.

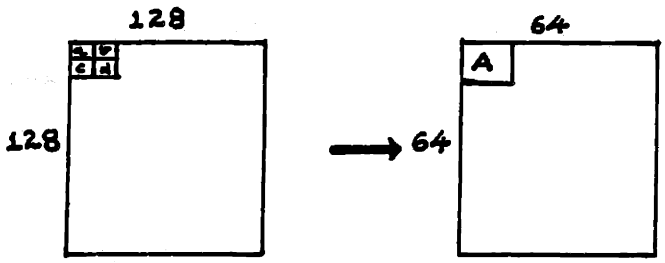


FIGURE 4.4: RESOLUTION AVERAGING

Thus, four different pixels will now have the same intensity. There are various techniques that could be used to achieve this picture compression. The reason why such a transformation can be justified is that the gray levels at all points in each picture are highly correlated. Since the function of system GREAF was to simulate as accurately as possible the effects of low resolution picture transmission, it was decided to adopt a simple sampling scheme where one pixel would reproduced a number of times over. Thus, in Figure 4.4, a,b,c and d would all be represented simply by a which has been spread out and so represented by A. This procedure requires little real time processing

and is therefore very reasonable from the practical view point. A little hardware in BOVAX can be used to accomplish this at fairly high speeds. Table 4.2 represents the various resolutions that different settings correspond to. Figure 4.5 shows pictures of various resolutions.

GRAYSCALE

The automation interface has 8 switches which can be set to provide the required grayscale. These switches were extended to the SPOX. Switches on SPOX can then be used to adjust the grayscale.

Since it was observed that there was not a significant amount of difference between 4 bits of gray and a greater number of bits of gray, only 4 bits of gray were used. That is, there were 4 switches on SPOX to produce 16, 8, 4 or 2 levels of gray. Figure 4.6 shows pictures with different gray scales. Table 4.3 shows the various gray scale settings available.

COMPLETE SYSTEM GREAF

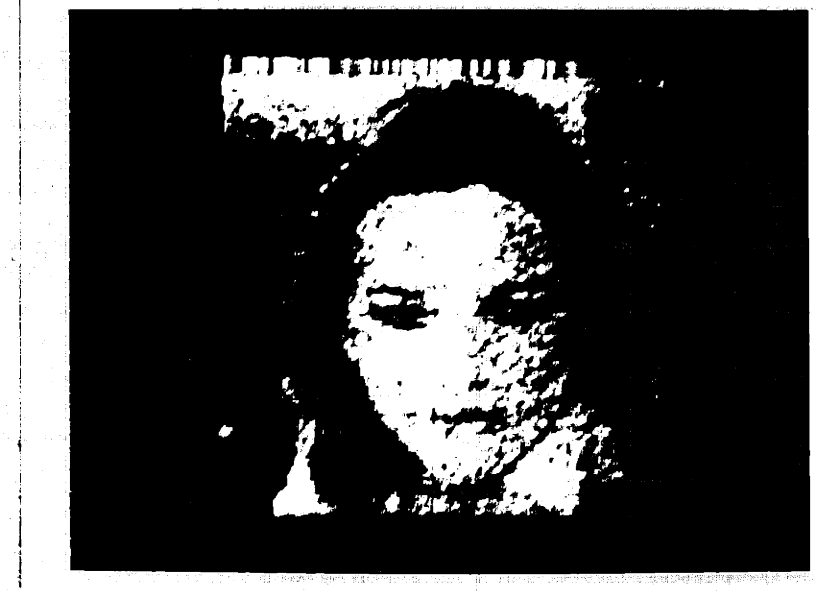
The complete system GREAF consisted of the camera, digitizer, SPOX and a monitor for display. Figure 4.7 shows the complete system.

0 = OFF

1 = ON

HOR				VER				RESOLUTION
1	1	1	1	1	1	1	1	128 x 128
0	1	1	1	0	1	1	1	128 x 128
1	0	1	1	0	1	1	1	64 x 128
1	0	1	1	1	0	1	1	64 x 64
1	1	0	1	1	1	0	1	32 x 32
1	1	1	0	1	1	1	0	16 x 16

TABLE 4.2: RESOLUTION SETTINGS



(a) 128 x 128

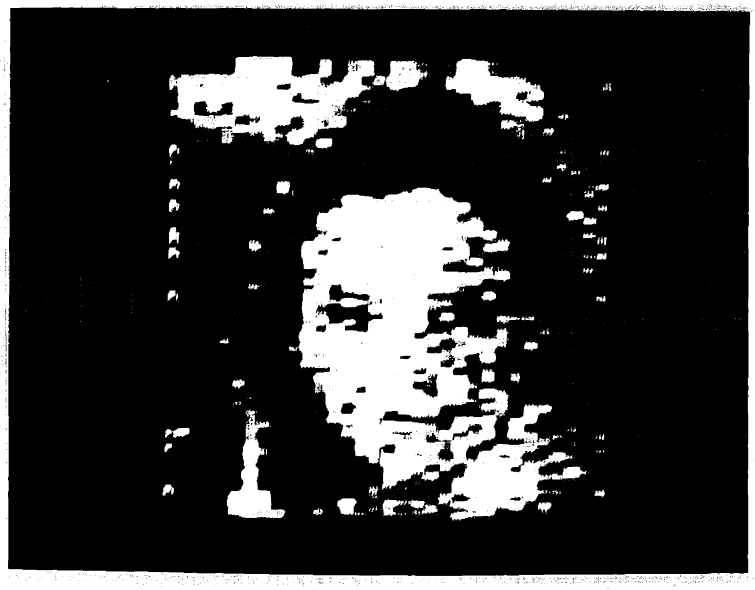
(b) 64 x 128

FIGURE 4.5: DIFFERENT RESOLUTIONS(FRAME RATE = 28 f/s, GRAYSCALE = 4 bits) GENERATED BY SPOX

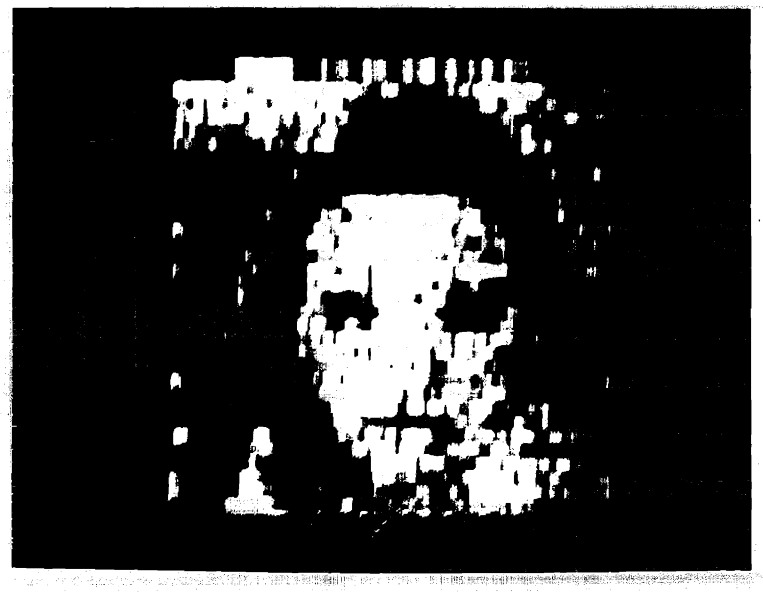
(c) 128x128

(d) 64x64

FIGURE 4.5: (Con't.)

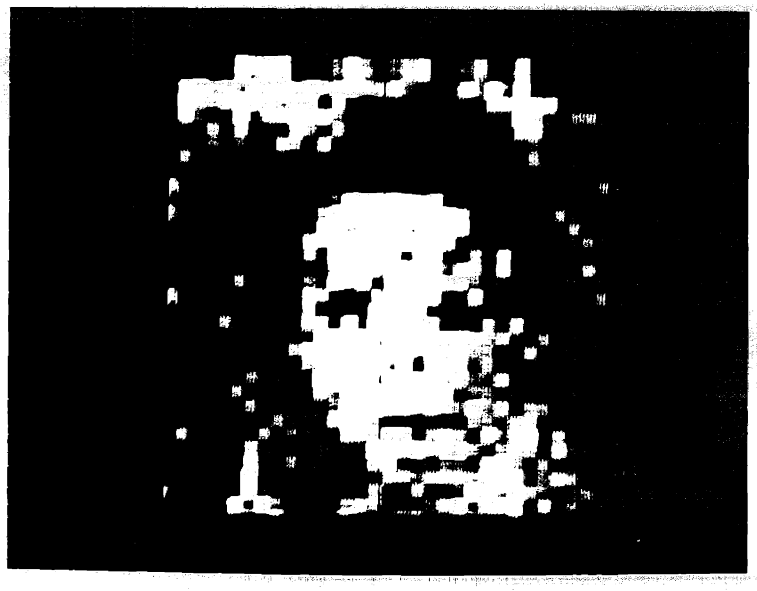


(e) 32x64

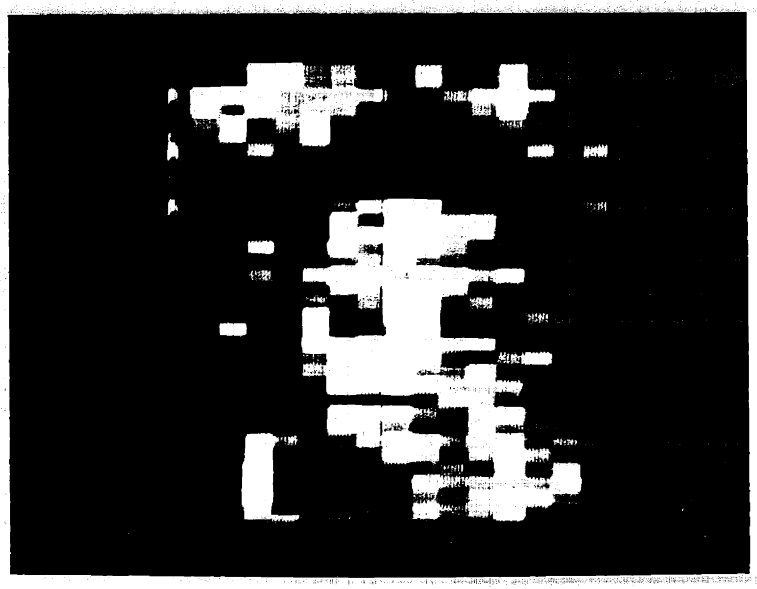


(f) 64x32

FIGURE 4.5: (Con't.)

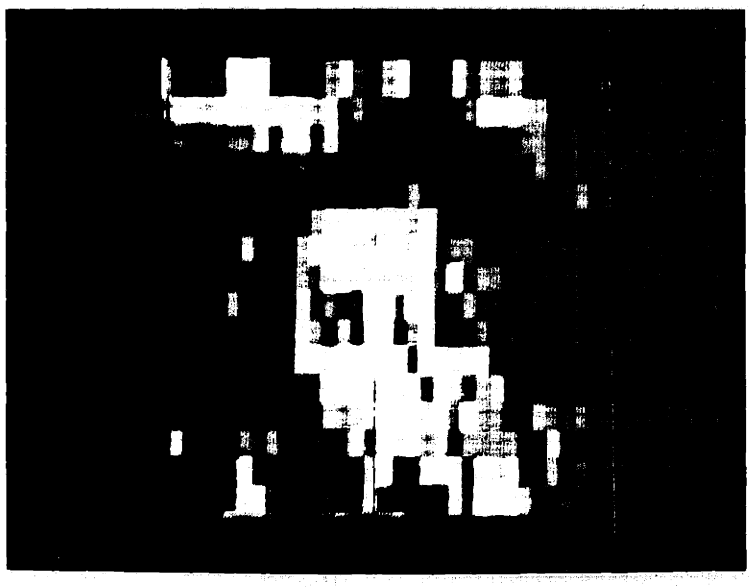


(g) 32x32

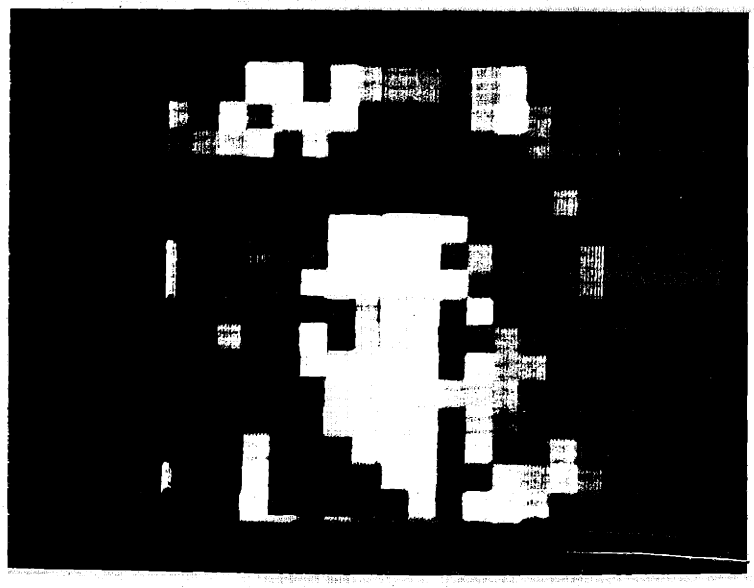


(h) 32x64

FIGURE 4.5: (Con't.)



(i) 32x16



(j) 16x16

FIGURE 4.5: (Con't.)

(a) 3 bits, gray, or 8 levels

(b) 2 bits fray, or 4 levels

FIGURE 4.6: LEVELS OF GRAYSCALE
(resolution: 128x128, framerate: 28 ft/s

0 = OFF

1 = ON

MOST SIGNIFICANT SWITCH				LEAST SIGNIFICANT SWITCH	# of Levels of gray
1	1	1	1		16
1	1	1	0		8
1	1	0	0		4
1	0	0	0		2

TABLE 4.3: GRAYSCALE SETTINGS

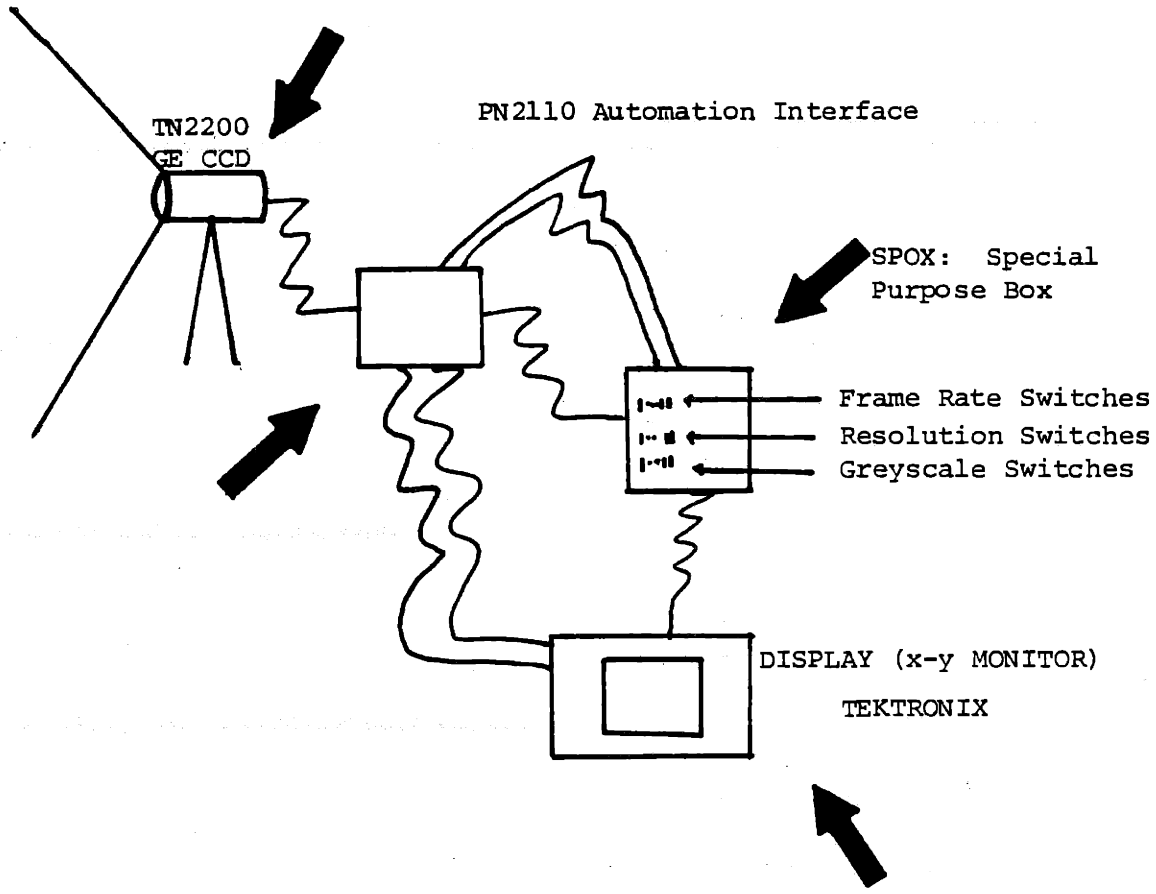


FIGURE 4.7: SYSTEM: GREAF

CHAPTER V

EXPERIMENTS, DATA AND RESULTS

To evaluate the performance of teleoperation under different viewing conditions, a manipulator was used with system GREAF and various tasks. The viewing conditions were varied by altering the three parameters which contribute to picture quality: Frame rate (F), Resolution (R), and Grayscale (G).

The three parameters were adjusted so that the FRG product was kept with current or near future technology in mind.

5.1 Limits of the Video System

System GREAF had the following limits of operation:

HIGHEST FRAME RATE: 28 frames/sec

LOWEST FRAME RATE: .109 frames/sec or 9 seconds/frame

HIGHEST RESOLUTIONS: 128 x 128 pixels

LOWEST RESOLUTIONS: 8 x 8 pixels

MAXIMUM GRAYSCALE: 4 bits (16 levels)

MINIMUM GRAYSCALE: 1 bit (2 levels)

The space within which the video system functions is illustrated in Figure 5.1. Note that these are merely the available settings and do not actually represent if teleoperation is actually possible at any particular point.

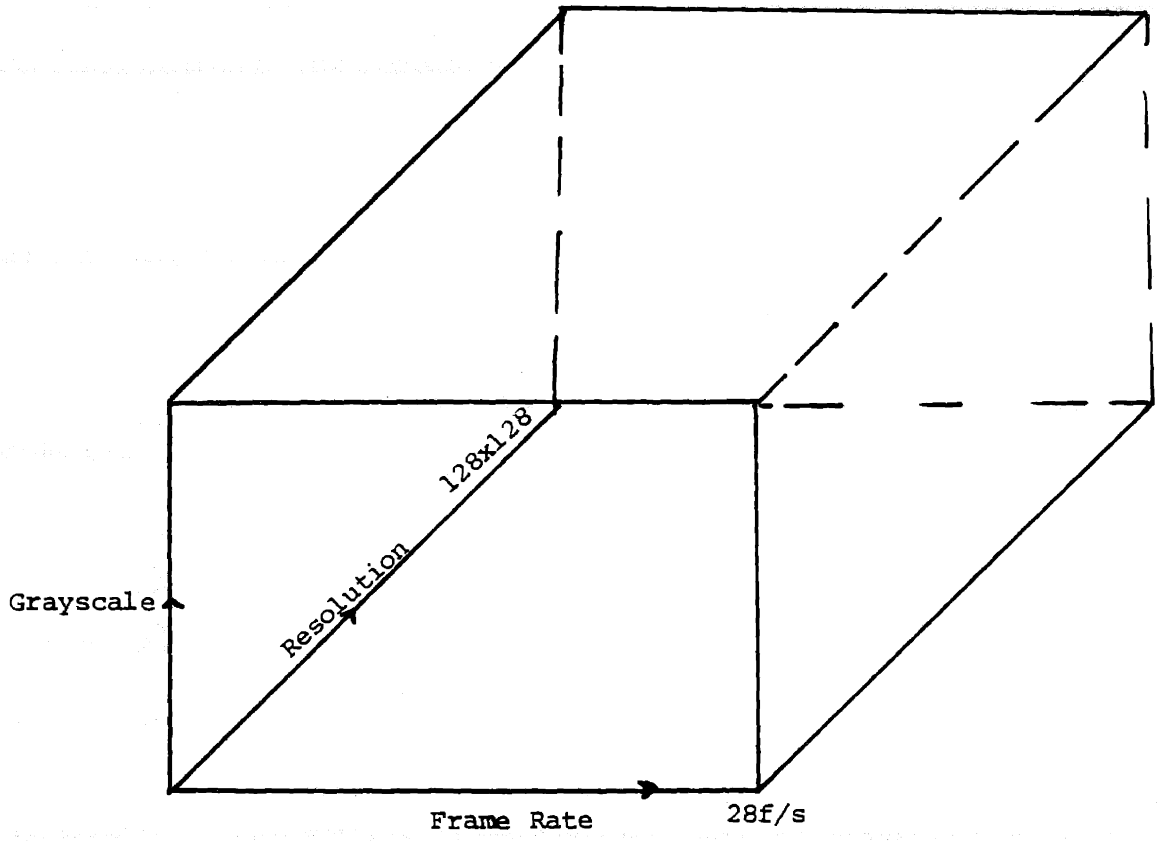


FIGURE 5.1: SPACE OF VIDEO OPERATION

5.2 Experimental Setup

The teleoperator system including SYSTEM GREAF was set up as in Figure 5.2. Figure 5.3 is a picture of the teleoperator laboratory used for the experiments.

A modified Argonne E2 master-slave manipulator was used. It is attached to an Interdata computer which initializes the manipulator arms. This arm can move in all six degrees of freedom plus grasp. Although the manipulator is equipped with force-reflection, this feature was disabled so that the operator's only feedback came through the visual channel. Further, it is known that force reflection with time delay causes special problems of its own, so that a valid experiment to study the effects of visual display variables on teleoperator performance should exclude the complications of force feedback.

5.3 Experimental Considerations

It has been the experience of previous investigators that the time required to complete a task under manual teleoperator control increases with the complexity of the task. It would be expected that deterioration of the picture being used would further increase the task completion time.

In an image transmission system the three parameters which most inlfuence picture quality are frame rate, resolution and grayscale. Picture quality is a term that is used subjectively. Not surprisingly, the higher the frame rate, resolution and grayscale, the clearer and more pleasing the picture appears to be.

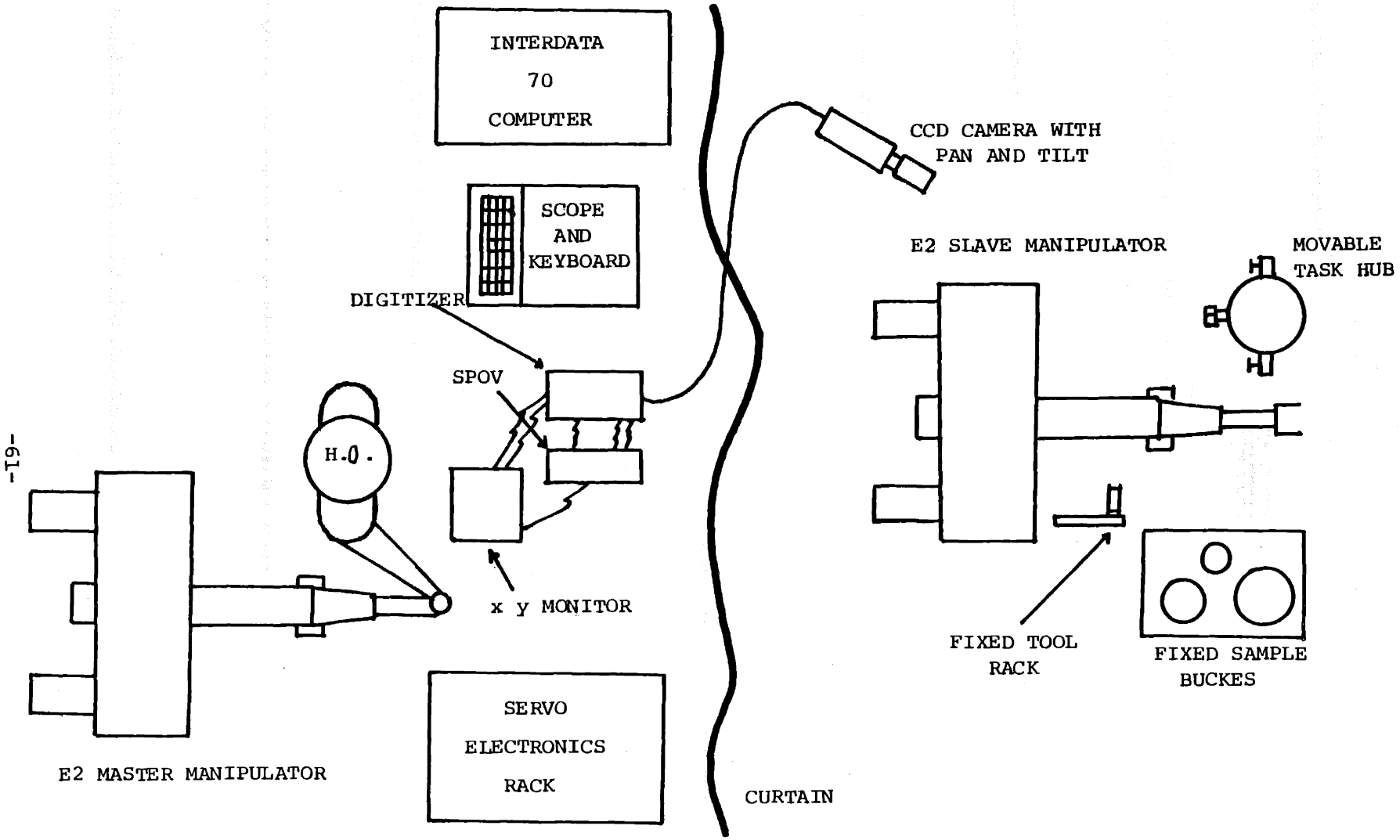


FIGURE 5.2: SCHEMATIC OF EXPERIMENTAL LAYOUT (Plan or Top View)

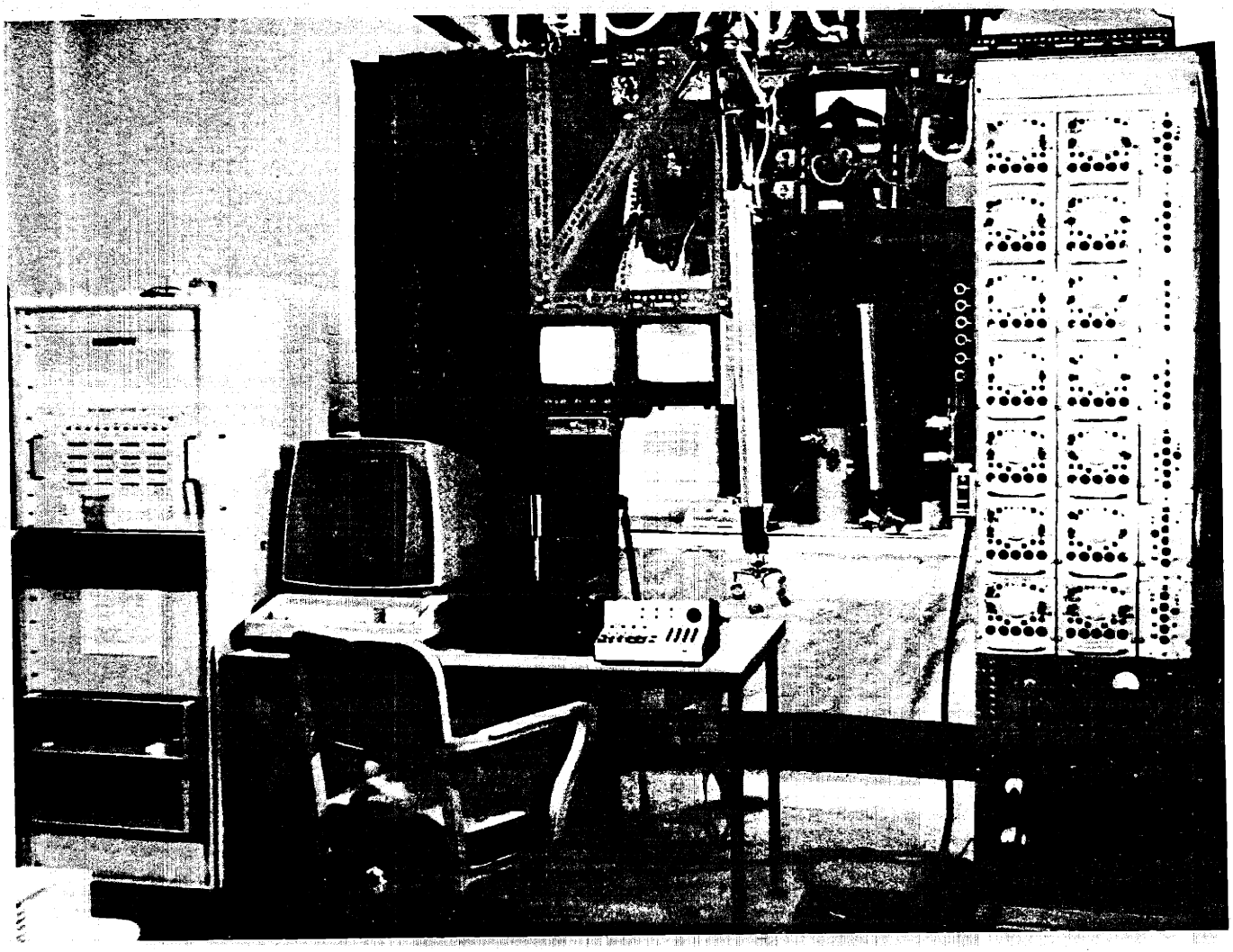


FIGURE 5.3: TELEOPERATOR LABORATORY IN MIT MAN MACHINE LABORATORY

Tasks had to be performable under degrading viewing conditions. The tasks chosen to test manipulator performance were selected to have two features:

[i] task be representative of undersea manipulations jobs

[ii] task performance be correlated with task accuracy

Some of the tasks that could be required of a manipulator in an undersea environment are [7,8,9,10,11]:

1) assessing damage (poking, prying, etc.)
2) bolting/unbolting
3) welding
4) connecting hoses
5) cutting pipes (pipes, wires, cable, etc.)
6) digging
7) drilling
8) tapping
9) fastening
10) lifting objects
11) pulling
12) recovery
13) reaching into confined spaces
14) threading cables
15) untangling cables
16) water jetting
17) wire brushing
18) opening/closing valves
19) sampling
20) coring

Since the purpose of this study was to determine the optimum conditions for video systems used in teleoperator systems generally, rather than to evaluate the particular manipulator system, the representative tasks chosen had to be fairly common. That is, the tasks had to be commonly known and used.

It was decided to use two different tasks to evaluate operator performance. A greater number of tasks would mean a significantly greater amount of time would be required - and in view of the numerous options that had to be experimented with, two tasks seemed an appropriate number.

TAKE-OFF-NUT (TON) TASK: A Useful Task

This task required the operator to locate a nut on a hub (Figure 5.4) and then proceed to take it off. It was important not to drop the nut after removing it. The general method used by the subject was to grasp the nut, turn 180°, pull back to test if the nut was off, and if not, release the grasp, reverse 180°; regrasp and repeat the operation.

The nature of this task is useful and at the same time representative of other useful tasks.

OBJECT PLACEMENT TASK (123T): Test Accuracy

Some preliminary experiments showed that evaluation of the video system being used should in some way incorporate tasks which required fine positioning movements. For this purpose it was decided

FIGURE 5.4: TAKE-OFF NUT (TON) TASK

to design a special purpose task which would allow easy measurement of the precise movements. A "Fitts' Law" task seemed to be an appropriate choice. The operator was asked to grasp a pen and to put dots in the two sets of lines as shown in Figure 5.5. To avoid making the task too rote in nature, this task was modified so that the operator had to touch a grasped marking pen to paper to place a dot in each of a series of circles in a prespecified order (Figure 5.6). Changing this order prevented the task from becoming mechanical - that is, where little or no visual feedback was required.

Preliminary experiments performed using this test showed that the operator tended to bang the pen into the vertical plane of the paper - a consequence of no force feedback. Furthermore, one of the crucial factors in evaluating a picture, depth perception, was not required to accomplish this task. For these reasons, a new task had to be considered.

Task 1-2-3:

As the name suggests, this task consisted of picking up a cork and placing it in squares 1, 2, 3 (Figure 5.7). The operator was given a number of 1-2-3 orderings:

A: middle -1-2-3-2-1-2-1-3-2-middle

B: middle -3-1-2-1-3-2-1-2-3-middle

C: middle -1-2-1-3-2-3-1-2-3-middle

And similarly, D,E. The above notation represents moving the cork in that sequence. For example, "middle-1-2" means the subject

FIGURE 5.5: CONVENTIONAL FITTS' LAW

On receiving a signal, the operator was to put dots between the two sets of lines as fast as possible. The motion was to be a back and forth one. The number of dots placed in a fixed period of time was the measured factor.

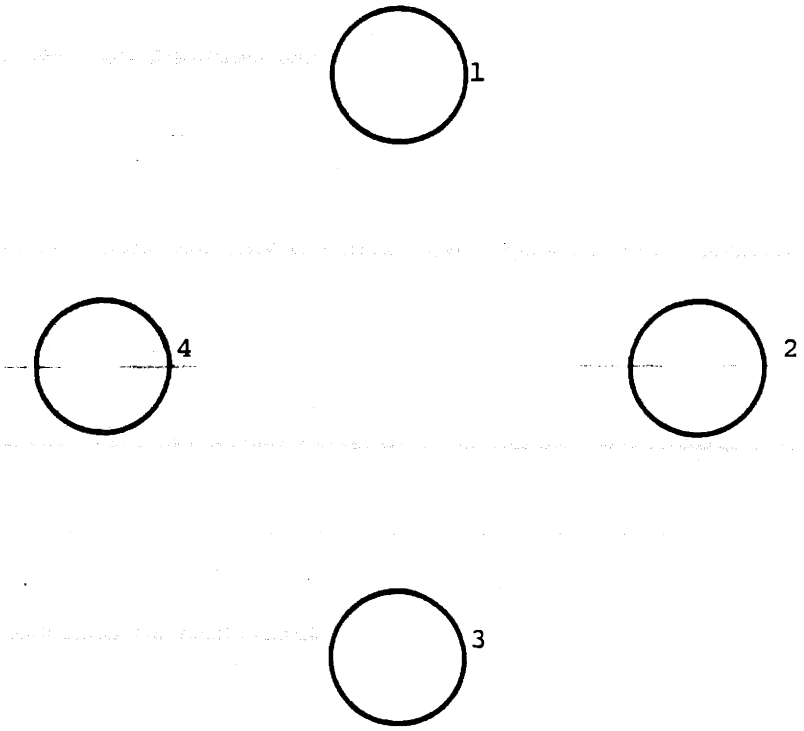


FIGURE 5.6: MODIFIED FITTS' LAW

On receiving a signal A,B,C,D the operator was to commence dotting the circles in the particular sequence

A:	1	2	3	4	2	3	1	2	1	3	2	1	4
B:	4	1	2	1	3	2	1	4	3	2	4	1	3
C:	1	4	3	2	1	4	3	2	3	1	4	3	1
D:	4	3	2	3	4	1	3	2	4	2	3	1	4

The factor measured was the amount of time required to run through the sequence.

picks up the cork from the middle and places it very accurately in square 1. The grasp on the cork is released and the procedure thereafter repeated for the next square in the sequence, square 2 in this example. The fact that the paper orientation was periodically changed and the A,B, C...orderings were random and different prevented the task from being too rote. Accurate perception of depth also became an important factor since the operator had to view motion and objects in the plane of the picture. The measured variable in this experiment was the time required to complete running through one of the sequencies. The subject would not get away with missing the squares during the course of the experiment since the author continuously monotired performance. Thus, in the case where the cork was not accurately placed-the author immediately notified the subject. Thus, the presence of the author served as a deterent to inaccurate task performance.

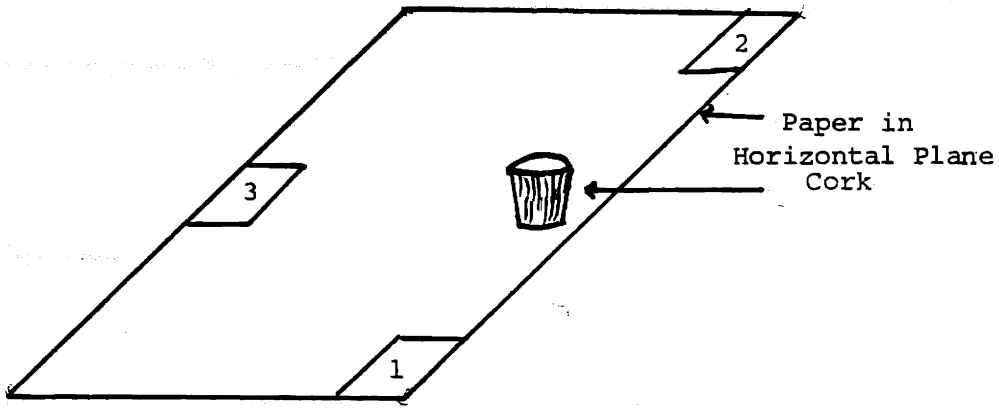
EXPERIMENTS AND SOUNDFEEDBACK:

The preliminary experiments showed that the subjects made considerable use of sound feedback - that is, they were able to hear the effects of their motion. For example, the sound of the manipulator jaws colliding with the task hub or the banging noise of the pen or the paper became valuable ways of position determination. In an ocean environment, sound feedback is not available. Fortunately, the "hum of the air-conditioners in the lab drowned out the sound feedback.

5.4 Subject Training

Two subjects were used as manipulator operators. It was believed important to compromise in the direction of well trained subjects rather than use more subjects who were untrained.

1-2-3 TASK



Pick up the cork and place it in the squares in order depending on signal (A,B,C..)

A: Middle - 1-2-3-2-1-2-1-3-2-middle

B: Middle - 3-1-2-1-3-2-1-2-3-middle

C: Middle - 1-2-1-3-2-3-1-2-3-middle

FIGURE 5.7: 1-2-3 TASKS SET UP

The training period was especially important, as during this time the subjects would be forming certain conceptions in their minds of the tasks and the methods for accomplishing them. After gaining familiarity with the manipulator, force feedback was removed and the subjects were asked to do the ton task with direct vision. After some practice, the subjects were asked to perform the same task using a good video picture. It was only when the subjects were comfortable with this that SYSTEM GREAF and its accompanying degraded quality digitized picture was used instead of the conventional high quality video picture.

To insure that the subjects made definite progress during the training sessions, their performance was continuously monitored. The subjects had the best possible picture from the digitizer during this period. Figures 5.8 and 5.9 are the learning curves for each subject.

In summarizing the results of these learning sessions: at the end of 10 intensive hours with SYSTEM GREAF, the learning curves leveled in comparable fashion and the two subjects were considered trained. The learning curve data is obtained on the basis of 10 trials for each each data point value for each subject.

5.5 Data and Results

The experiments were done in a certain order so that two of the three variables (frame rate, resolution and grayscale) were kept at a constant level while the third was varied. The sections that follow discuss the results for each of the three cases.

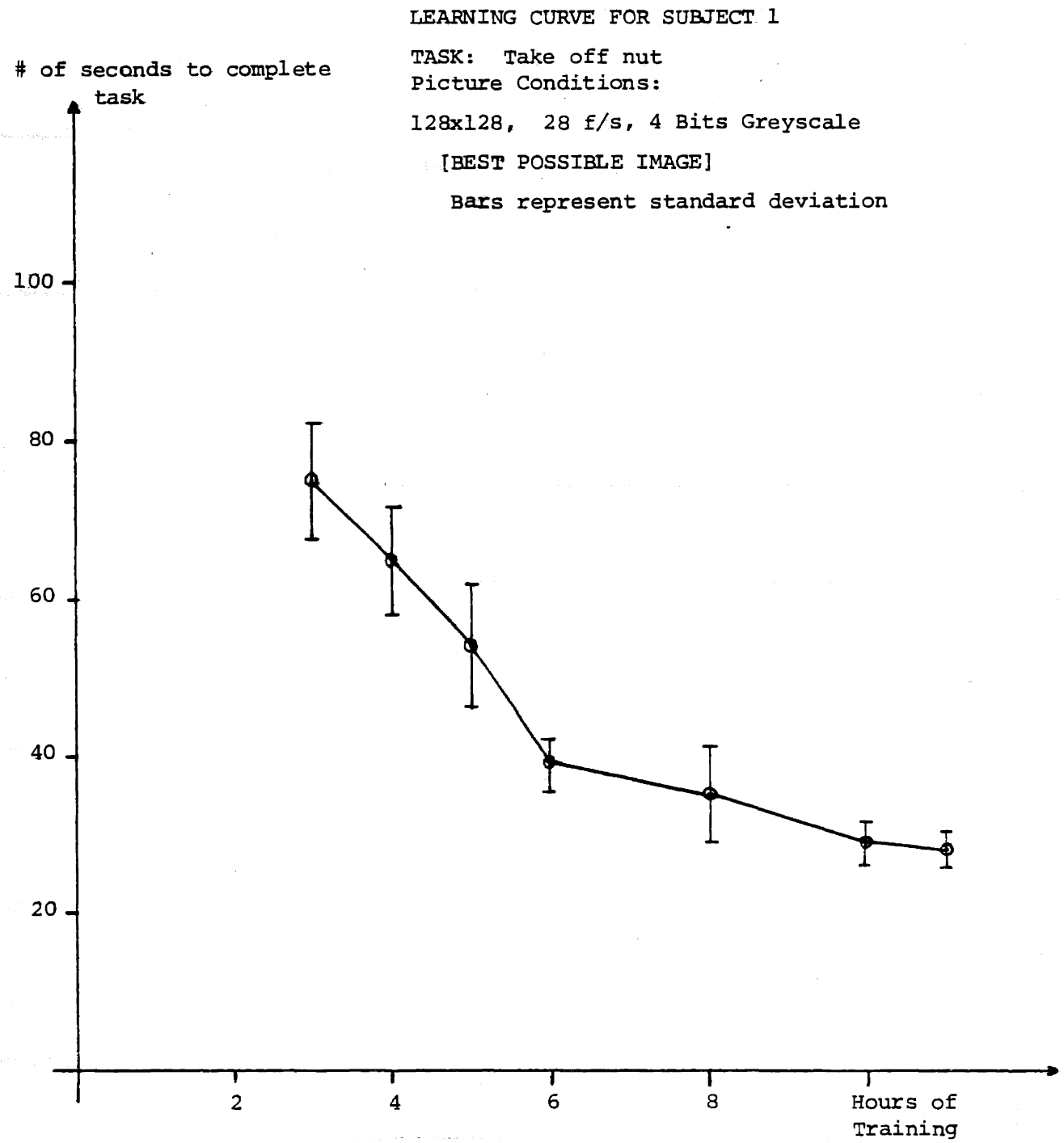


FIGURE 5.8: LEARNING CURVE FOR SUBJECT 1

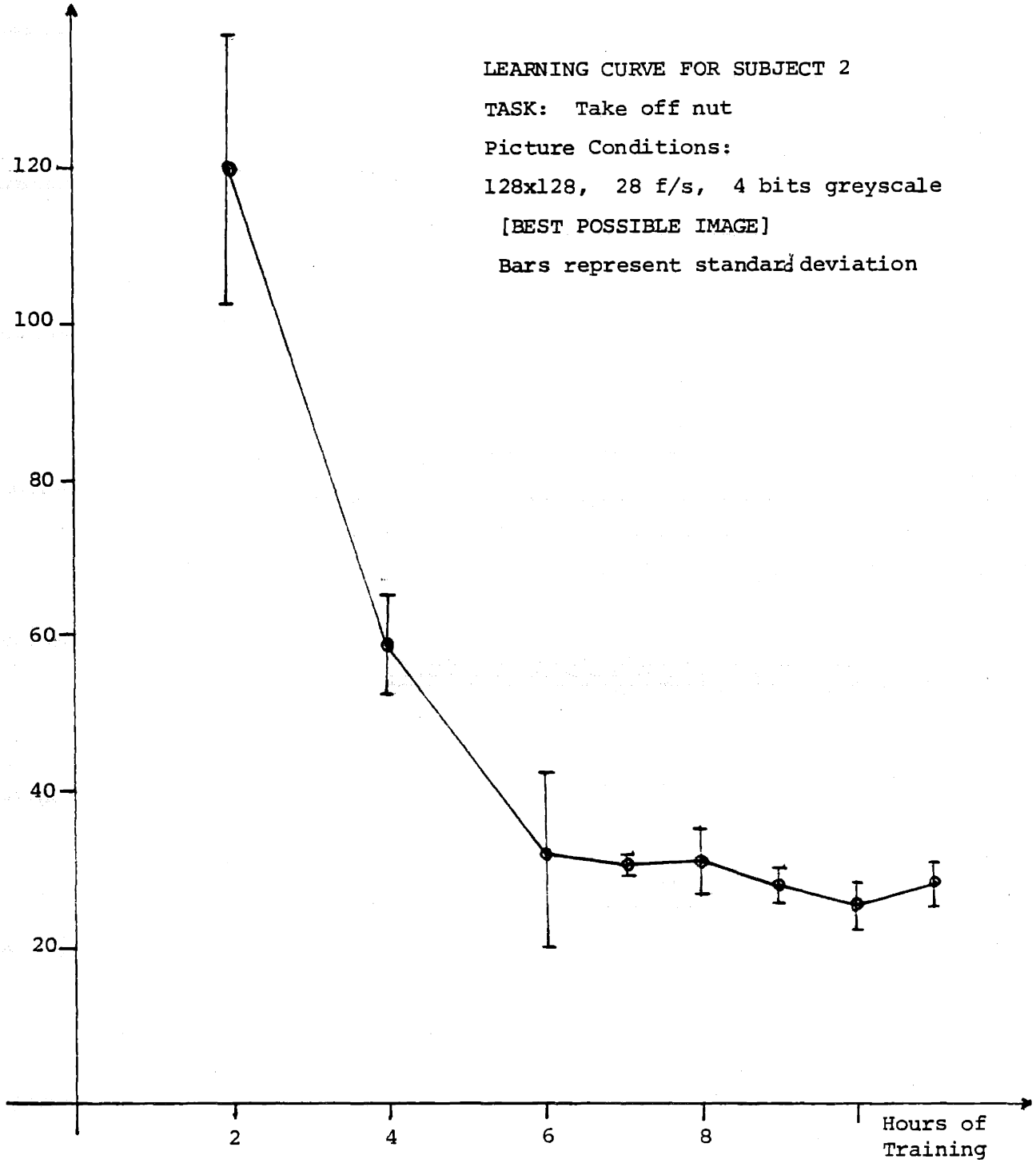


FIGURE 5.9: LEARNING CURVE FOR SUBJECT 2

As a performance baseline the best possible image conditions were used. All results were then compared with respect to this case. The best possible case has the conditions:

128 x 128 pixels resolution
4 bits Grayscale
28 frames/sec Frame Rate

Each subject was allowed to practice for awhile on each new image conditions. Twelve readings (i.e., time to accomplish task) on each were then taken and the last six readings used as data. The ton task hub and 1-2-3 task paper were periodically reorinted to prevent the task from becoming rote.

The average data for the best case (BCD) was surprisingly:

TASK	SUBJECT #1 TIME	SUBJECT #2 TIME
Take-off-Nut (ton)	28 seconds	28 seconds
1-2-3 Task (1-2-3T)	28 seconds	28 seconds

In order to be able to suitably represent the data, a performance factor was defined:

If

PERFORMANCE FACTOR = P

and t_n = Ton time, t_a = 1-2-3T time

t_{BCD} = best case data = 28 seconds for both tasks

{all values average over six measurements}

$$p_n = \frac{28}{tn} \times 100 \quad \text{for Performance in TON task}$$

$$p_a = \frac{28}{ta} \times 100 \quad \text{for Performance in 1-2-3T task}$$

$$P = \frac{P_n + P_a}{2} \quad \text{for overall Performance.}$$

5.5.1 Variable Framerate Case

System GREAF was used so that the grayscale and resolution of the picture were kept constant while the framerate was varied. This was carried out for two different grayscale-resolution combinations.

Table 5.1 summarizes the data gathered for Subject 1 while 5.2 shows the data for Subject 2. Note the similarities in the data for the two subjects. Standard deviations are indicated by parentheses.

Tables 5.1 and 5.2 indicate that 28 frames/second (for 128 x 128 resolution and 4 bits gray) gives the same performance as 14 frames/second. Furthermore, performance degraded ~34% from 28 f/s to 3 f/s and ~28% from 3 ft sec to 1.6 f/s. Thus, at low frame rates not only are performances bad and variances high, but performance also degrades very fast. This is true for both subjects.

The data from Table 5.1 is represented in Figure 5.10 and from Table 5.2 in Figure 5.11. Figure 5.12 is a plot of the performances, p_n, p_a and P as defined earlier. Figure 5.12(a) is for Subject 1 while 5.12(a) is for Subject 2.

SUBJECT #1:

RESOLUTION: 128x128

GRAYSCALE: 4 bits

NUMBERS IN BRACKETS REPRESENT STANDARD DEVIATIONS

Case #	Frames/sec	tn	ta	Overall Performance (P)
1	28	28(3)	28(1)	100%
2	14	28(3)	28(1)	100%
3	9	38(4)	30(1)	83%
4	5.6	42(4)	33(3)	76%
5	3	51(15)	36(2)	66%
6	1.6	91(24)	60(3)	39%
7	856	110(27)	89(5)	28%
8	.43	306(97)	262(42)	10%
9	.22	1200(480)	1000(400)	3%

TABLE 5.1: VARIABLE FRAME RATE DATA FOR SUBJECT 1

SUBJECT #2:

RESOLUTION: 128x128 pixels

GRAYSCALE: 4 bits

NUMBERS IN BRACKETS REPRESENT STANDARD DEVIATIONS

Case #	Frames/sec	tn	ta	Overall Performance (P)
1	28	28(5)	28(3)	100%
2	14	28(6)	28(3)	100%
3	9	39(8)	32(2)	80%
4	5.6	44(6)	35(4)	72%
5	3	56(20)	36(4)	65%
6	1.6	100(30)	64(6)	36%
7	.856	120(30)	94(6)	26%
8	.43	400(150)	300(50)	8%
9	.22	1250(500)	1200(500)	2%

TABLE 5.2: VARIABLE FRAME DATA FOR SUBJECT 2

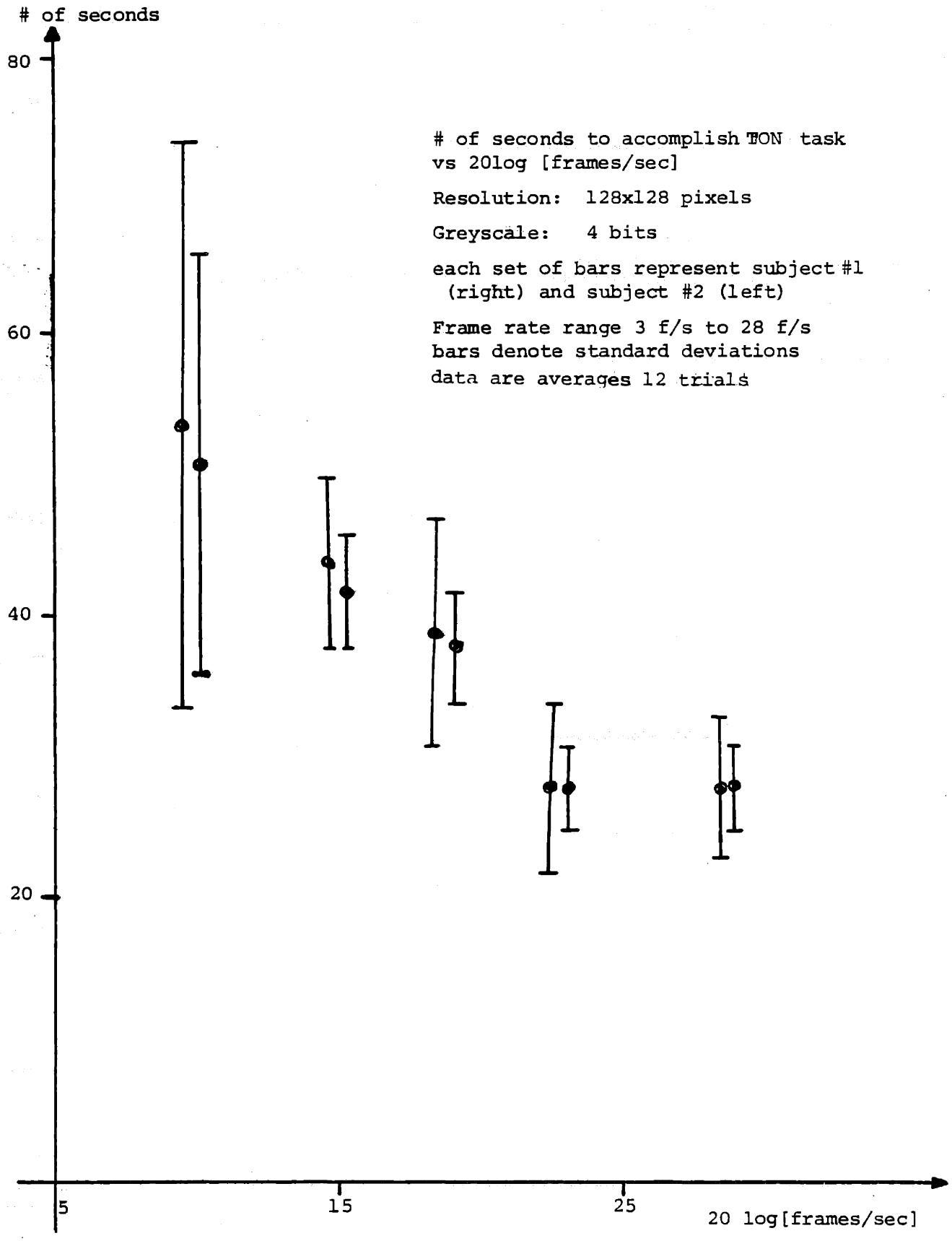


FIGURE 5.10(a): NUMBER OF SECONDS TO ACCOMPLISH TON TASK VS 20log [frames/sec]

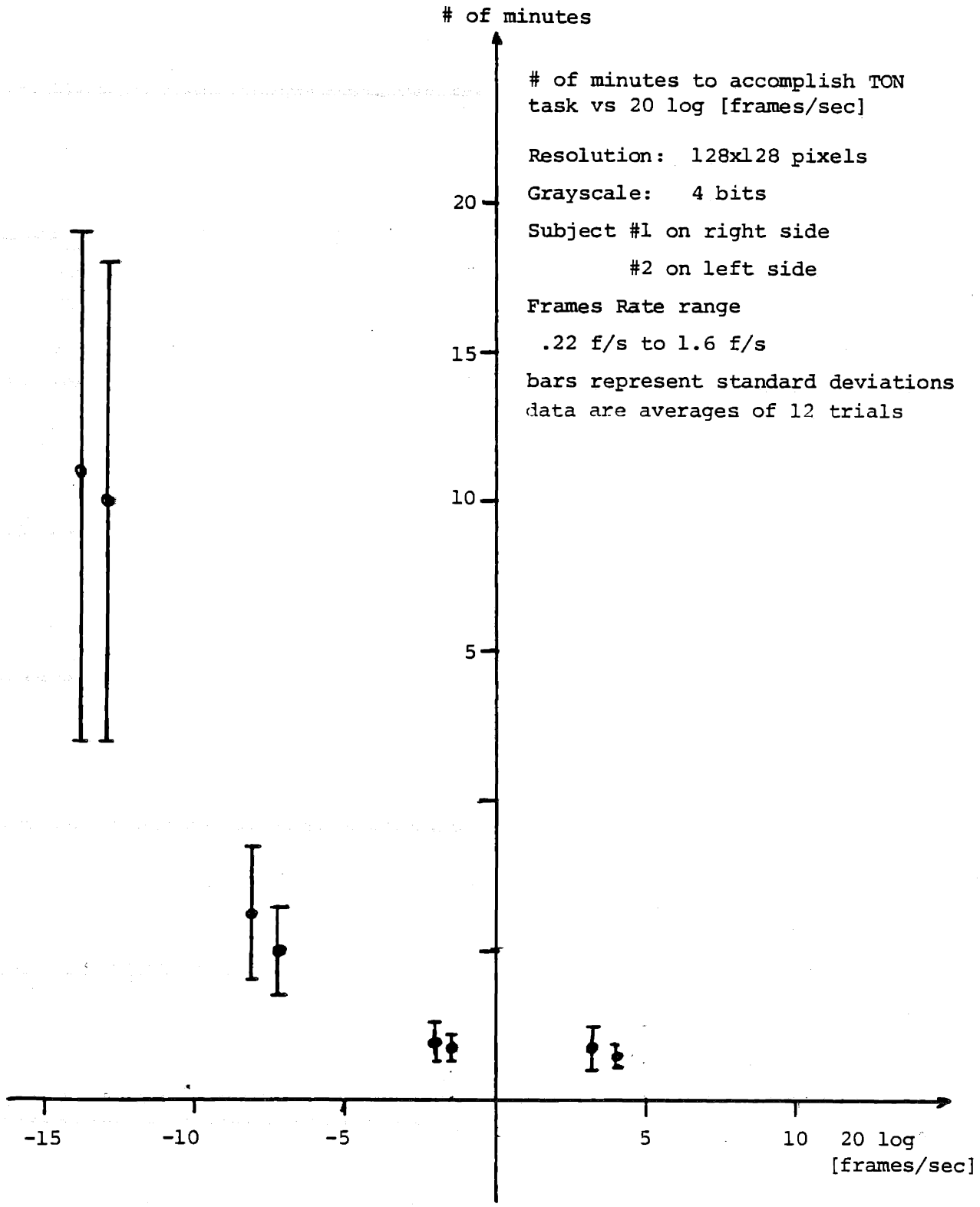


FIGURE 5.10(b): (Con't.)

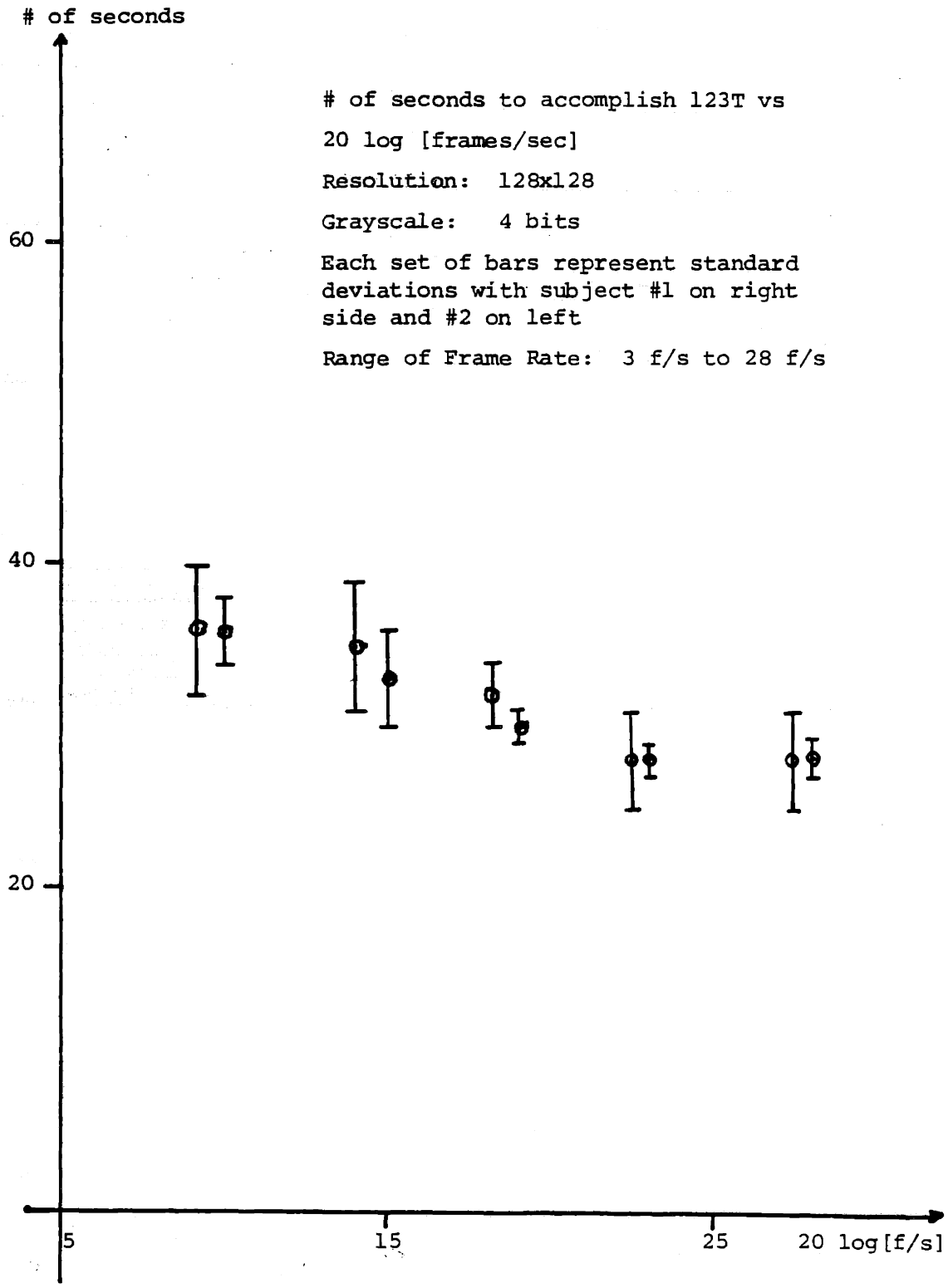


FIGURE 5.11(a): NUMBER OF SECONDS TO ACCOMPLISH 123T VS 20 LOG [frames/sec]

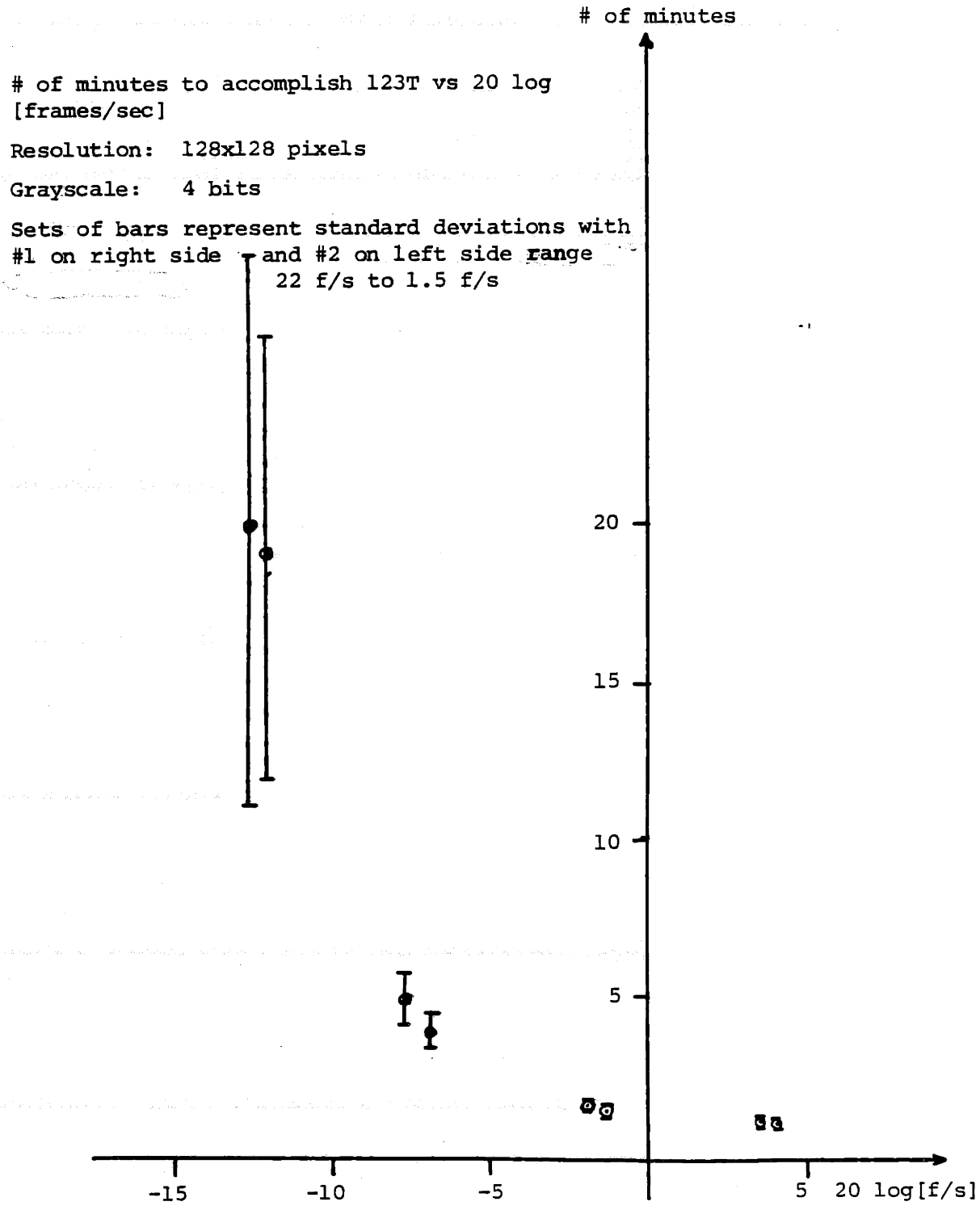


FIGURE 5.11(b): (Con't.)

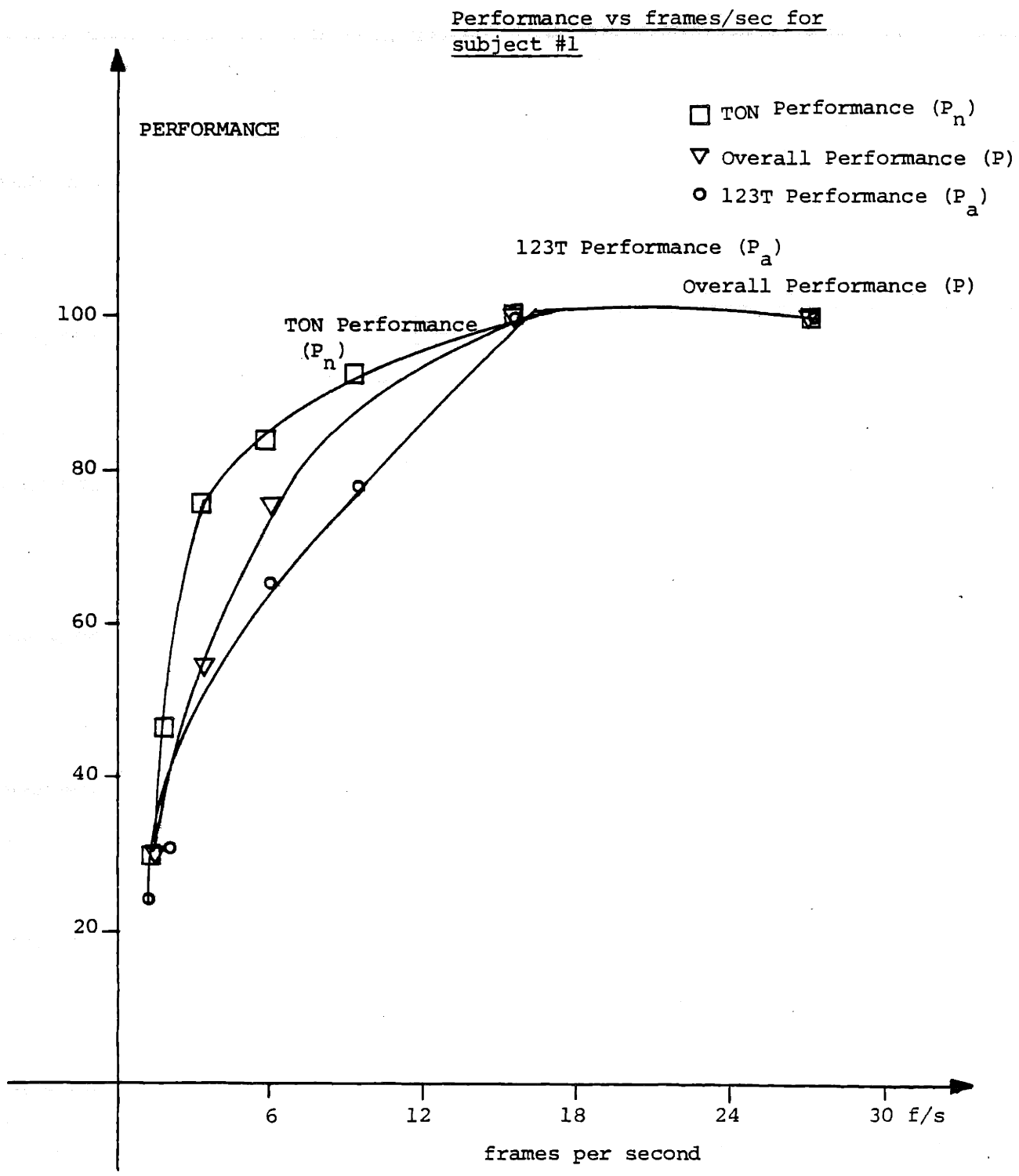


FIGURE 5.12(a): PERFORMANCE VS FRAMES/SEC FOR SUBJECT #1

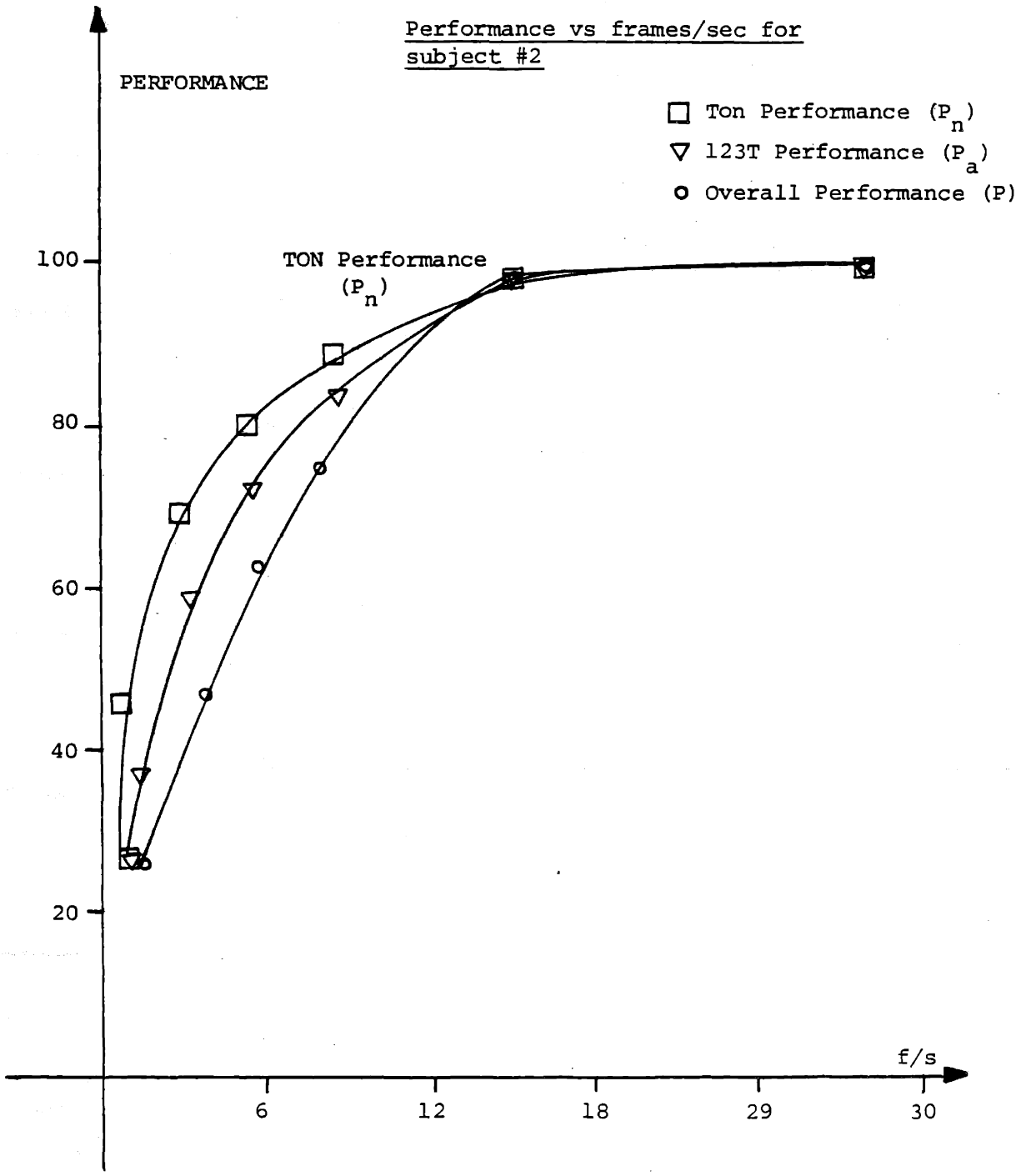


FIGURE 5.12(b): PERFORMANCE VS FRAMES/SEC FOR SUBJECT #2

TABLE 5.3:

VARIABLE FRAME RATE RESULTS FOR SUBJECT #1:

Resolution: 64x64 pixels

Grayscale: 2 bits

Case #	Frames/sec	tn	ta	Overall Performance (P)
1	28	43(4)	90(22)	48
2	14	52(12)	100(25)	41
3	9	85(30)	*	
4	5.6	108(62)	*	

* could not be accomplished owing to a bad quality picture.

TABLE 5.4:

VARIABLE FRAME RATE RESULTS FOR SUBJECT #2:

Resolution: 64x64 pixels

Grayscale: 2 bits

Case #	Frames/sec	tn	ta	Overall Performance (P)
1	28	45(6)	100(30)	45%
2	14	70(10)	730(40)	22%
3	9	89(27)	*	*
4	5.6	120(73)	*	*

* could not be accomplished owing to a bad quality picture.

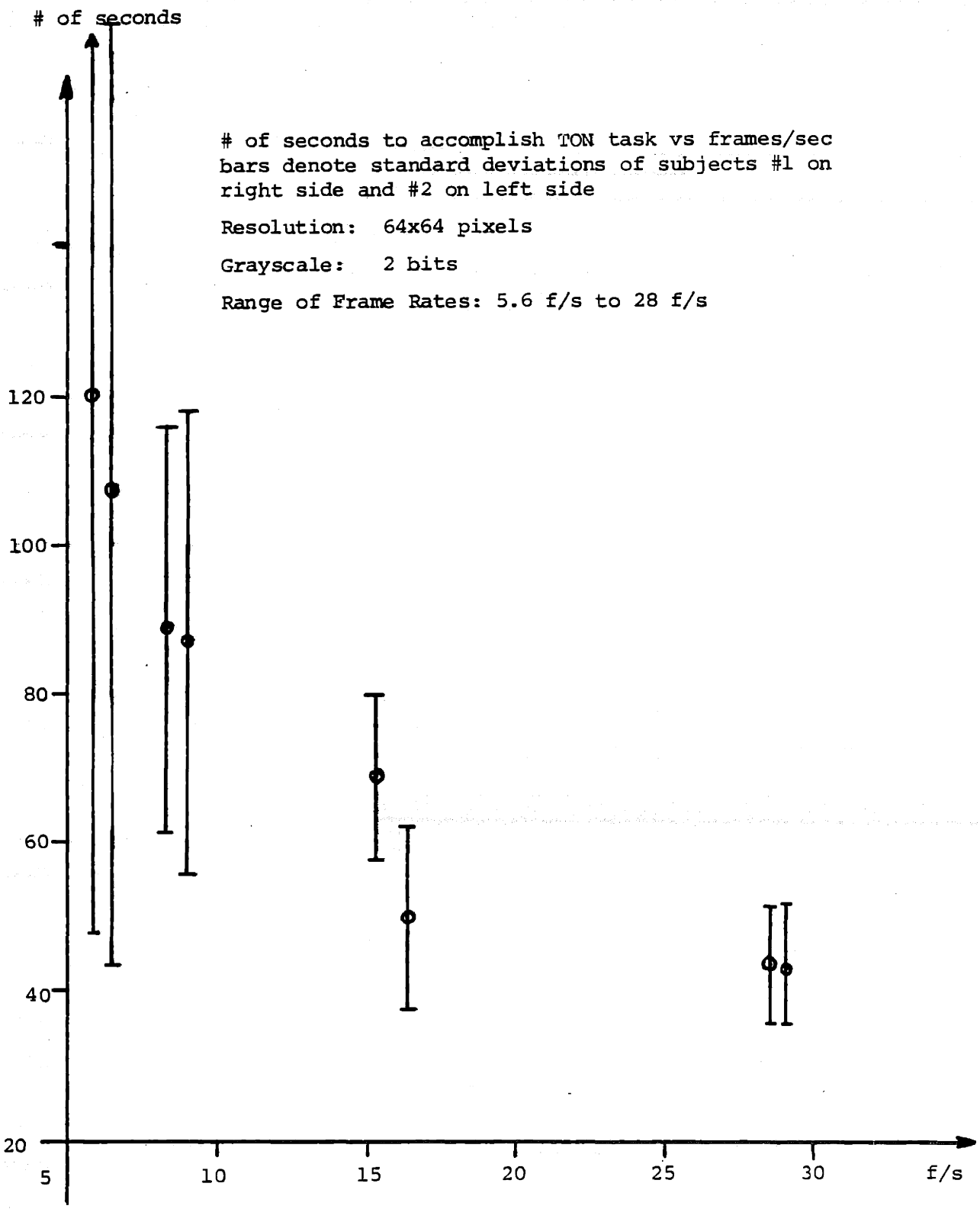


FIGURE 5.13: NUMBER OF SECONDS TO ACCOMPLISH TON TASK VS FRAMES/SEC

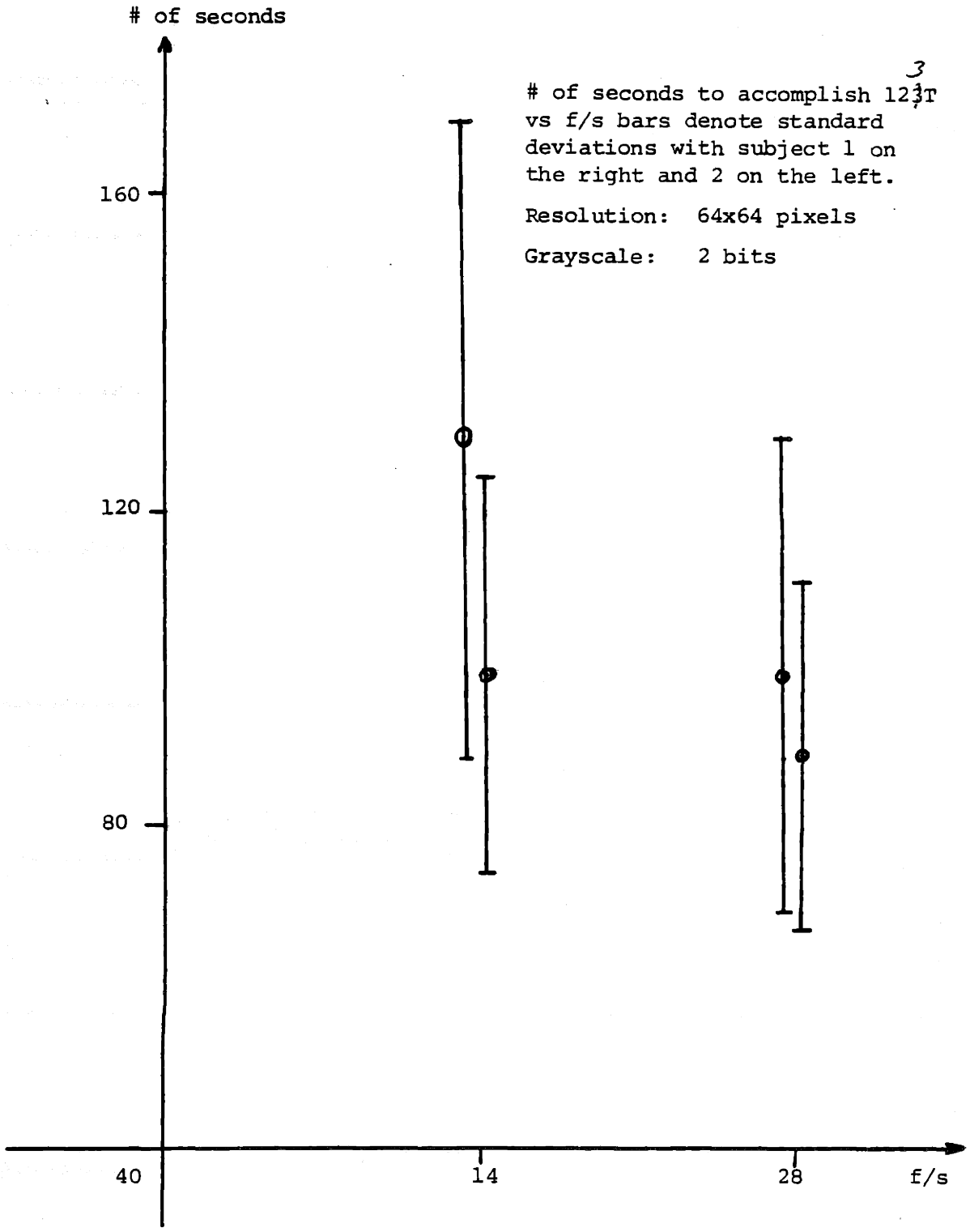


FIGURE 5.14: NUMBER OF SECONDS TO ACCOMPLISH 123T VS FRAMES/SEC

The same experiment was repeated with a different set of grayscale and resolution:

GRAYSCALE: 2 bits

RESOLUTION: 64 x 64 pixels

The frame rate was varied as before. Results for the two subjects are tabulated in Tables 5.3 and 5.4.

In this case, it was impossible to perform the 123T at frame rates below 14 f/s. This is because picture quality appeared to have degraded to the extent of making the lines in the 123 task picture undistinguishable. The results of this experiment are represented in Figure 5.13 and Figure 5.14.

To enable a comparative study of the results, Figure 5.19 shows the performance curves for the TON task for both subjects at the two different settings.

From the results obtained so far it is clear that slower frame rates make a significant difference in performance below 5.6 frames/second. At very low speeds the subjects had to use a move and 2nd wait strategy. This meant that after each discrete motion the operator had to wait to observe the position the manipulator arm finally took as a result of the previous control action. Faster motions were missed. Naturally, the move and wait strategy is time consuming but even at low speeds such as 1 frame every 5 seconds, with sufficient resolution the operator was able to accomplish the task successfully. However, low resolutions

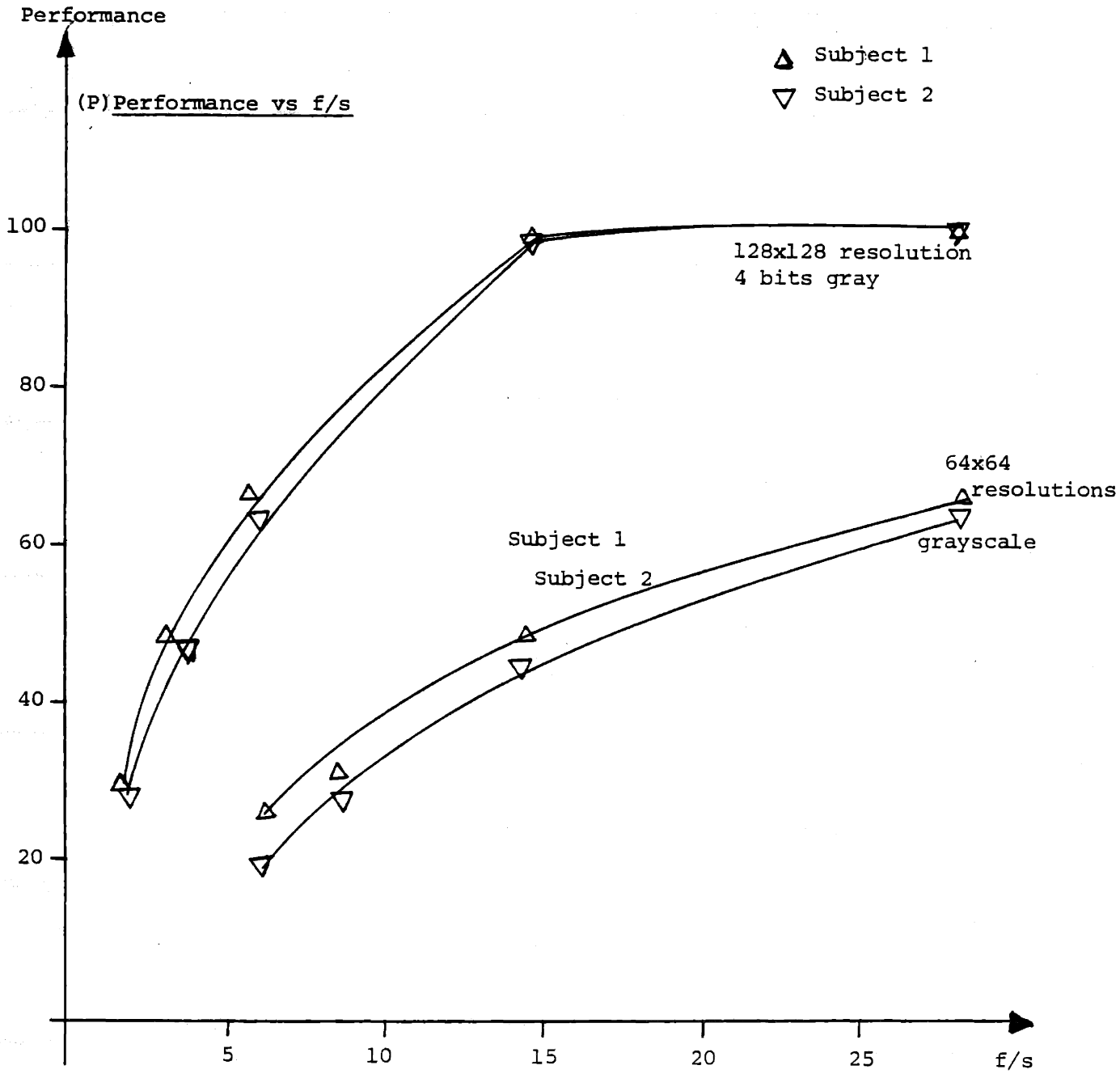


FIGURE 5.15: PERFORMANCE VS FRAMES/SEC

and grayscales cause excessive noise problems.

In the "take-off-nut" task, the most important job was to position the manipulator correctly. The operator achieved this by commanding the jaws of the manipulator to close. If the jaws of the slave did not close, then the operator had succeeded in grasping the nut. After this, the long hours of training paid off, since one 180° turning motion could be accomplished with minimal visual feedback. The repositioning that followed this motion however, once again required depth perception on the part of the operator. At low frame rates subjects had to fight the temptation to move without waiting for a sufficiently long period of time for the previous picture to be displayed. The large deviations in task completion times for the low frame rates can be attributed to the fact that positioning the manipulator was to some extent a "hit and miss" affair: luck played an important role. Thus, on some occasions the operator was able to quickly locate the nut while on other occasions it took a longer period of time.

The 1-2-3 task was easier in some ways and more difficult in others. Lower frame rates had less of an effect on task completion time in the beginning; however, as frame rates got lower, the performance began to degrade faster. At very low frame rates, performances for both tasks converged.

5.5.2 Variable Resolution Case

In this set of experiments the resolution was varied while keeping the frame rate and grayscale a constant. Operator performance

As a performance baseline the best possible image conditions were used. All results were then compared with respect to this area. The best possible case has the conditions:

128 x 128 pixels resolution

4 bit Grayscale

28 frames/sec Frame Rate

Each subject was allowed to practice for while on each new image condition. Twelve readings (i.e., time to accomplish task) on each were then taken and the last six readings used as data. The TON task hub and 1-2-3 task paper were periodically reoriented to prevent the task from becoming rote.

The average data for the best case (BCD) was surprisingly:

TASK	SUBJECT #1 TIME	SUBJECT #2 TIME
Task-off-Nut (TON)	28 seconds	28 seconds
1-2-3 Task (1-2-3T)	28 seconds	28 seconds

In order to be able to suitably respresent the data, a performance factor was defined:

If

PERFORMANCE FACTOR = P

and t_n = Ton time, ta = 1-2-3T time

t_{BCD} = best case data = 28 seconds for both tasks

{all values average over six measurements}

TABLE 5.5:

VARIABLE RESOLUTION DATA FOR SUBJECT #1:

Frame Rate: 28 f/s

Grayscale: 4 bits

Case #1	Resolution	tn	ta	Overall Performance (P)
1	128x128	28(3)	28(1)	100%
2	64x128	31(6)	30(1)	92%
3	64x64	33(10)	32(15)	86%
4	32x64	48(12)	*	*
5	32x32	85(37)	*	*

*Not possible to accomplish task.

TABLE 5.6:

VARIABLE RESOLUTION DATA FOR SUBJECT #2:

Frame Rate: 28 f/s

Grayscale: 4 bits

Case #	Resolution	tn	ta	Overall Performance
1	128x128	28(5)	28(3)	100%
2	64x128	30(6)	30(4)	93%
3	64x64	32(10)	32(6)	88%
4	32x64	50(10)	*	*
5	32x32	90(42)	*	*

*Not possible to accomplish task.

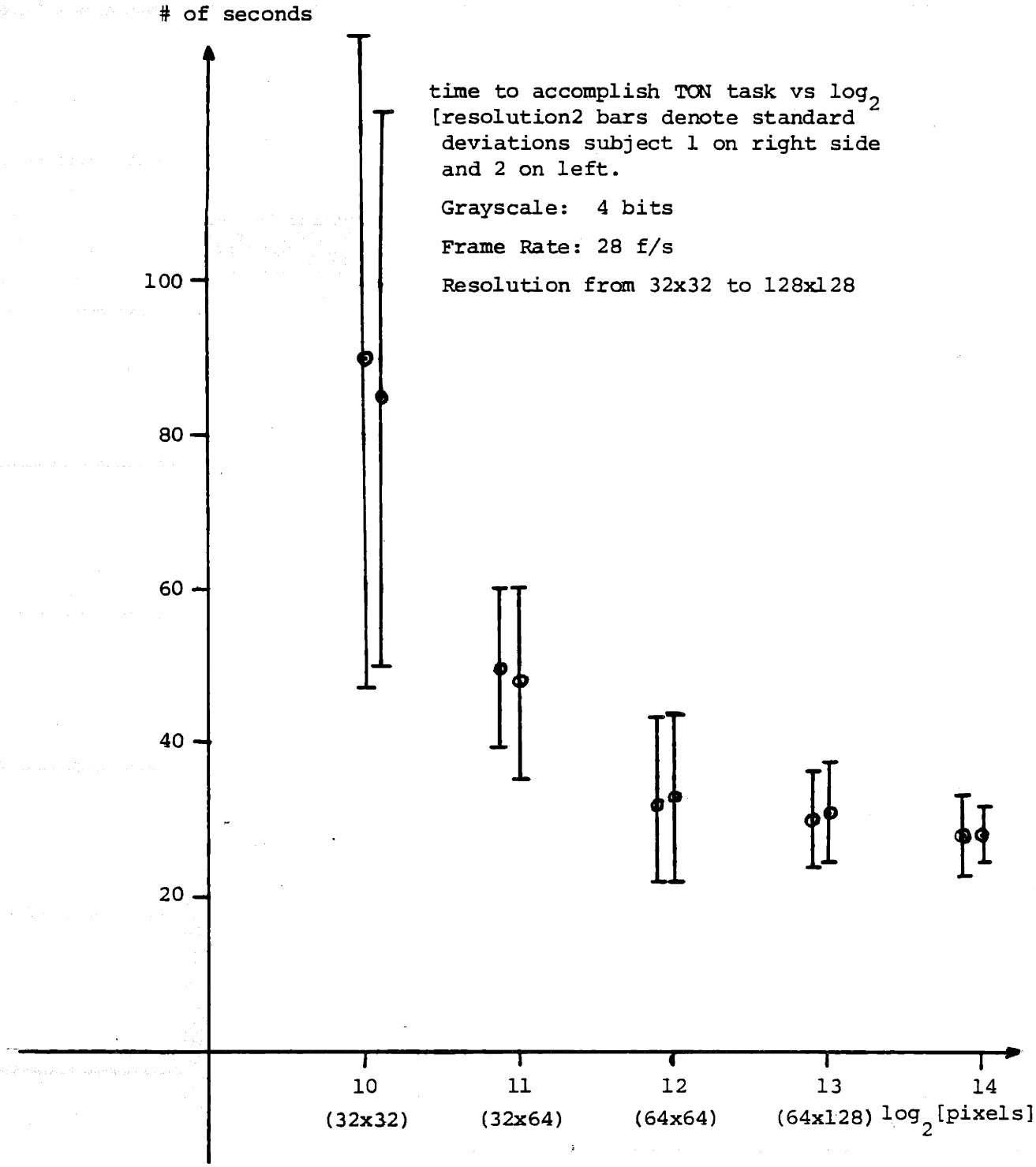


FIGURE 5.16: TIME TO ACCOMPLISH TON TASK VS LOG_2 OF RESOLUTION

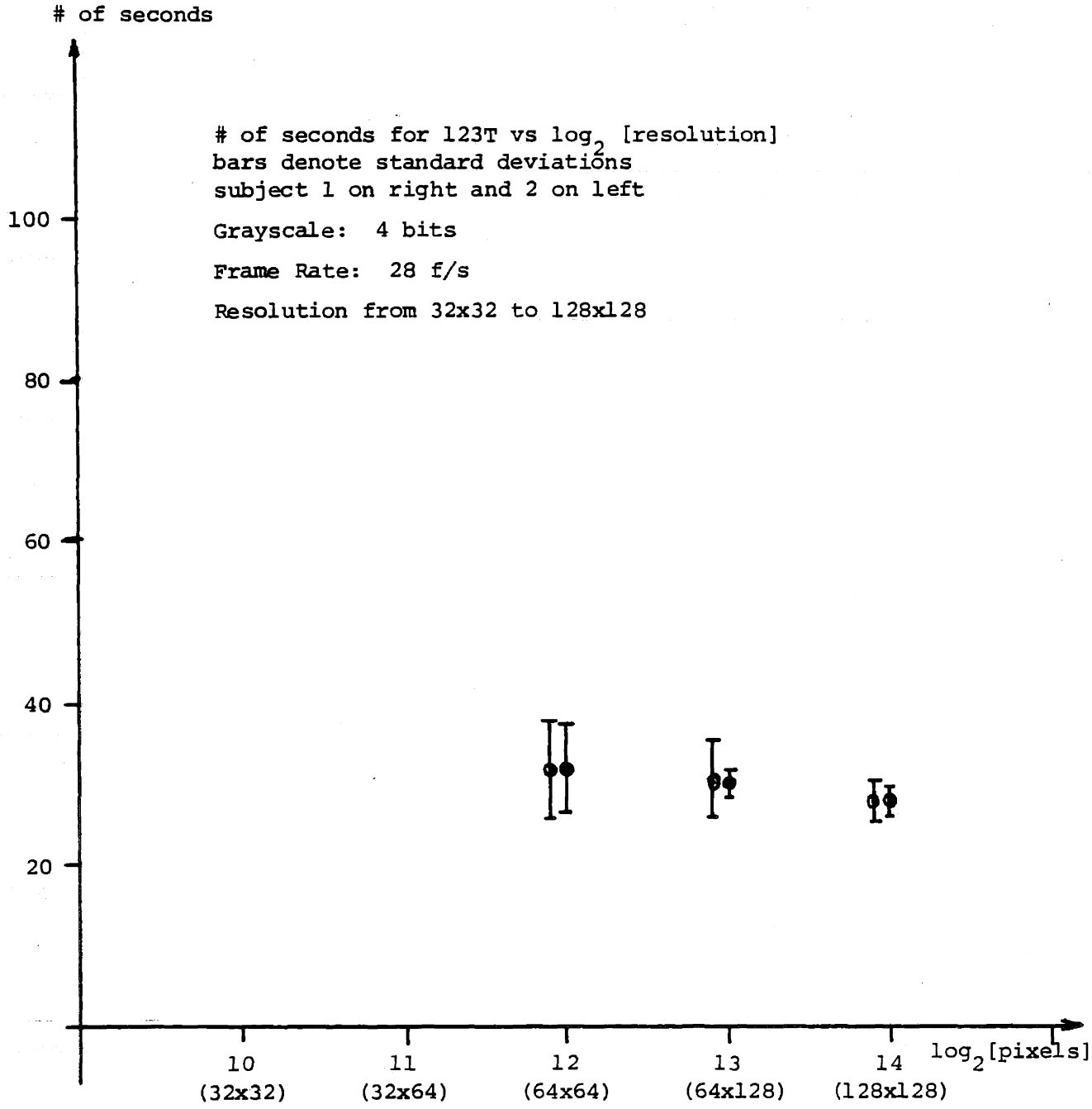


FIGURE 5.17: NUMBER OF SECONDS FOR 123T Vs LOG_2 RESOLUTION

was measured for the various resolutions possible.

The experimental procedure was kept the same as described earlier. Tables 5.5 and 5.6 summarize the data for a frame rate of 28 frames/second and for 4 bits of grayscale. The resolution was not always symmetric ($\ell x \ell$) but was sometimes ($\ell x \ell/2$) as seen in the table. The 1-2-3 task could not be performed for resolutions below 64x64 since it was no longer possible to distinguish between the lines forming the squares into which the peg had to be placed. Figures 5.16 and 5.17 indicate this.

The experiment was then repeated with 2 bits of grayscale. It was attempted to also decrease the frame rate; however, this caused noise problems. For this reason the frame rate was kept at 28 f/s.

Once again the 1 - 2 - 3 task could not be performed for resolutions below 64 x 64. Tables 5.7 and 5.8 summarize the data for the two subjects. This is illustrated in Figures 5.18 and 5.19.

The same data is illustrated in terms of performance in Figure 5.20. Since it was not possible to accomplish the 123 task at resolutions below 64 x 64, the performance curve is for the TON task only.

The fact that the subjects were able to perform the nut removal tasks at resolutions as low as 32 x 32 is of special significance. The variances increase rapidly with decreasing resolution. This is because there is a certain element of luck involved in correctly positioning the manipulator. The process is, however, not entirely luck dependent: the data showing that high resolutions result in better

TABLE 5.7:

VARIABLE RESOLUTION DATA FOR SUBJECT #1:

Grayscale: 2 bits

Frame Rate: 28 f/s

Case #	Resolution		123( )	Overall Performance
1	128x128	23(3)	28(1)	100%
2	64x128	33(4)	28(3)	92%
3	64x64	41(8)	35(8)	74%
4	32x64	59(16)	*	*
5	32x32	105(60)	*	*

*Could not accomplish task.

TABLE 5.8:

VARIABLE RESOLUTION DATA FOR SUBJECT #2:

GRAYSCALE: 2 bits

FRAME RATE: 28 f/s

Case #	Resolution	(TON)	123T( )	Overall Performance
1	128x128	28(5)	28(3)	100%
2	64x128	34(6)	30(4)	88%
3	64x64	44(8)	38(8)	69%
4	32x64	60(10)	*	*
5	32x32	120(36)	*	*

*Could not accomplish.

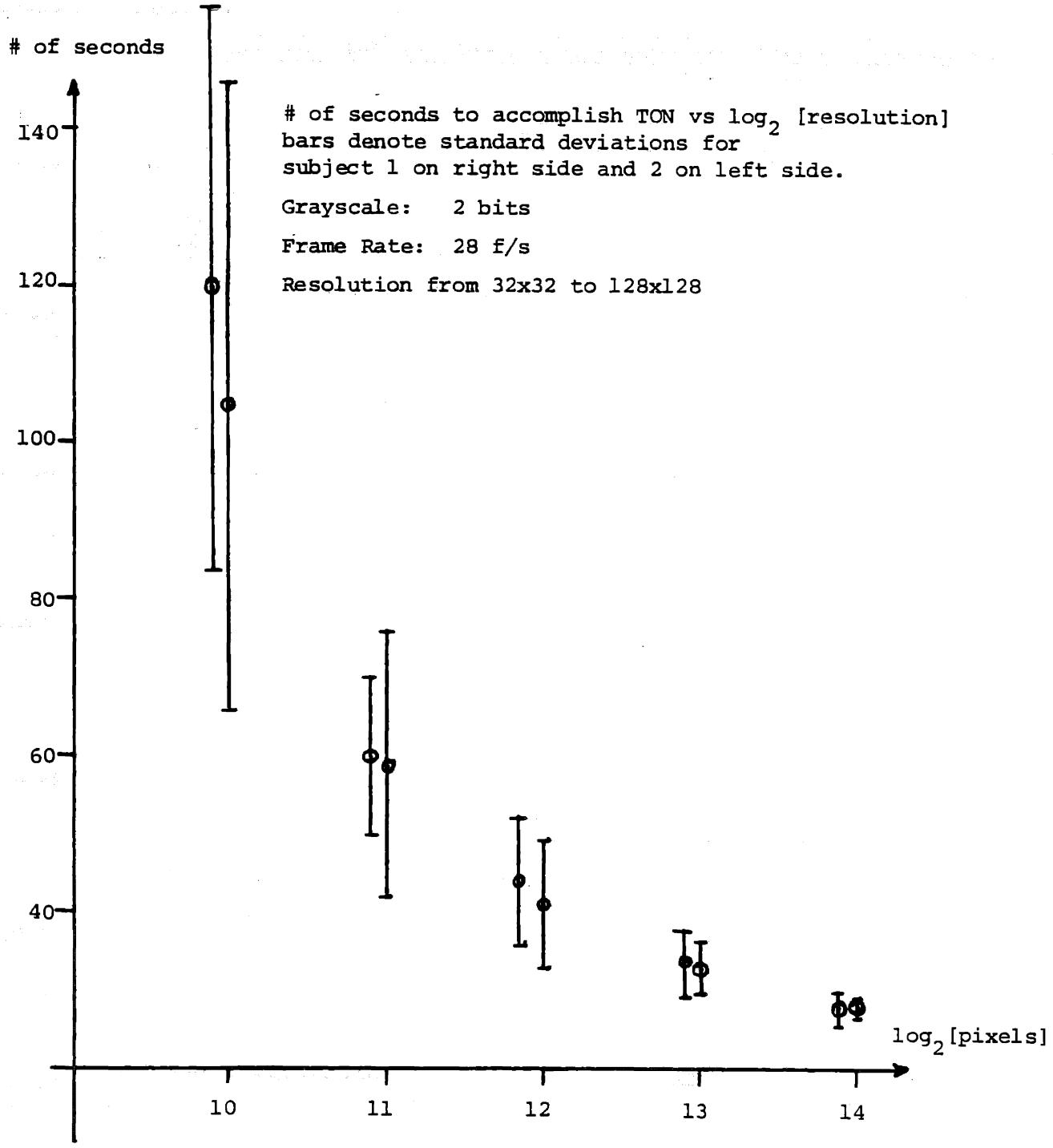


FIGURE 5.18: NUMBER OF SECONDS TO ACCOMPLISH TON VS LOG_2 OF RESOLUTION

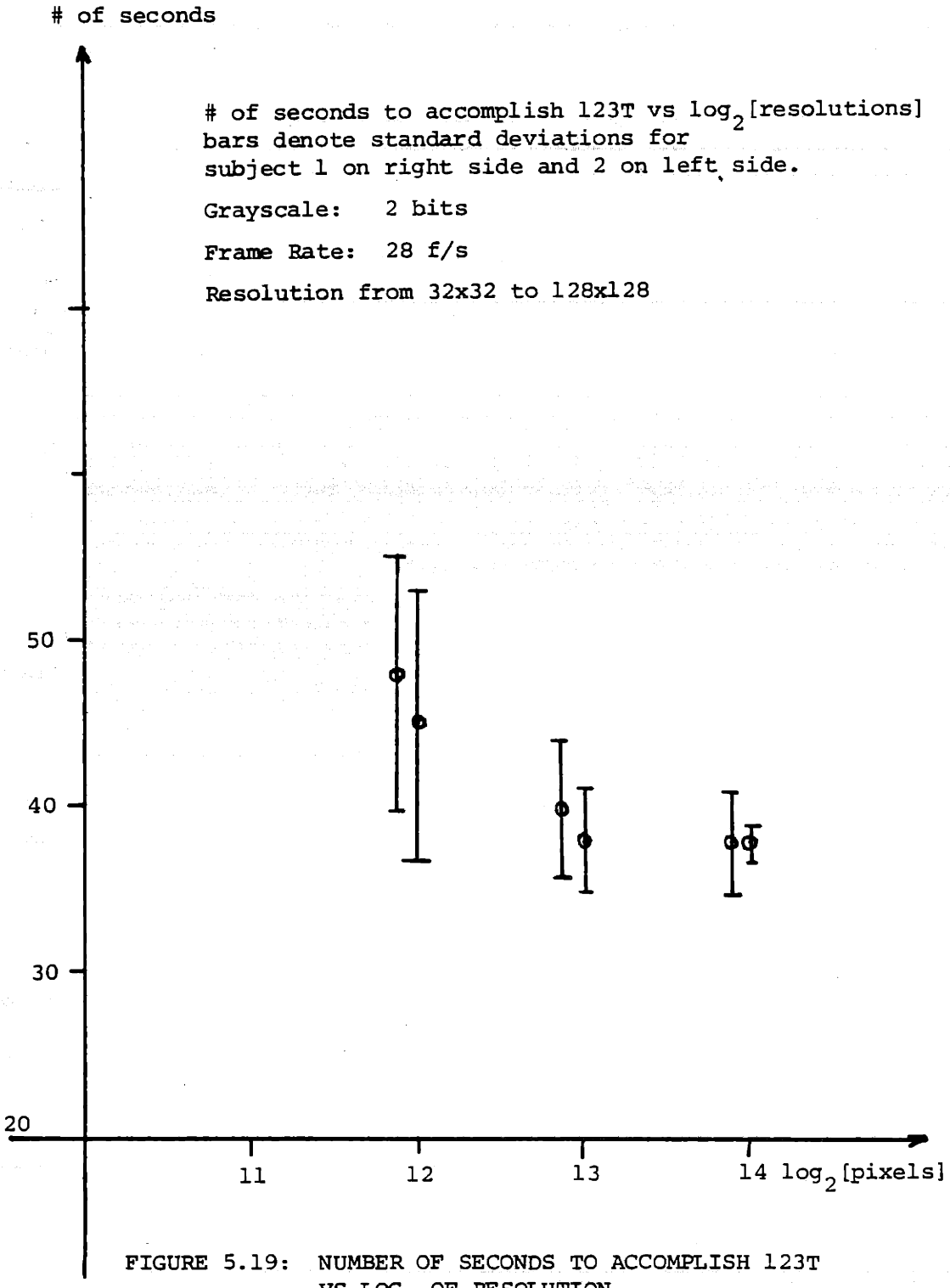


FIGURE 5.19: NUMBER OF SECONDS TO ACCOMPLISH 123T VS LOG_2 OF RESOLUTION

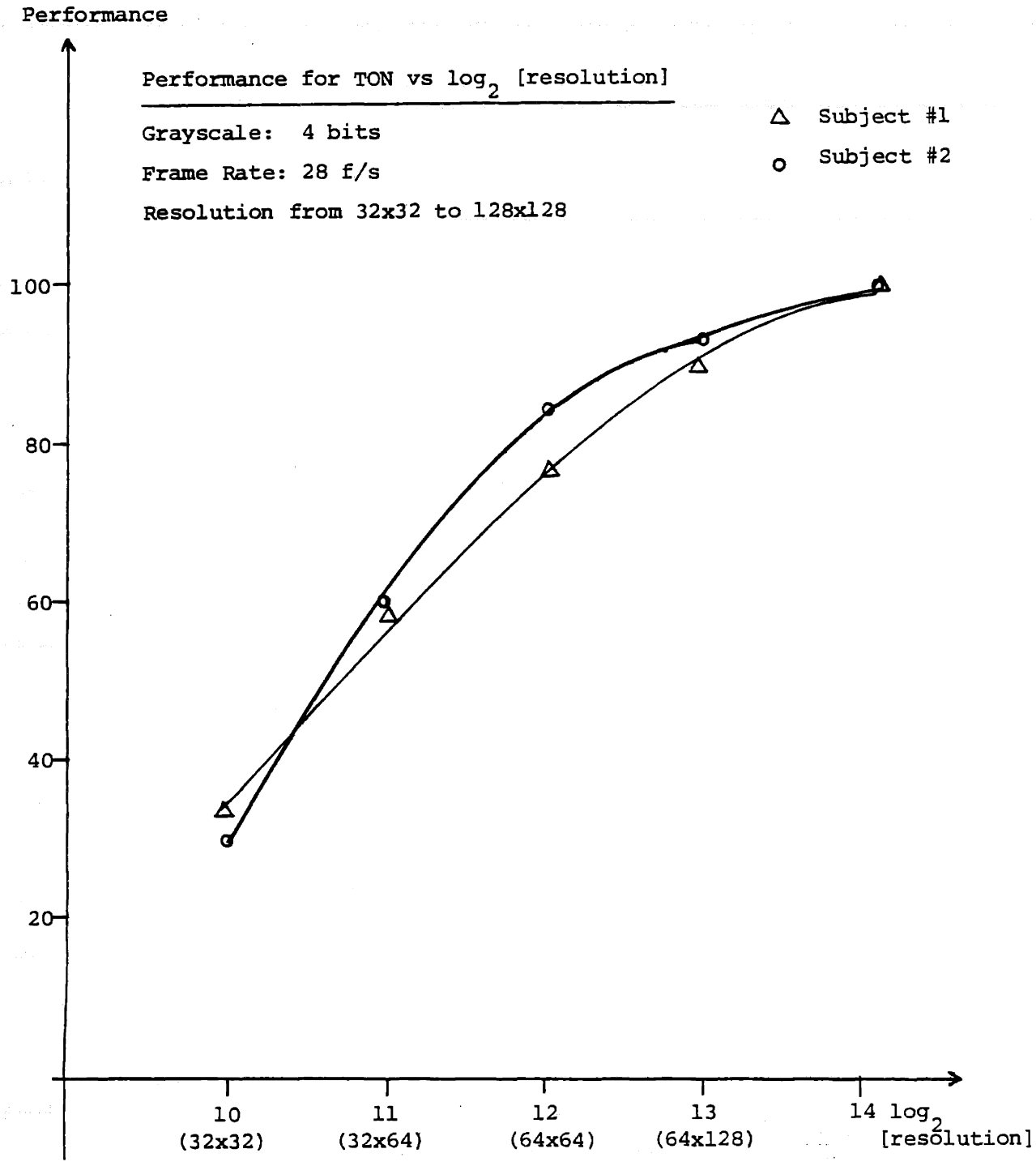


FIGURE 5.20: PERFORMANCE FOR TON VS LOG_2 OF RESOLUTION

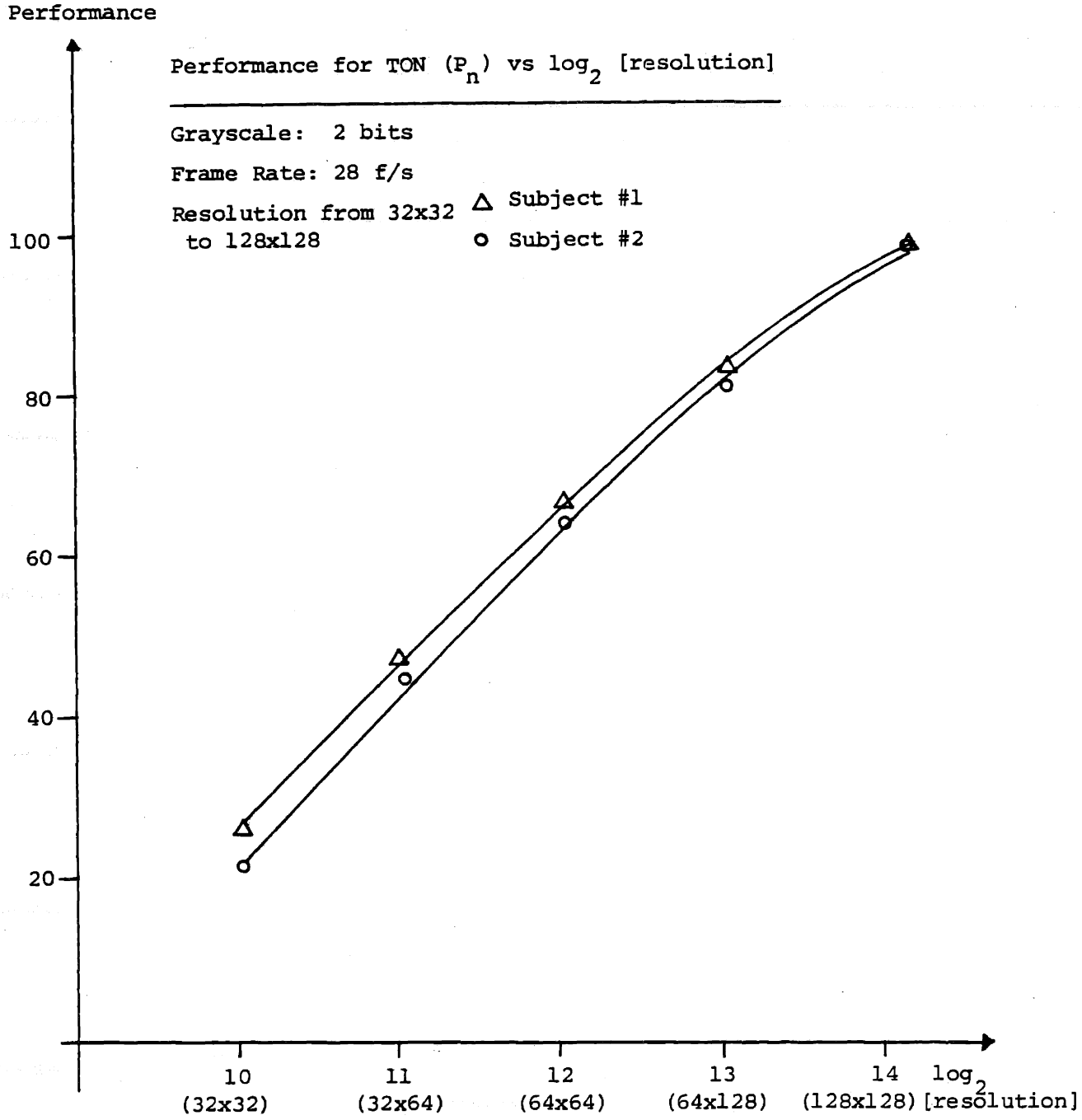


FIGURE 5.21: PERFORMANCE FOR TON (P_n) VS LOG_2 OF RESOLUTION

performance is testimony to this fact. As the resolution decreases, luck plays a greater role in the experiment. It is also noticed that 64 x 128 pixels result in considerably better performance than do 64x64 just as 64x32 result in better performance than do 32x32. This fact is resonant since it shows that picture symmetry is not as important as the total number of pixels.* Symmetry is not entirely unimportant since a resolution of 64x64 is preferred to 128x32 even though both have the same total number of pixels.

At lower resolutions, adequate operator training is important. A well trained operator can perform a useful task (once the manipulator is positioned) with little visual feedback.

It is apparent that a well trained operator can perform tasks familiar to him/her at even very low resolutions (32x32). If the task requires fine perception - such as the identification of a line, low resolutions cannot be used.

5.5.3 Variable Grayscale Case

Keeping the frame rates and resolutions constant, the grayscale was varied. The two different tasks were performed as before while performance was monitored and evaluated. Tables 5.9 and 5.10 tabulate the results for the two subjects when resolution was 128x128 and frame rate was 28 f/s.

---

*Naturally, this would not be true to say a picture with horizontal stripes!

These tables show that with 28 f/s frame rate and 128x128 resolution, grayscale can be reduced to the two bit level without affecting performance. Figures 5.22 and 5.23 illustrate the data in Tables 5.9 and 5.10.

The experiment was repeated for a frame rate of 14 f/s and the data is tabulated in Tables 5.11 and 5.12. With a lower frame rate, lowering the grayscale does affect performance.

5.6 *Discussion of Results

From the data gathered for these two specific tasks, it is apparent that the three parameters F, R and G could each be degraded keeping the other two constant, without much effect on performance up to a certain point where performance then degrades rapidly. Under the specified conditions: F = 28 f/s, R = 128 x 128, G = 4 bits, reducing the frame rate by a factor of 4 (2 bits) affects performance by only 20%. Similarly, reducing the grayscale alone by a factor of 2 bits performance by 25%. In the case of the resolution however, two bit reduction degrades performance for the TON task by 70%, while making the 123 task impossible to accomplish. It is especially useful to consider these facts in terms of the number of bits per second to be transmitted:

*Data in this section has been combined for the two subjects.

TABLE 5.9:

VARIABLE GRAYSCALE DATA FOR SUBJECT #1:

Resolution: 128x128

Frame Rate: 28 f/s

Case #	Grayscale	tn	ta	Overall Performance (P)
1	4	28(3)	28(3)	100%
2	3	28(2)	28(3)	100%
3	2	28(2)	28(4)	100%
4	1	42(5)	36(3)	74%

TABLE 5.10:

VARIABLE GRAYSCALE DATA FOR SUBJECT #2:

Resolution: 128x128

Frame Rate: 28 f/s

Case #	Grayscale	tn	ta	Overal Performance (P)
1	4	28(2)	28(4)	100%
2	3	29(3)	29(3)	100%
3	2	28(4)	28(4)	100%
4	1	50(6)	38(4)	65%

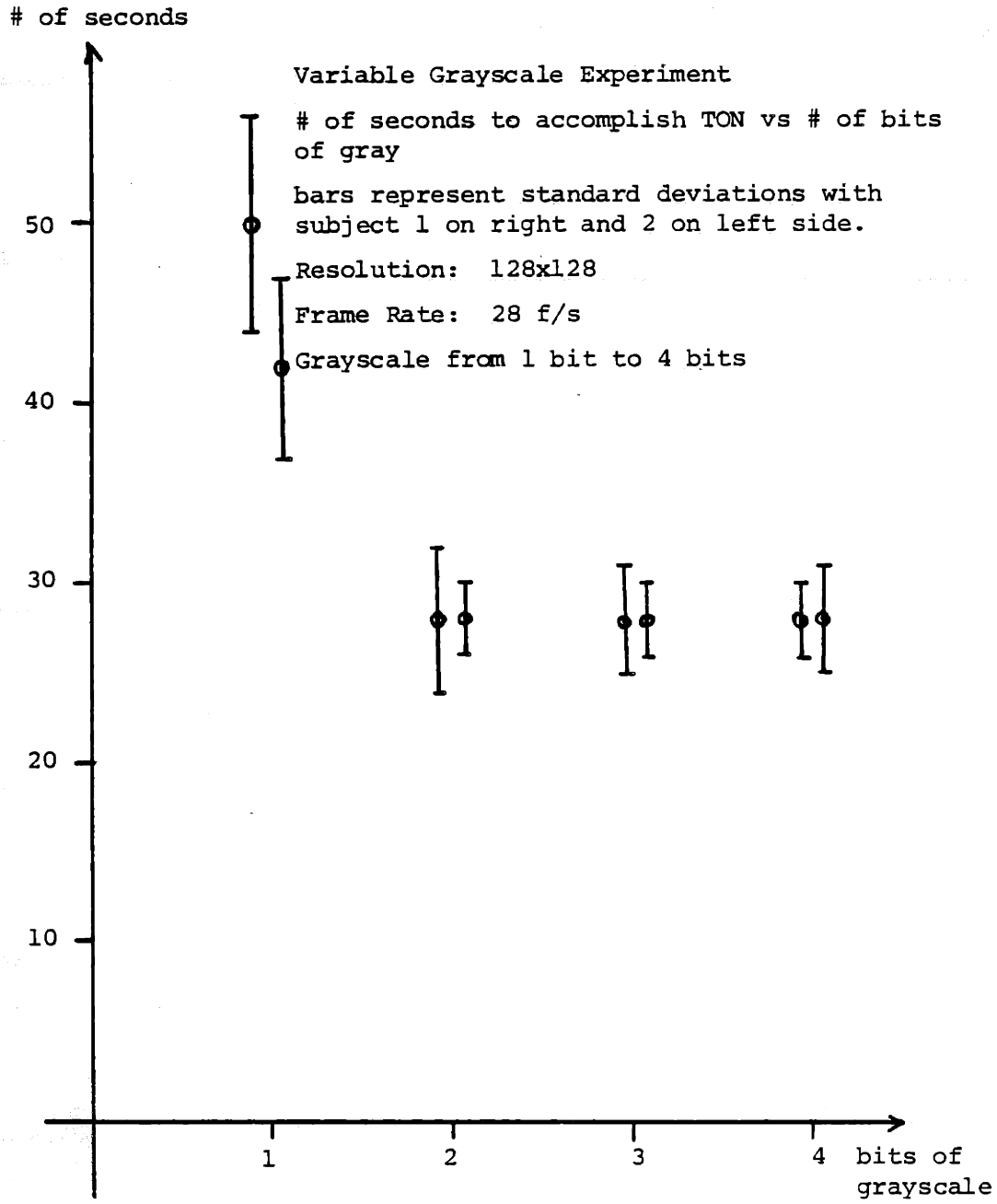


FIGURE 5.22: VARIABLE GRAYSCALE EXPERIMENT

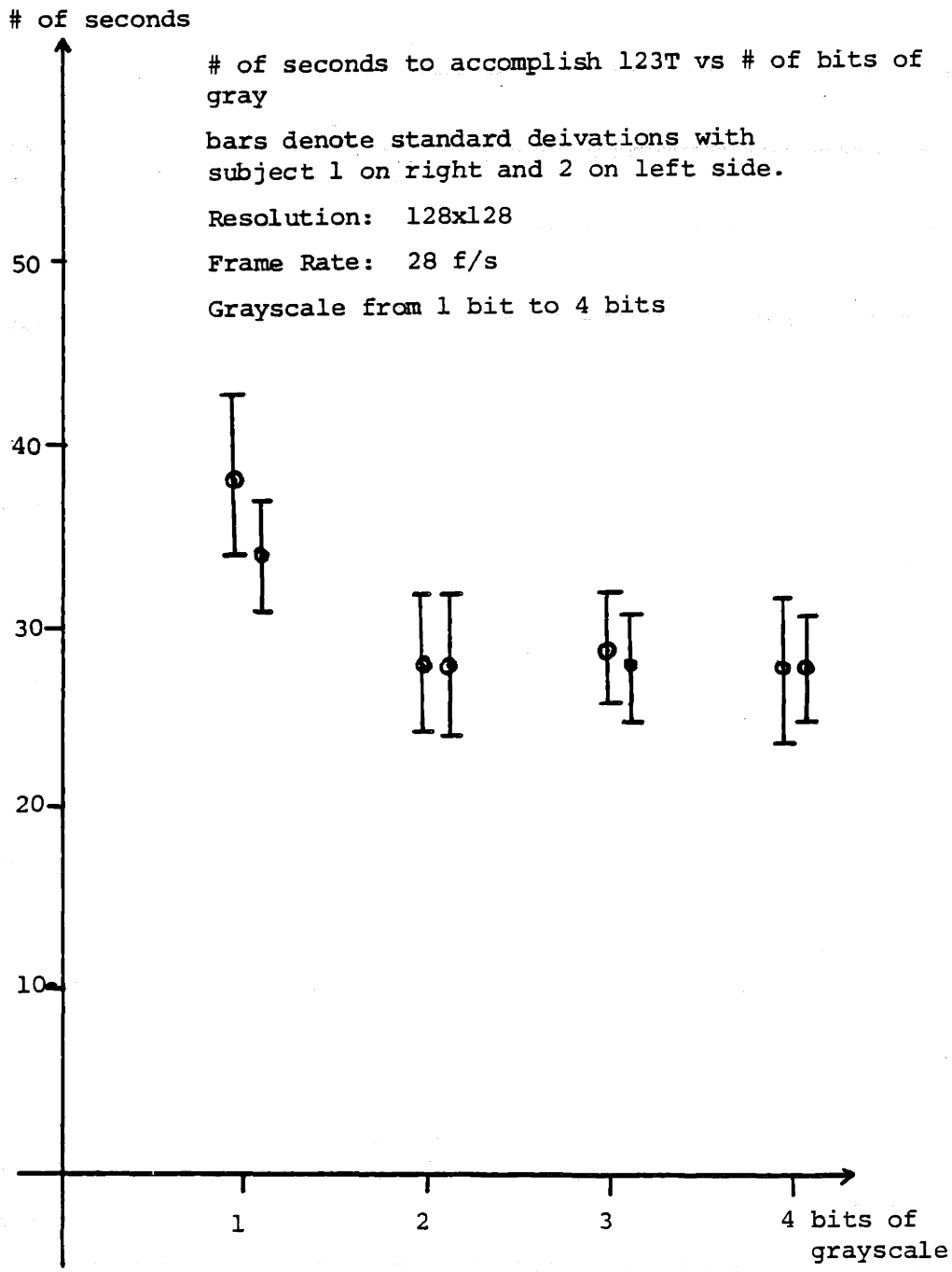


FIGURE 5.23: NUMBER OF SECONDS TO ACCOMPLISH 123T VS NUMBER OF BITS OF GRAY

TABLE 5.11:

VARIABLE GRAYSCALE DATA FOR SUBJECT #1:

RESOLUTION: 128x128

FRAME RATE: 14 f/s

Case #	Grayscale			Overall Performance
1	4	28(3)	28(1)	100%
2	3	32(8)	34(3)	85%
3	2	42(6)	36(4)	68%
4	1	53(10)	41(5)	60%

TABLE 5.12:

VARIABLE GRAYSCALE DATA FOR SUBJECT #2:

RESOLUTION: 128x128

FRAME RATE: 14 f/s

Case #	Grayscale			Overall Performance
1	4	29(2)	28(1)	100%
2	3	36(7)	39(4)	75%
3	2	46(8)	43(4)	63%
4	1	55(10)	49(8)	52%

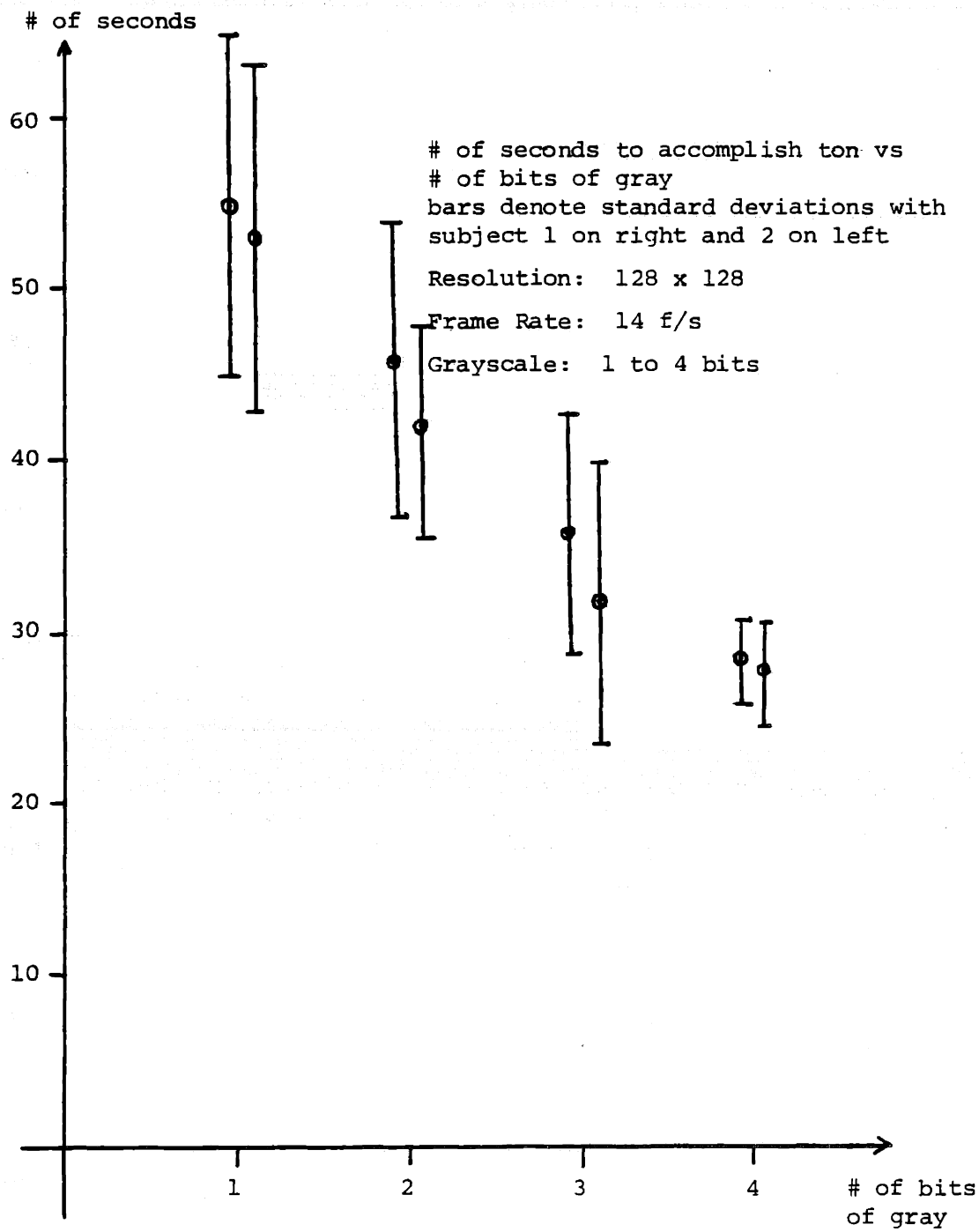


FIGURE 5.24: NUMBER OF SECONDS TO ACCOMPLISH TON VS NUMBER OF BITS OF GRAY

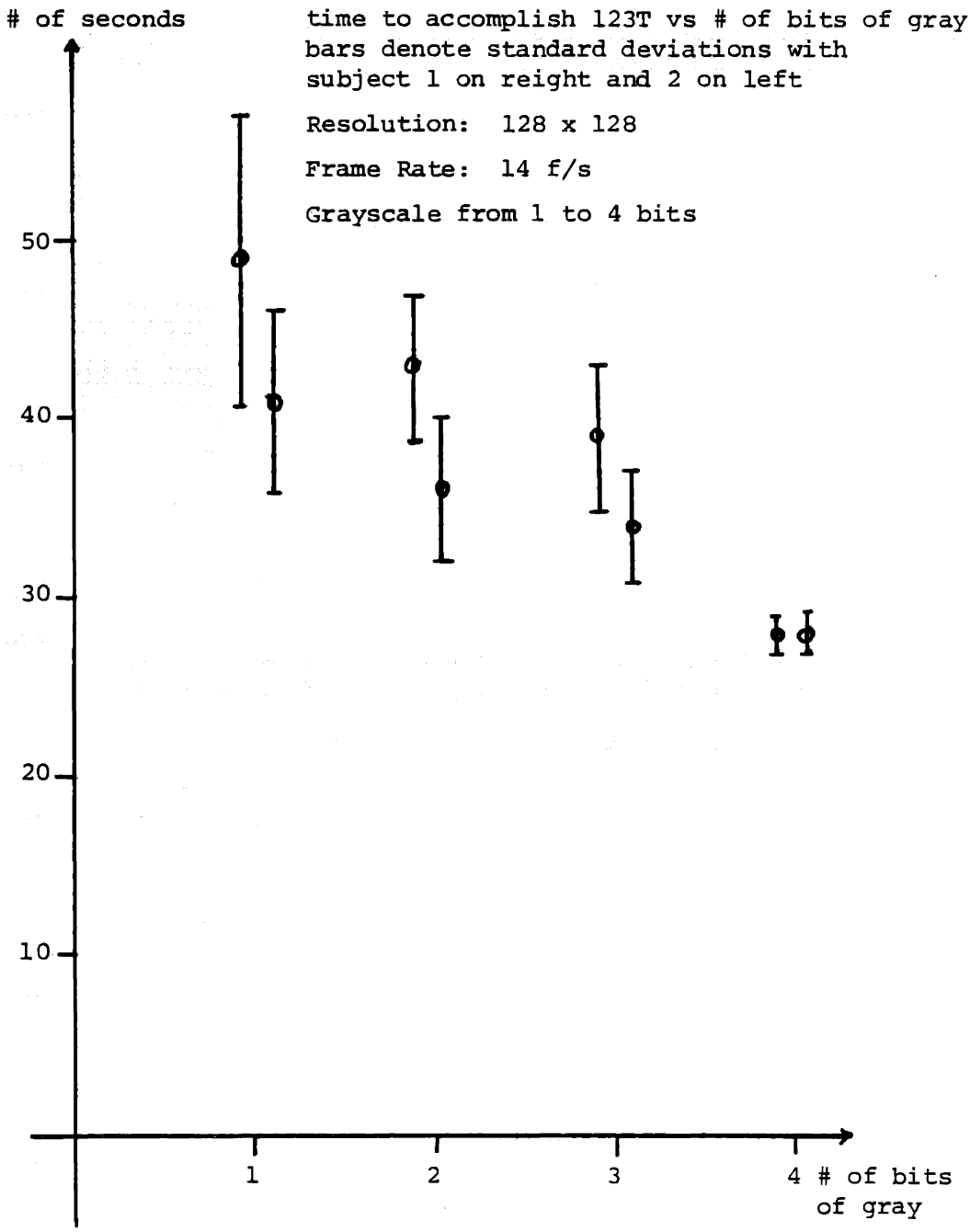


FIGURE 5.25: TIME TO ACCOMPLISH 123T VERSUS NUMBER OF BITS OF GRAY

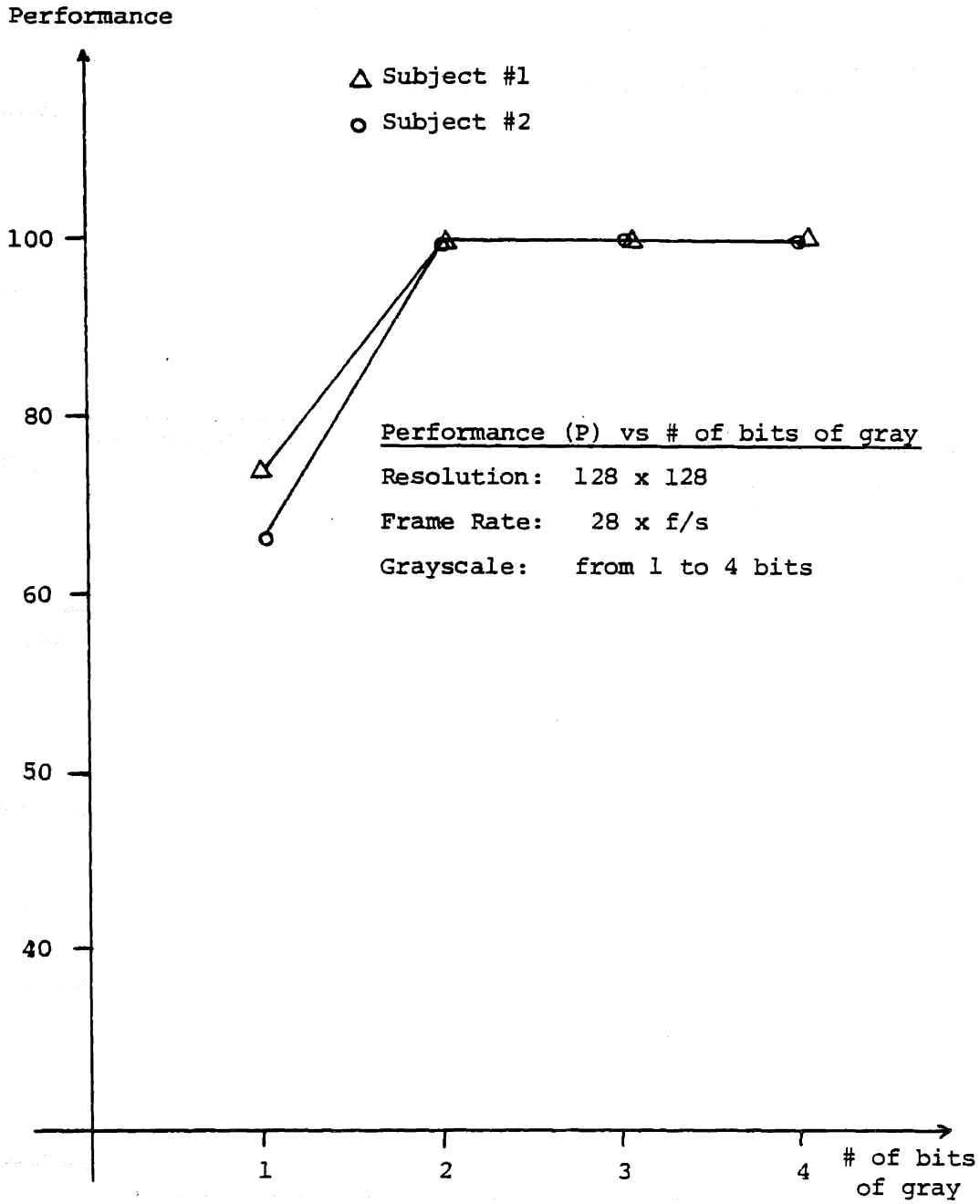


FIGURE 5.26: PERFORMANCE (P) vs NUMBER OF BITS OF GRAY

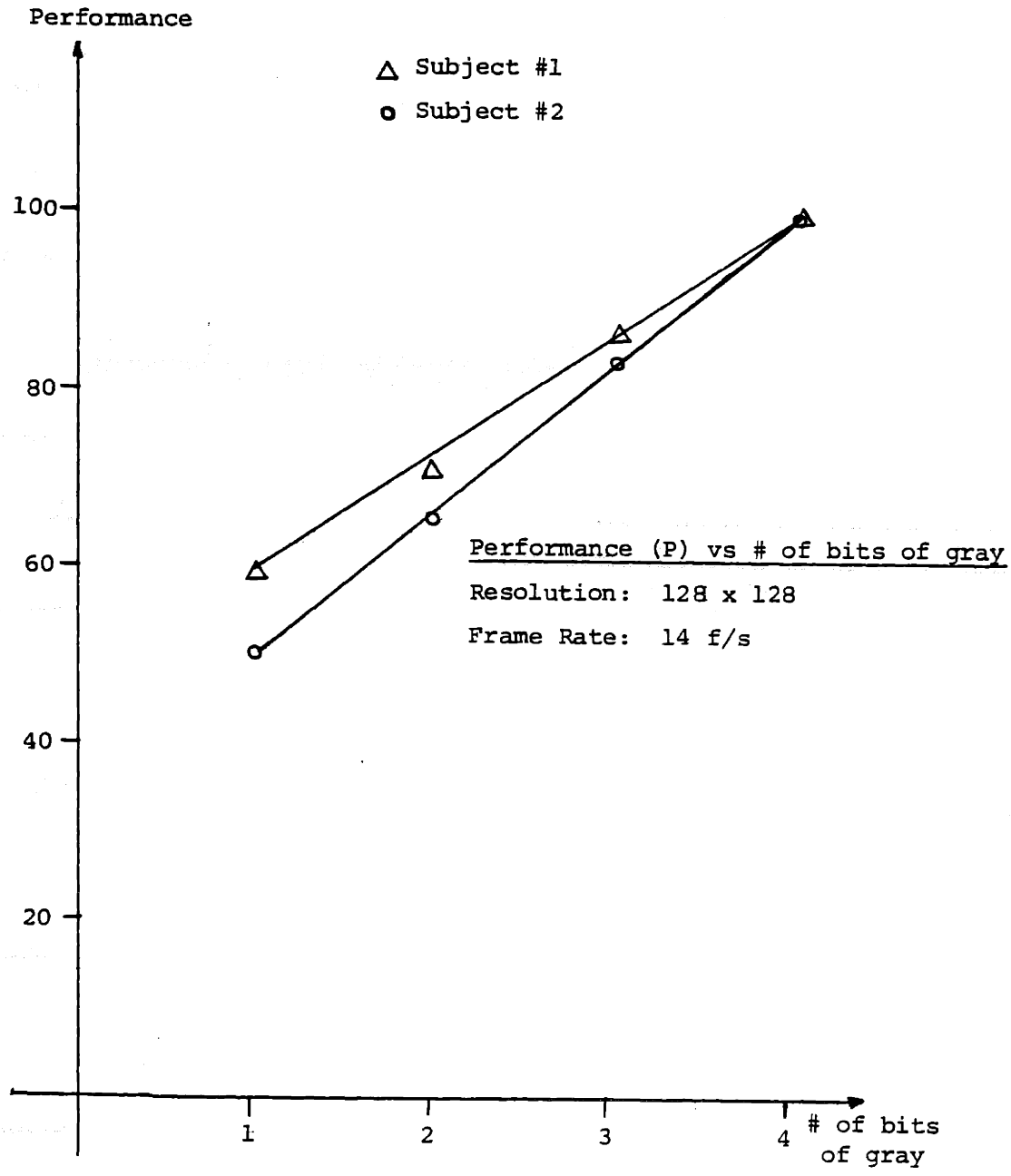


FIGURE 5.27: PERFORMANCE (P) VERSUS NUMBER OF BITS OF GRAY

FRAME RATE*	RESOLUTION	GRAYSCALE	PERFORMANCE**	# OF BITS/sec
28 f/s	128x128	4	100%	1,835,000
7 f/s	128x128	4	80%	458,750
28 f/s	128x128	1	75%	458,750
28 f/s	32x32	4	*33%	458,750

It is clear from this data that the number of bits per second could be kept the same and yet produce different performance for different combinations of F, R & G.

For the purpose of better illustrating the various trade-offs for these particular tasks, curves called "ISOPERFORMANCE CURVES" were constructed for combinations of F, R & G along which the performance is almost the same. Curves along which the information transmission rate is the same will be called ISOTRANSMISSION curves and were drawn for combinations of F, R and G along which the performance is almost the same.

5.6.1 Constant Resolution Isoperformance Curves

Along these curves, the resolution is the same. That is, the frame rate and grayscale are varied so as to keep the performance at a certain level. Consider the data:***

*Data combined for two subjects.

**-TON performance, P not defined.

***This data have been combined for the two subjects.

FRAME RATE	GRAYSCALE	TON	123T	PERFORMANCE
28 f/s	1	46	36	72%
	2			
	3			
5.6	4	33	34	76%

The third point necessary to get this curve can be estimated from the following data:

RESOLUTION - 128x128

FRAME RATE	GRAYSCALE	PERFORMANCE
14	1	60%
14	2	68%
14	3	85%
14	4	100%

Thus, the third point can be estimated since ~75% performance will result for 14 f/s when the grayscale is between 2 and 3 bits. This curve along with a isotransmission curve is shown in Figure 5.28.

Similarly, another isoperformance curve for constant resolution can be obtained with the data.

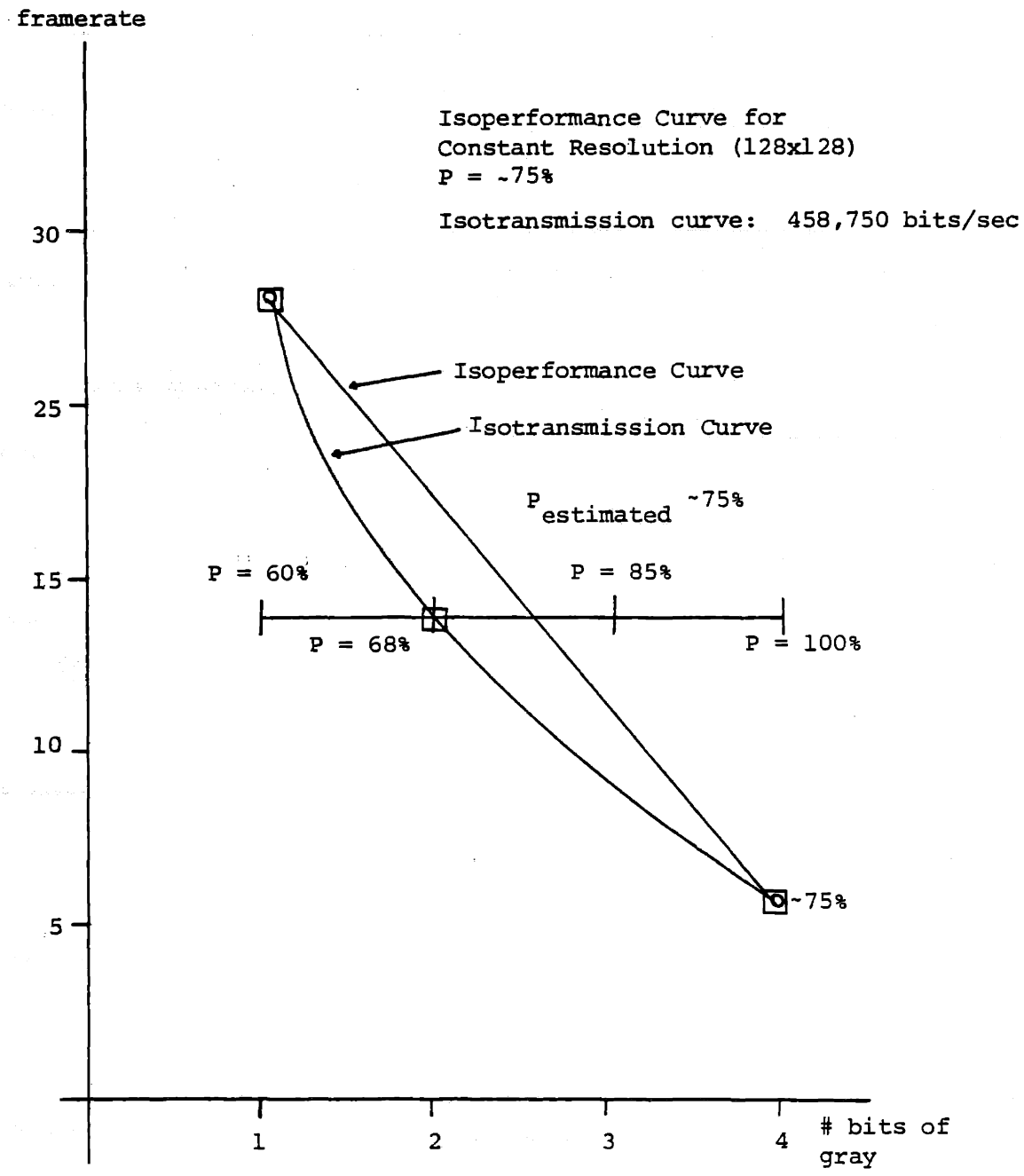


FIGURE 5.28: ISOPERFORMANCE CURVE FOR CONSTANT RESOLUTION (128x128)

RESOLUTION = 128x128

FRAME RATE	GRAYSCALE	TON	123T	PERFORMANCE
14 f/s	1	53	41	60%
	2			
	3			
2	4	52	42	60%

As before the third point can be estimated and is found to be between 2 and 3 bits of gray for about 4 frames/sec. This curve is shown in Figure 5.29.

The curves show that for the 75% performance curve, the extremities are the best points of operation. For the 60% curve however, the low framerates are more beneficial.

5.6.2 Constant Grayscale Isoperformance Curves

The grayscale was kept constant and the resolution and framerates varied so as to maintain performance at a constant level. The following data was used to plot isoperformance curves for a constant grayscale:

FRAME RATE	RESOLUTION	PERFORMANCE
28 f/s	64x64	~85%
14 f/s	64x128	~85%
9 f/s	128x128	~85%

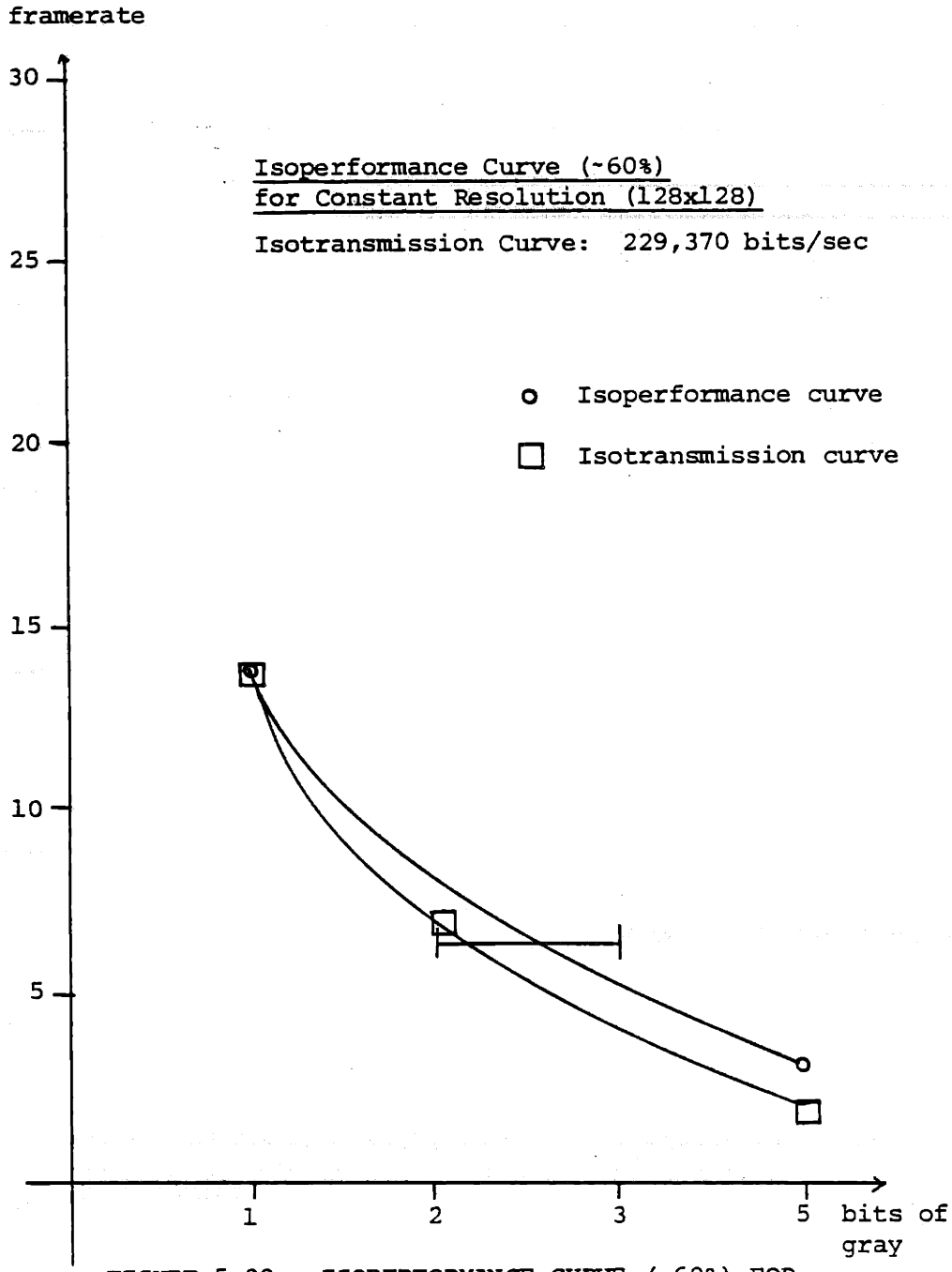


FIGURE 5.29: ISOPERFORMANCE CURVE (~60%) FOR CONSTANT RESOLUTION (128x128)

This curve along with its corresponding isotransmission curve is shown in Figure 5.30. Another isoperformance curve was derived from the data:

FRAME RATE	RESOLUTION	PERFORMANCE
14 f/s	64x64	~75%
9 f/s	64x128	~75%
9 f/s	128x128	~75%

This isoperformance curve and its corresponding isotransmission curves are shown in Figure 5.31.

Notice how closely the isoperformance curves follow the isotransmission curves. Thus, in this case there is a very strong correlation between resolution and framerate.

5.6.3 Constant Framerate Isoperformance Curve

The framerate was kept a constant and grayscale and resolution varied so as to maintain the same performance level. The specific points for this curve were more difficult to obtain owing to the discrete nature of the grayscale and resolution. The following data were used:

FRAMERATE: 28 f/s

CASE #	GRAYSCALE	RESOLUTION	PERFORMANCE
1	3 bits	64x64	~75%
2	2 bits	64x128	~90%
3	1 bit	64x128	~60%
4	1 bit	128x128	~75%

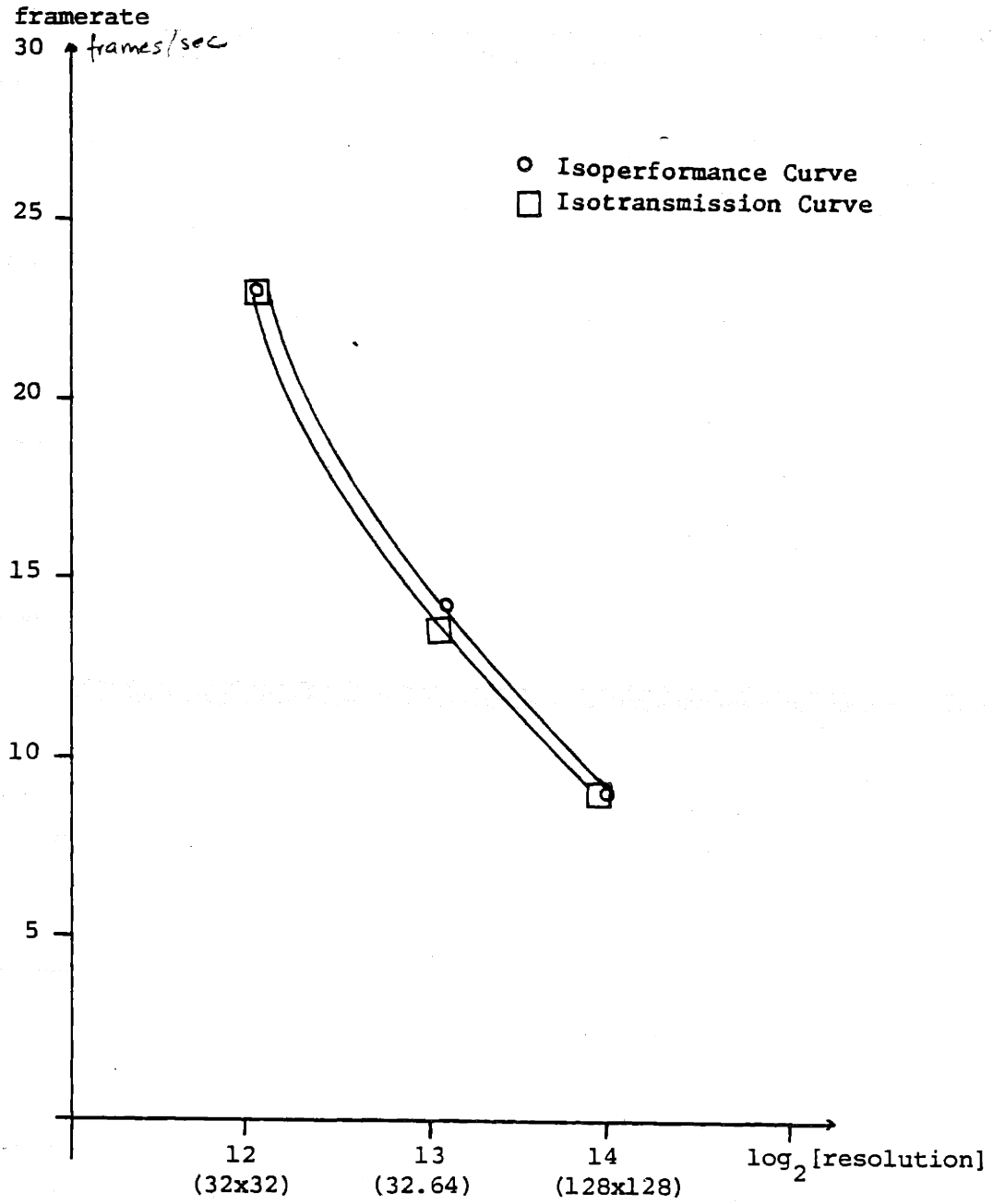


FIGURE 5.30: ISOPERFORMANCE CURVE (85%) FOR CONSTANT GRAYSCALE (4 bits)

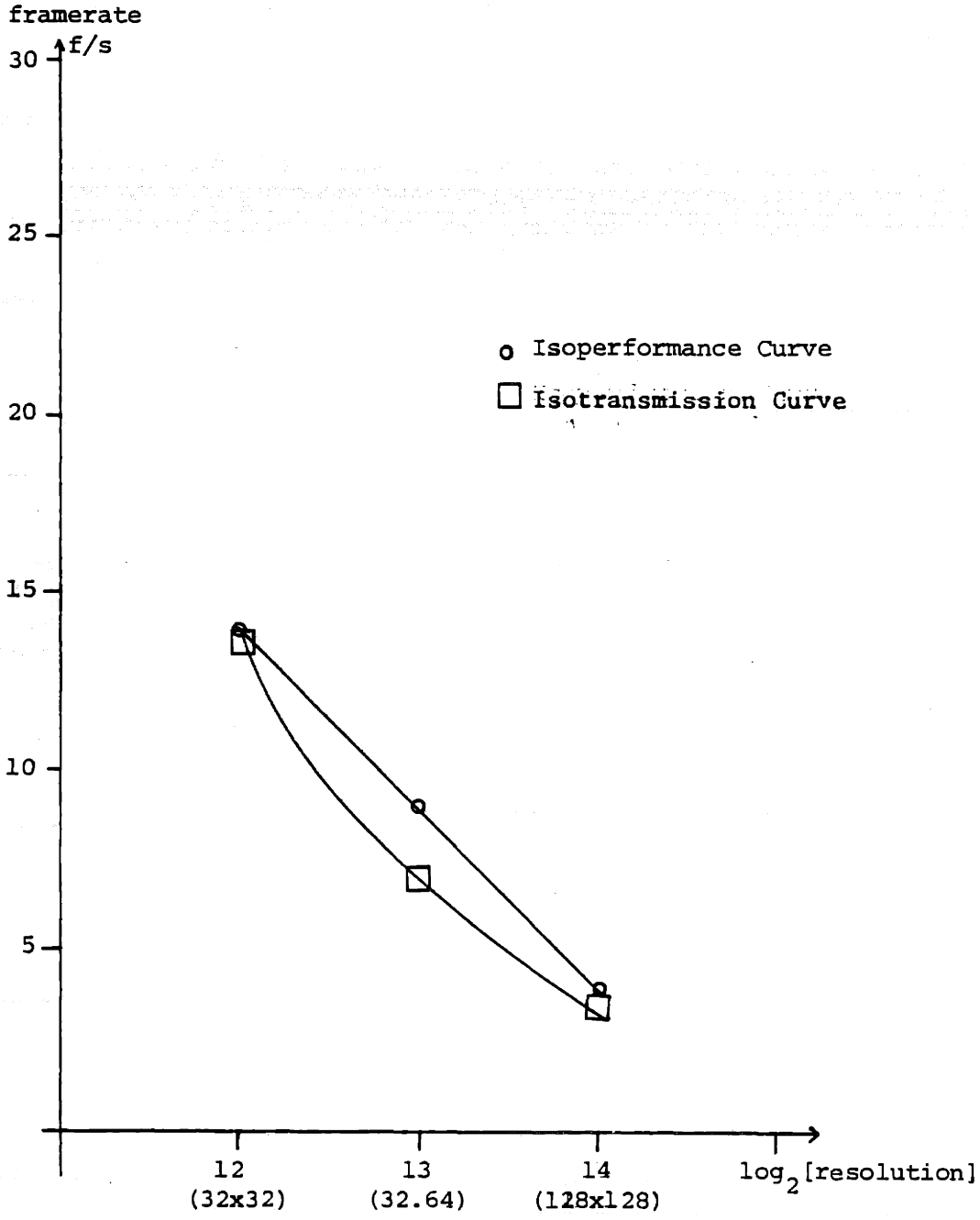


FIGURE 5.31: ISOPERFORMANCE CURVE (~75%) FOR CONSTANT GRAYSCALE (4 bits)

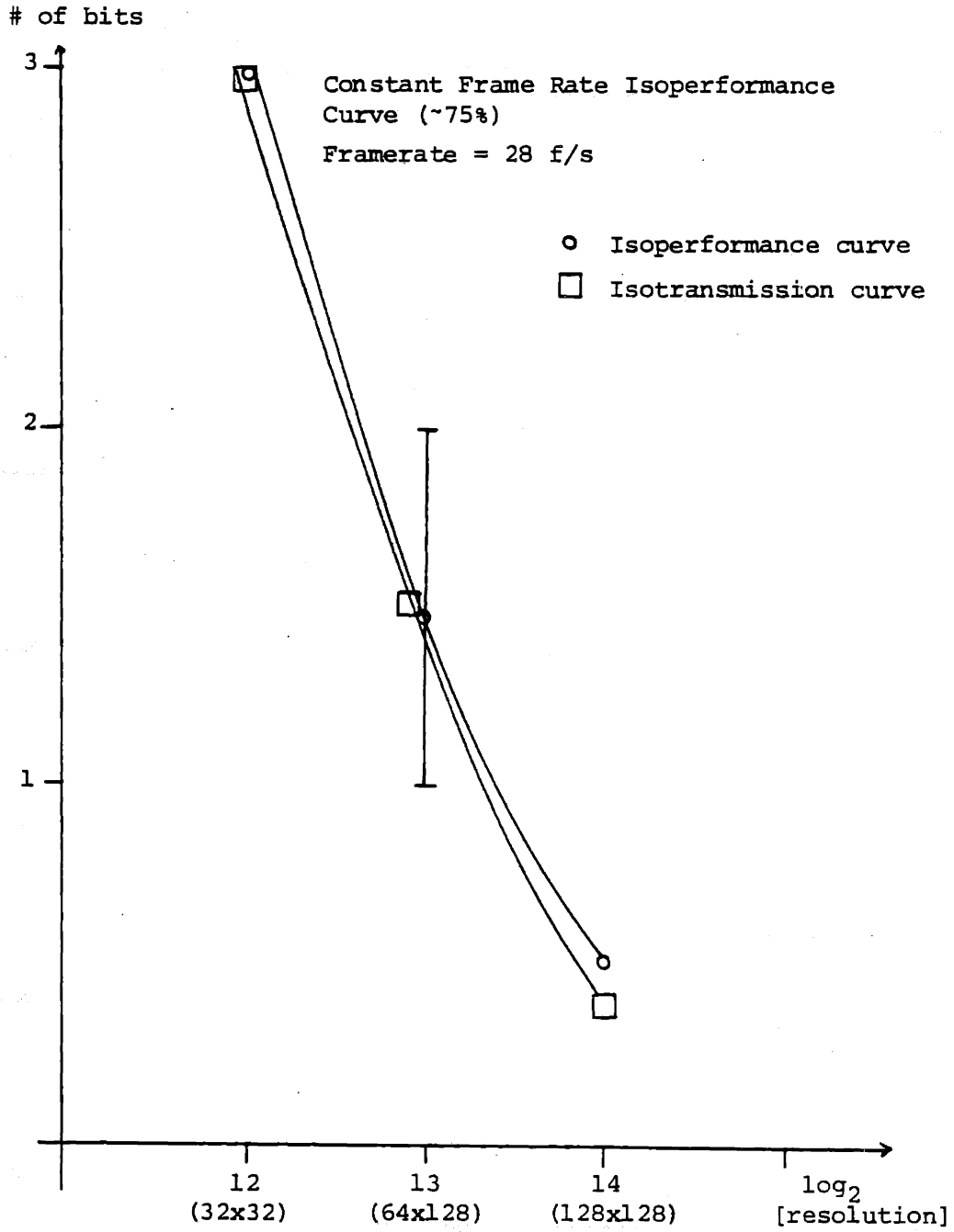


FIGURE 5.32: CONSTANT FRAME RATE ISOPERFORMANCE CURVE (~75%)

from Case #2 and 3 the third point can be estimated since if resolution is kept at 64x128 the grayscale will be between 1 and 2 bits. This 75% isoperformance curve is shown in Figure 5.34 with its corresponding iso-transmission curve.

Notice from this figure how closely the two curves follow each other and in fact coincide for higher gray levels.

5.7 Noise Problems at Low Framerates

It was noticed that at low framerates there was more noise in the picture. As a result of this, the operator's task was additionally complicated with a slower picture. Having examined the hardware of system GREAF, it was clear that the noise problems were not because of the change coupled device in the system. It was therefore unclear whether:

(i) the noise was present in the slow images

(ii) there appeared to be more noise in the slow pictures

The eye is often thought of as a low pass filter. This is in fact the reason why pictures (television, movies...) are shown at no faster than 30 frames/second.

In the case of SYSTEM GREAF, the fastest sampling rate is 28 f/s. Each of these frames is made up of signal and noise. In each consecutive frame the signal is the same* but the noise changes. At the higher frequencies the visual nervous system averages the noise. At lower frame rates however, each frame is displayed several times. Thus,

*For a stationary picture.

the frequency of the noise is now much lower. For example, at 28 f/s the noise has a frequency of 28 hz. At 1 f/s, the frequency is 1 Hz. Owing to the low pass nature of the eye, the signal to noise ratio increases with frame rate.

MATHEMATICAL MODEL:

Assume that the human eye averages over a period of τ seconds.

Let n = number of frames/τ seconds

F_n = nth frame in a particular τ period

S_n = signal in nth frame all 1's white

W_n = noise in nth frame all 0's black

$F_n = S_n + W_n(\sigma_n^2)$ where W_n^2 = variance of noise. For the τ seconds period the variance will be σ_n^2/n since there were assumed to be n frames/τ seconds. Signal to Noise Ratio (SNR) = $\frac{\text{variance of signal}}{\text{variance of noise}}$

as $\frac{\sigma_n^2}{n}$ decreases, SNR increases

and $\frac{\sigma_n^2}{n}$ decreases as n increases

Thus, for a higher n (i.e., for a higher frame rate) there will be a higher SNR so that there appears to be less noise in the picture. In other words, at slow frame rates, the perceptual system "forgets" the information between frames. At high frame rates the signal is constant but the noise changes from frame to frame. When successive frames are averaged (integrated) the signal will appear to show out more distinctively.

CHAPTER VI

CONCLUSIONS

The first conclusion is that human operations can (much to their own disbelief) perform fairly complicated tasks familiar to them with a course, intermittent digitized picture requiring as little as 50,000 bits/second.

Furthermore, for a picture at 128x128 resolution, 28 f/s frame rate and 4 bits of grayscale, each of these three can be lowered considerably individually, without preventing the operator from accomplishing his/her task.

In the range of operation*, frame rate and grayscale could be degraded by greater factors than resolution before making task accomplishment impossible.

For the given tasks and manipulator threshold points exist for all three parameters:

For FRAME RATE: 3 f/s when resolution=128x128, grayscale= 4 bits

For RESOLUTION: 64x64 when frame rate = 28 f/s, grayscale = 4 bits

For GRAYSCALE: 1 bit

Any further degradation of a parameter beyond these points (while holding the other two constant) results in degradation in performance such that

---

*R = 128x128, F = 28 f/s, g = 4 bits

the task cannot be completed.

It was also observed that lowering the sampling rate created more problems than making the display slower. There appeared to be more noise in images at low frame rates. This was believed to be because of the low pass nature of the visual nervous system.

CHAPTER VII

INFERENCES AND RECOMMENDATIONS FOR ADDITIONAL WORK

From the data it is apparent that certain F, R and G combinations result in better performance than others even for the same transmission rate. It is also believed that some F, R, G combinations are better for certain tasks than they are for others. It is therefore the case that in some instances different permutations of F, R, G result in roughly the same performance. This is fortunate since for a fixed bandwidth transmission system it would enable the operator to have some choice of F, R and G. Typically, for a task requiring several subtasks, the operator might find it to his/her advantage to have different F, R, G settings for the different subtasks. For example, a task consisting of discovery, identification and "perform job" is made up of three subtasks:

SUBTASK 1: DISCOVERY

This could require the operator to spot a certain object while ranging widely over the terrain.

SUBTASK 2: IDENTIFICATION

The operator has to correctly identify the specific data is of object and categorize. This categorization could lead to further actions.

SUBTASK 3: PERFORM JOB

Based on the information from the previous subtask, the operator might have to do a specific manipulation involving dynamic visual feedback.

This task (made up to these three subtasks) is fairly representative of undersea teleoperation and any task could be similarly decomposed into subtasks.

Of special interest is the kind of visual feedback necessary in each case. For SUBTASK 1, a high framerate is necessary since this would enable the operator to quickly scan through the particular scenario. Since the operator is merely to "spot" a certain formation, resolution and grayscale might not be crucial.

On having "spotted" the object of concern, further identification would require a high resolution picture. Since the picture is now stationary, the framerate can be lowered and emphasis placed on grayscale and resolution.

If the job that needs to be performed following its categorization is very familiar to the operator than any combination of F, R and G could be accomplished fairly rapidly. On the other hand, if the job is new to the operator then a high resolution picture might be necessary for accuracy, or a high frame rate for dynamic control.

Thus, this tasks consisting of three subtasks needs different ratios of F, R and G for each subtask.

The veracity of this notion could be tested by the design of a simple lab experiment. Unfortumately, this experiment could not be performed owing to severe time constraints, but is recommended for future study.

RECOMMENDATIONS:

A subject's task might be to locate a number on a background wall. The subject has then to read the number and proceed to take-off-a-nut from the task hub. The nut must then be placed in the pail with that number on it.

This is a complete task consisting of three subtasks. Locating the number on the wall could be thought of as SUBTASK 1. Reading it is SUBTASK 2 and performing the subsequent operations is* SUBTASK 3. The experiment should be performed in two situations:

SITUATION 1: In this situation the operator has to perform the experiment with constant F, R, G combination.

SITUATION 2: This time the subject has at his/her disposal three different F, R, G** combinations. The operator has only to mention which of these combinations is required and the experimenter would make adjustments to bring about the required transition during the task.

Performing the experiment in situations 1 and 2, "total time to accomplish task successfully" could be measured as a gage of performance. The time required for transitions in SITUATION 2 should be deduced from the total time.

---

*Subtask 3 could itself be further divided.

**The combinations should require the same transmission rate.

A more advanced version of this experiment could be performed by attaching the computer to system GREAF such that the computer can be used to bring about transitions (in F R G settings) while keeping the transmission rate at a constant.

The outcome of this study MIGHT justify the need for an image transmission system in which it is possible to adjust the F, R and G as required.

APPENDIX A: SYSTEM GREAF CAMERA

The camera used was a commerically available charge coupled device from General Electric: TN2200 [3]. It is a totally solid state device which translates an optical image into a precise electrical signal. The height to width aspect ratio is 1.1. The C.I.D. microsensor contains not only the 16,384 pixels (picture elements) of the 128x128 TN200, but all of the circuit logic necessary to perform a sequential raster scan to generate synchronization signals. The C.I.D. arrays are fabricated as a silicon P-MDS device similar in many respects to some microprocessor and memory arrays. The cameras require only ± 15 volt power and a 0 to 5 volt clock signal to produce nominal 1 volt (peak-to-peak) analog video signals and C-MOS level synchronization signals.

TECHNICAL SPECIFICATIONS: TN2200

ELECTRICAL	INPUTS
Power Source	± 15V at 50ma
Clock Source	0 to 5 volt (TTL) clock signal. Must be in high state for 50 ms minimum.
·Clock Frequency	50 kHz to 5MHz

Even though this is the manufacturer's specification, experiments showed that at too low frequencies there is excessive noise integration. This is a characteristic of charge coupled devices since the charge flowing in and out of "charge wells" does so at a slow rate. The noise from each flow transaction builds up and is more visible at lower frequencies.

ELECTRICAL - OUTPUTS

Video	0 to 1 volt(nominal) sampled and held analog levels, white positive.
	.8 to 1.5 volts at white (saturation)
synchronization	C-MOS signal levels 0 volts false,
	+15 volts true

MECHANICAL

Weight:	Fixed Head Cameras	18 oz (510 gr)
	Remotable Head Cameras	21 oz (596 gr)
	Remote Head Only	8 oz (227 gr)
	Remote Body Only	13 oz (396 gr)

Lens Mount:	Standard "C" Mount
	1 inch (25.4 mm) 32 Thread
Camera Triped Mount:	1/4 x 20 Tapped Hole
Connector:	Dust Tight Aluminum Extrusion

ENVIRONMENTAL

Ambient Temperatures:	0° to 50°C (32° to 123°F)
Altitude:	50,000f (15,240m)

Signal-to-Noise Ratio

Temporal SNR	256:1 typical
Spatial SNR	25:1 typical

Blemishes

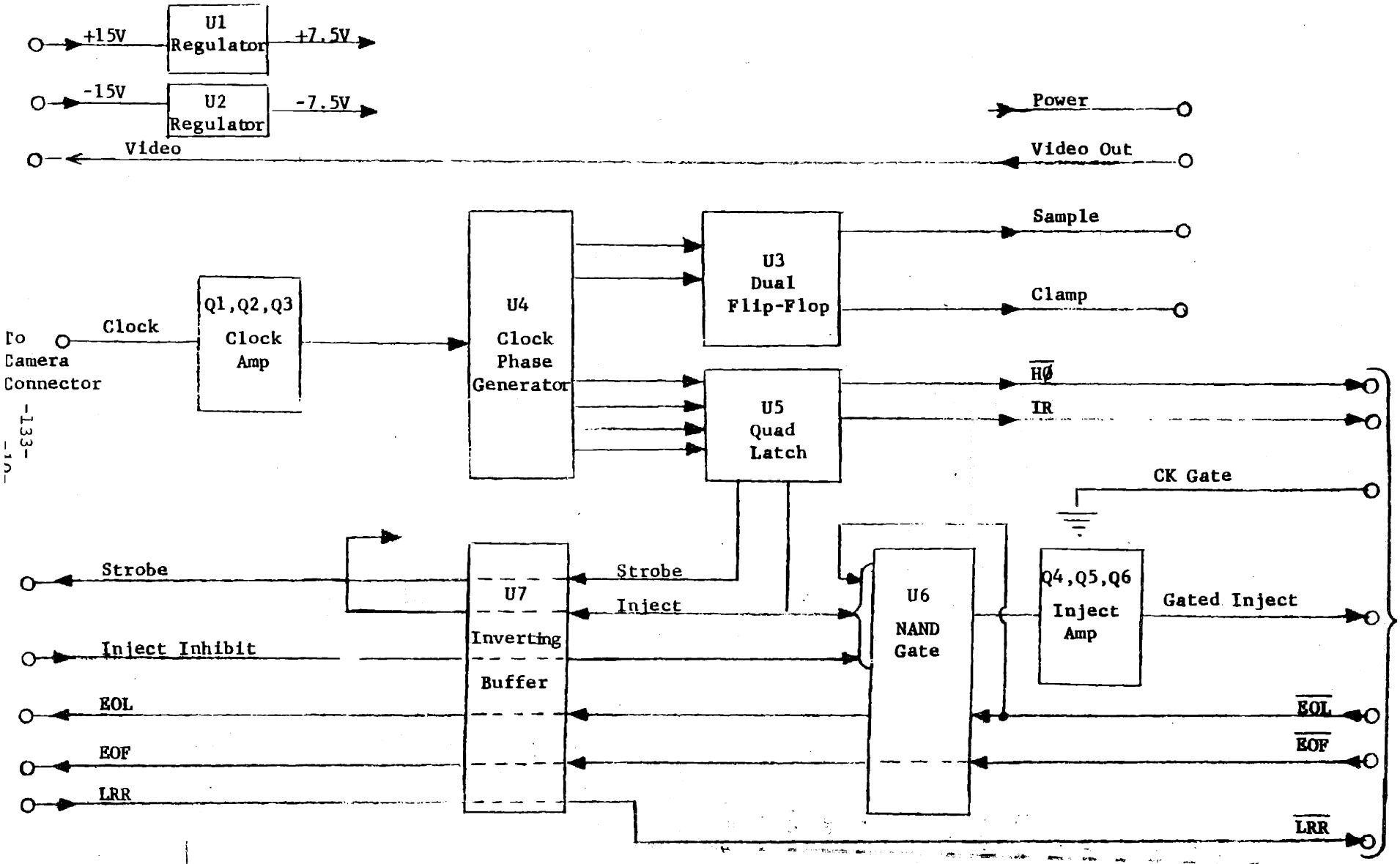
Maximum 8 black or white pixels at 30 frames/second 25°C (77°F) ambient.

Dark Current

Silicon sensors such as the TN2200 are subject to a doubling in leakage current for each 9-10°C size in temperature. Typically, the TN2200 will reach one-half saturation from leakage current in 1 to 2 seconds. For very slow operation, or operation above room temperature, the constant offset of dark current should be taken into account.

Figures A-1 throught A-5 are various diagrams of relevance. For more information refer to [3].

FIGURE A.1: SYNC GENERATOR
BLOCK DIAGRAM



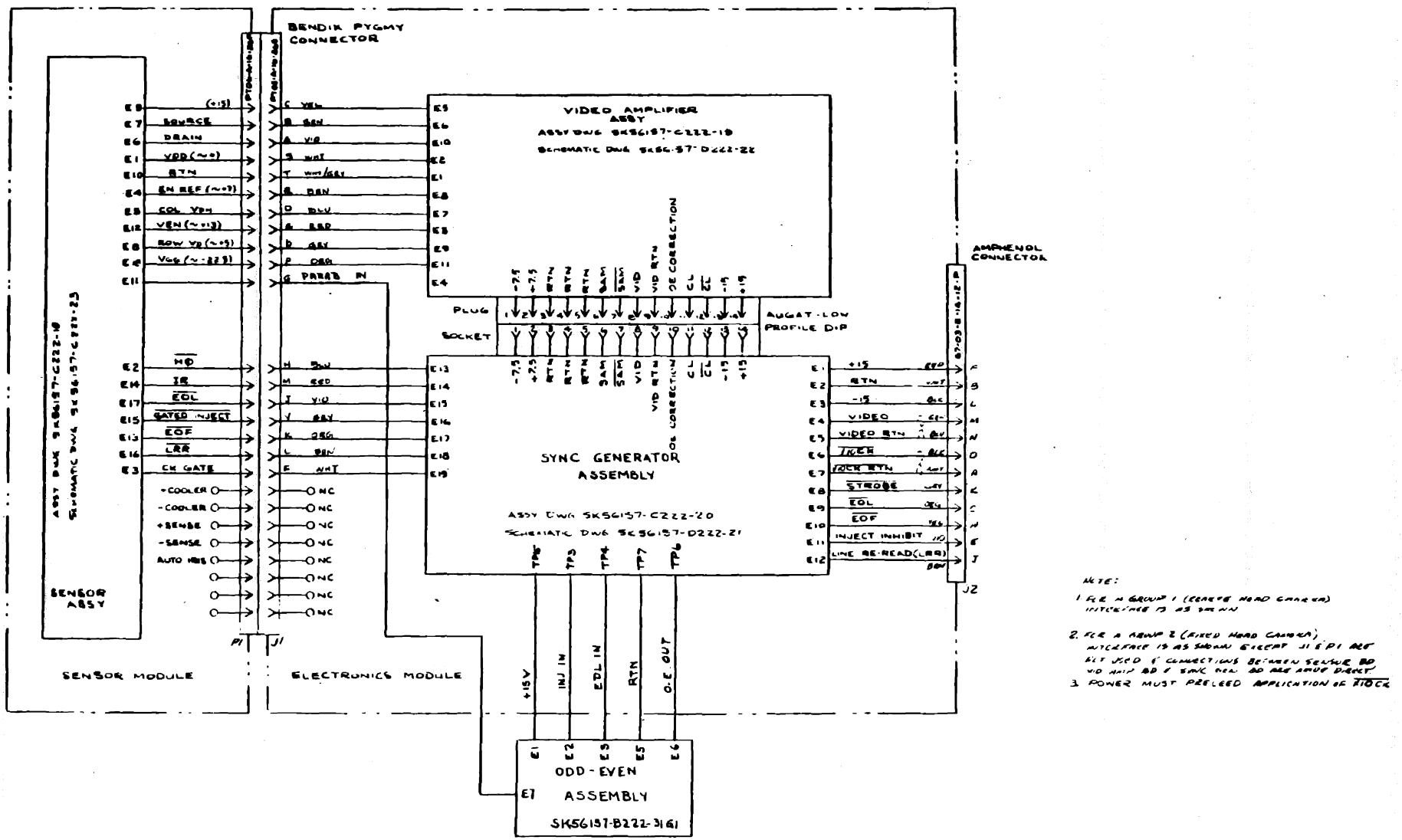
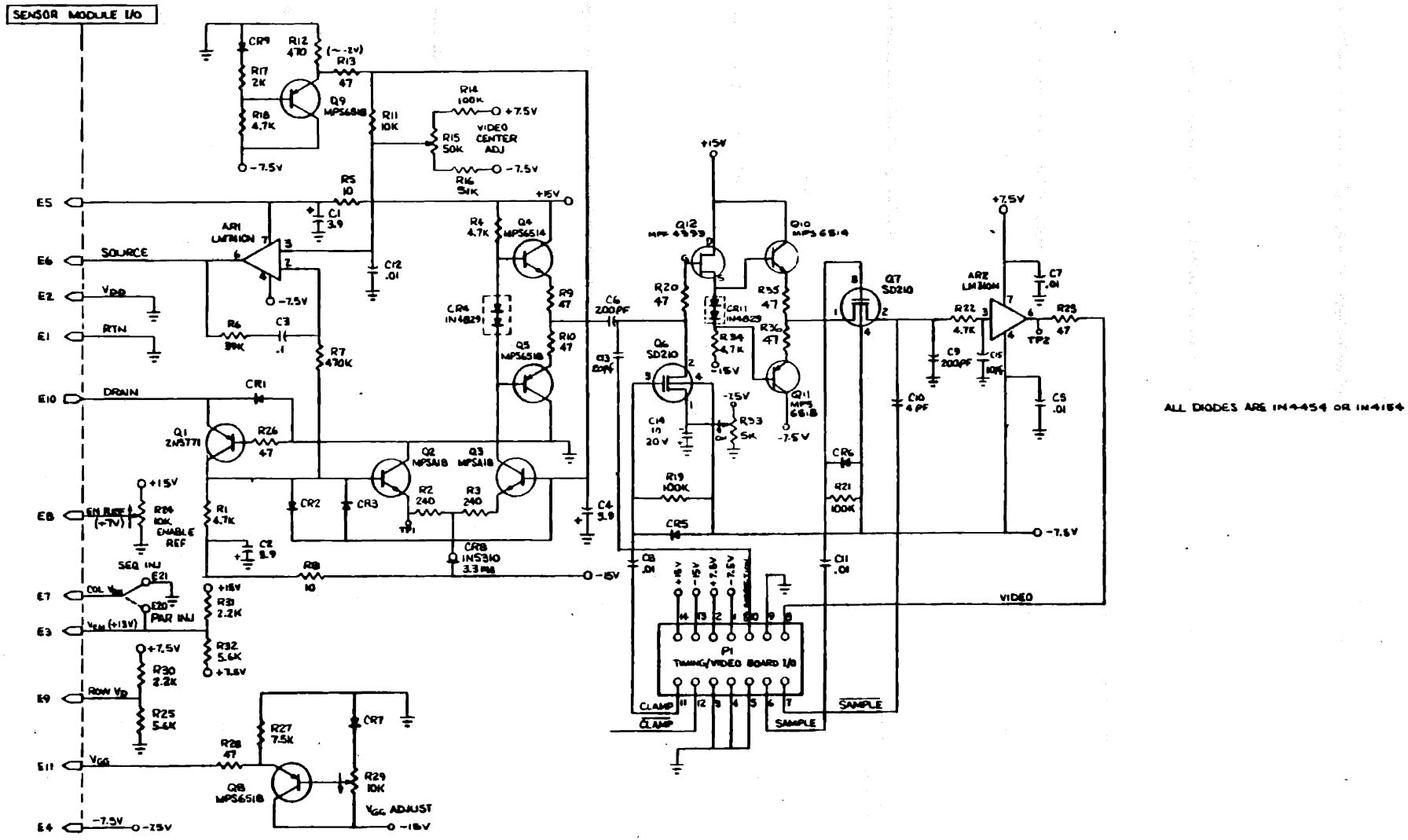


FIGURE A.2: **Interface Diagram TN2200/2201 Camera**



A-3

FIGURE A.3: **Schematic Diagram TN2200/2201 Video Amplifier.**

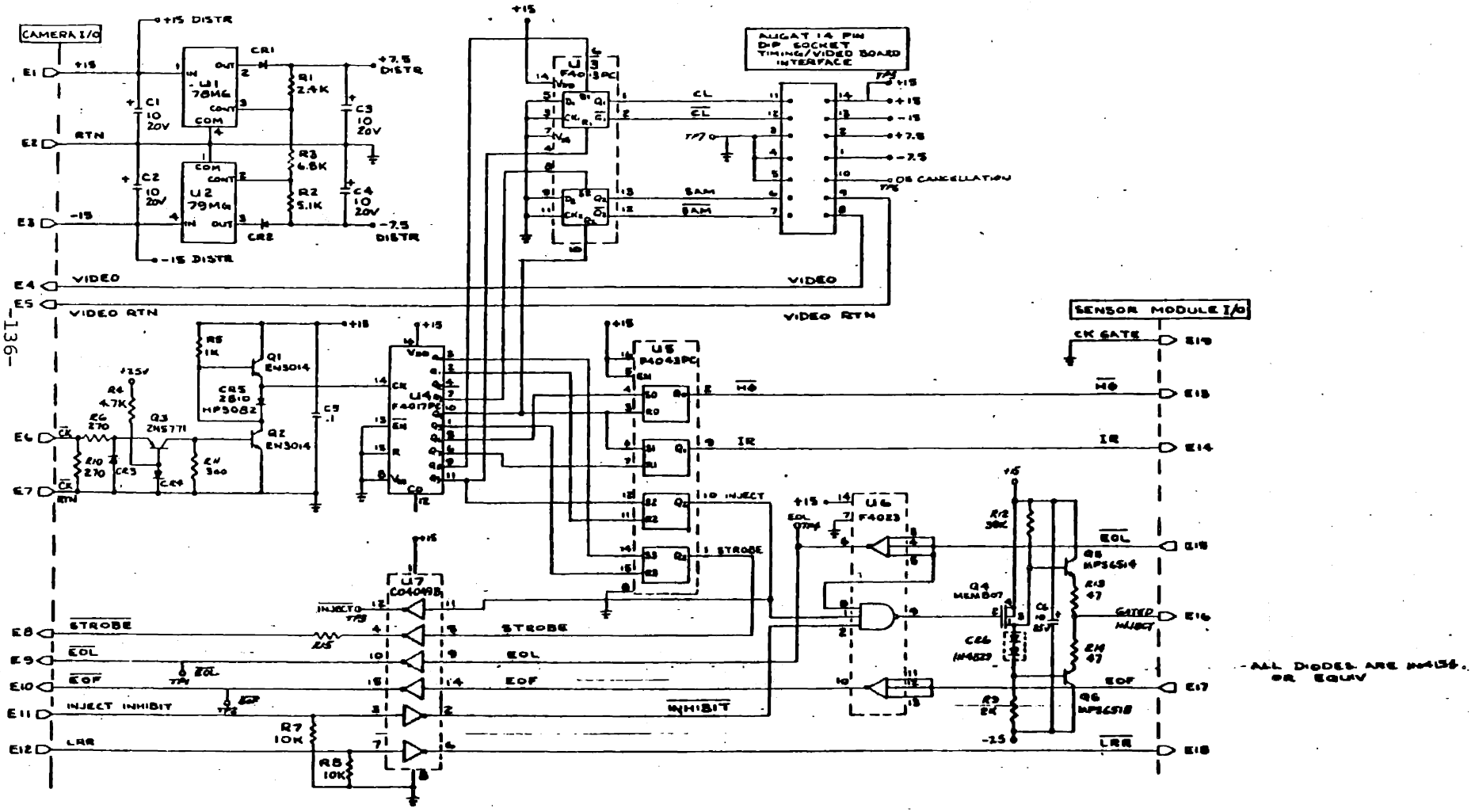


FIGURE A.4: **Schematic Diagram TN2200/2201 Sync Generator**

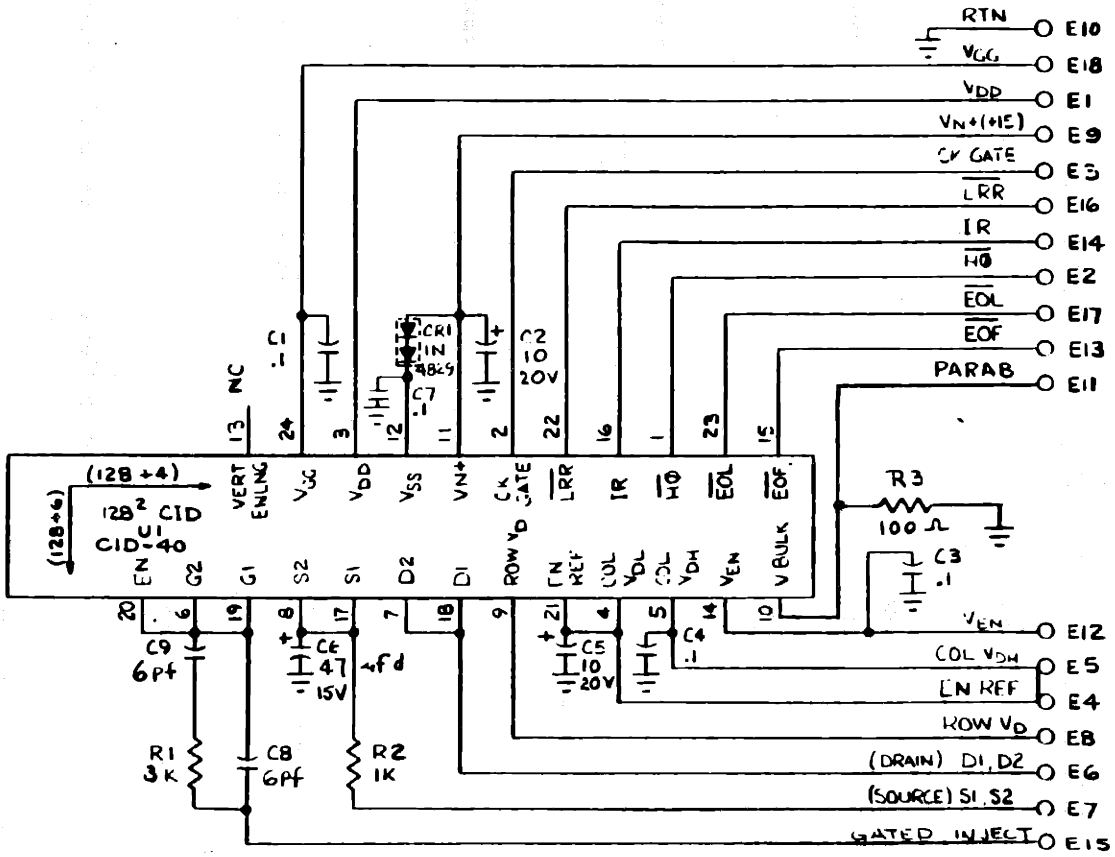


FIGURE A.5: **Schematic Diagram TN2200 Sensor Board**

A-J

APPENDIX B: DIGITIZER

A commercially available automation interface was used as the system digitizer. The specifiec component was the PN2110A Automation Interface made by Electronic Systems Division of General Electric.

The PN2110A is specifically designed to support the General Electric TN2200 solid state Automation Camera. It provides analog and digitized video, threshold video, power, clock signal, TTL signal level buffering and conversion, and analog sweeps for CRT display presentation. This is shown in the functional block diagram of Figure B.1.

A multi-conductor cable connects the Automation Interface to the camera itself. The camera may be driven by a clock signal from an external TTL source, or from a variable oscillation within the interface itself. In either mode, a buffer clock driver is provided between the clock source and camera. Clock frequency may be changed from the normal 10MHz setting by means of an 8-pole-dip-switch, located inside the unit, which controls a frequency divider. This switch can be replaced by hardwiring or remote switching.

Display monitor outputs in the form of analog sweeps and video are furnished on standard BNC connectors. Monitor sweeps maintain constant picture size, regardless of frame time.

Video output signals are available in the form of analog video, 8-bit-digitized video, or thresholded binary video. An

8-pole-dip-switch is provided within the unit to set the reference level for the threshold binary output. This switch also may be replaced by hardwiring or remote switching.

A multi-pin output connector implements the system interface. Outputs appearing at this connector include the 8-bit-digitized video, thresholded binary video, data rate clock (strobe), end of line (EOL) pulse, and end of frame (EOF) pulse. Also included are the display monitor outputs, in parallel with the standard BNC outputs previously mentioned. Camera control inputs may be provided through this connector from the externally interfaced computer or signal processor, to implement the line reread or injection inhibit functions and to establish digital output signal levels.

Use of this interface unit greatly simplifies applications of the GE TN2200 to automation systems.

Technical Specification Summary

Electrical

INPUTS

AC Voltage & Fuse	110V, 60 Hz - AGC 1/2 220V, 50/60 Hz - AGC 1/4
AC Input Power	18 watts
External Clock (optional)	TTL Level, 39 kHz to 10 MHz
Camera Control Signals	TTL Level

OUTPUTS

DC Power (camera)	± 15V, 50 MA
Clock Frequency	39 kHz to 10 MHz, switch switchable in 256 steps
Synchronization and Digitized Video Signals	TTL or Tristate Level, selectable by camera control input signal logic levels
Data Clock (Strobe)	TTL level
X and Y Analog Display Sweeps	IV Peak Sawtooth
Analog Video	IV Peak, White High

Mechanical

Weight	3 lb. (1.36 kg)
Size	10.7 in. x 10.7 in. x 3.2 in. (27 cm x 27 cm x 8 cm)
Camera Connector	DB25-S (ITT CANNON)
XYZ Display Monitor Connectors	BNC
Output Connector	DB25-P (ITT CANNON)
Case	Formed Aluminum

Environmental

Ambient Temperature	32° to 123°F (0° to 50°C)
Altitude	50,000 ft. (15,240 meters)

Performance

Picture Element (Pixel) Period	PN2110A = 1 to 256 µs (in 1 µs increments using internal clock)
	TN2200 = > 2 µs
	TN2201 = > 3 µs
A/D Converter	8 bits at 2 µsec/pixel period
	4 bits valid if pixel period is less than 2 µsec.

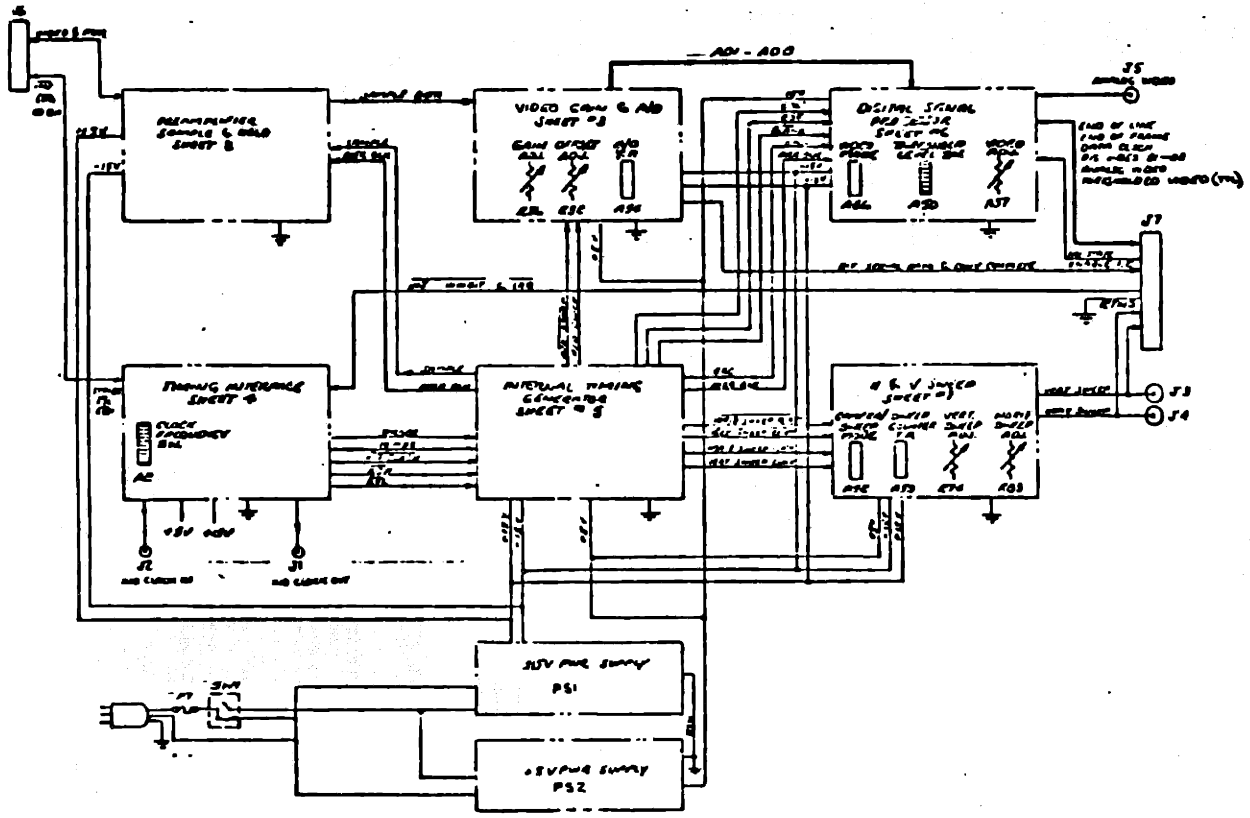
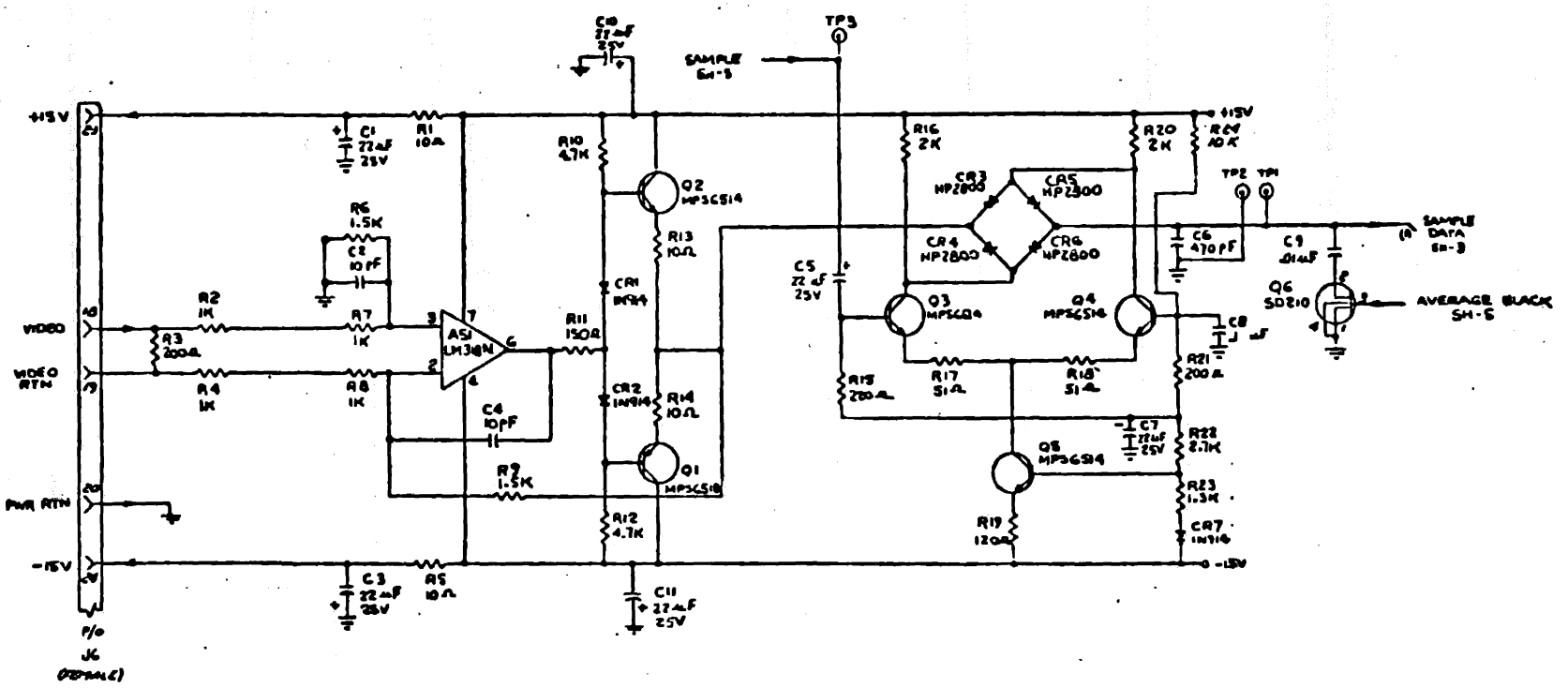


FIGURE B.1: PN2110A1 Schematic Organization



NOTES: 1. ALL RESISTERS ARE ¼ WATT UNLESS OTHERWISE SPECIFIED.

FIGURE B.2: **Preamp Sample and Hold**

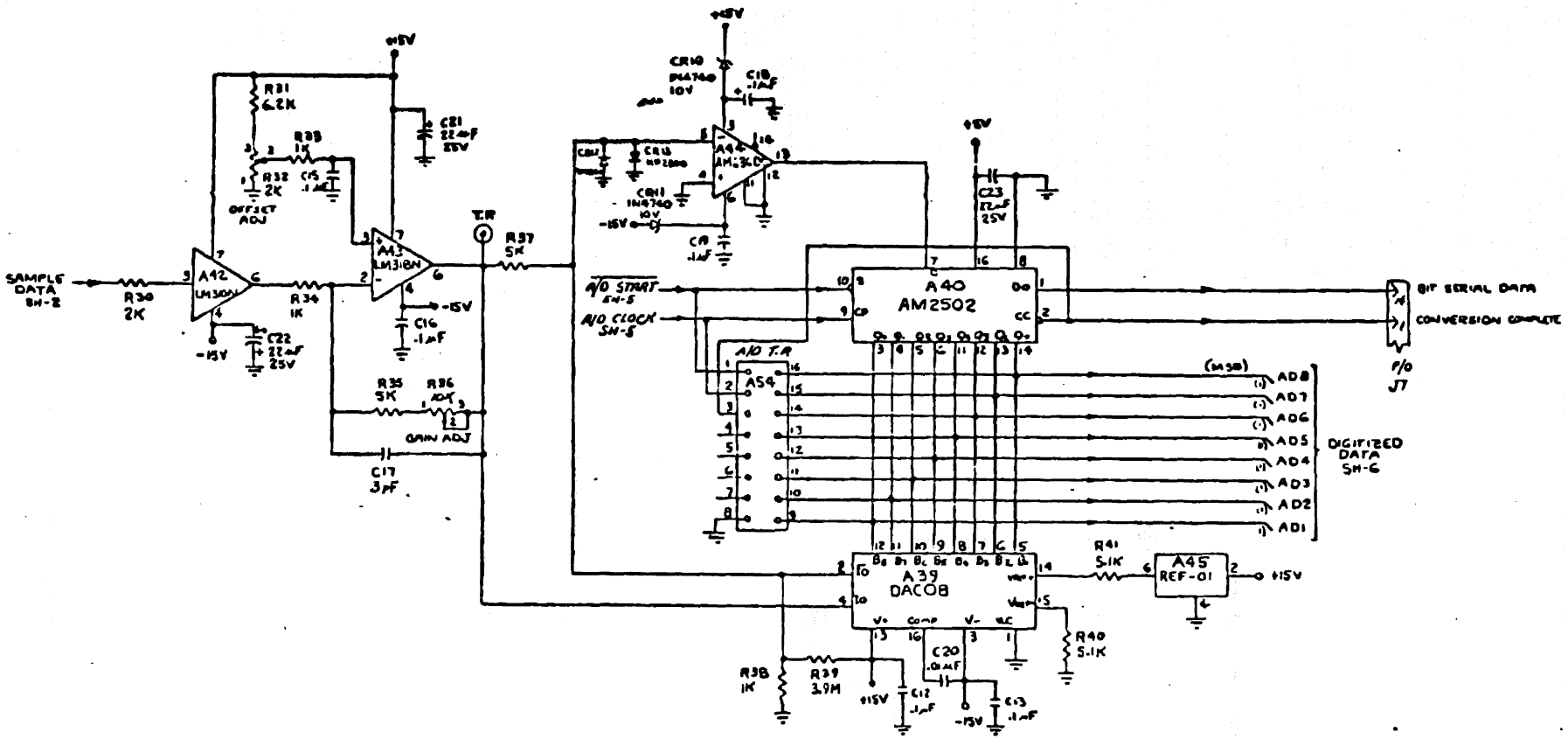


FIGURE B.3: Video Gain and A/D

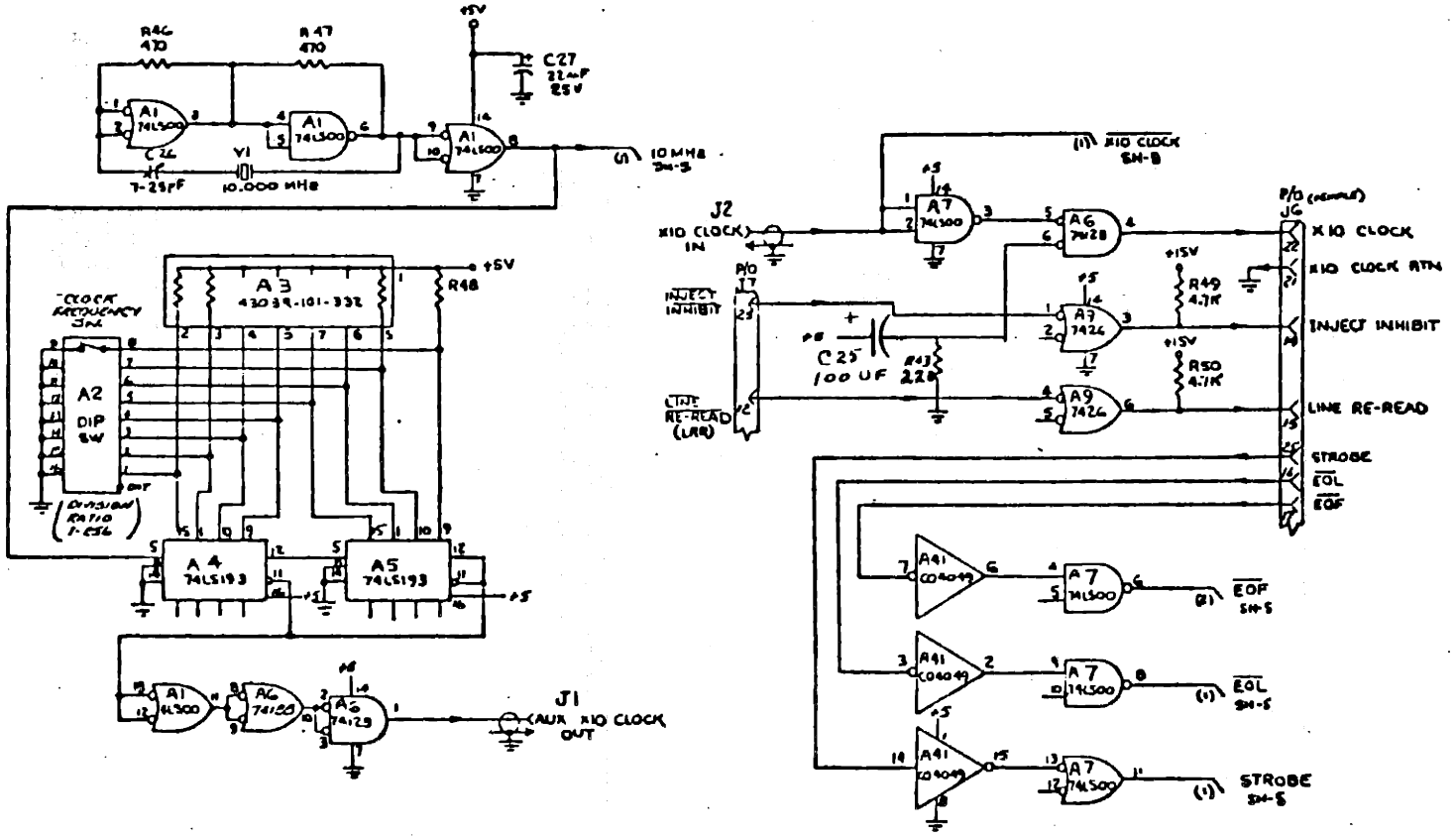


FIGURE B.4: Timing Interface

B-4

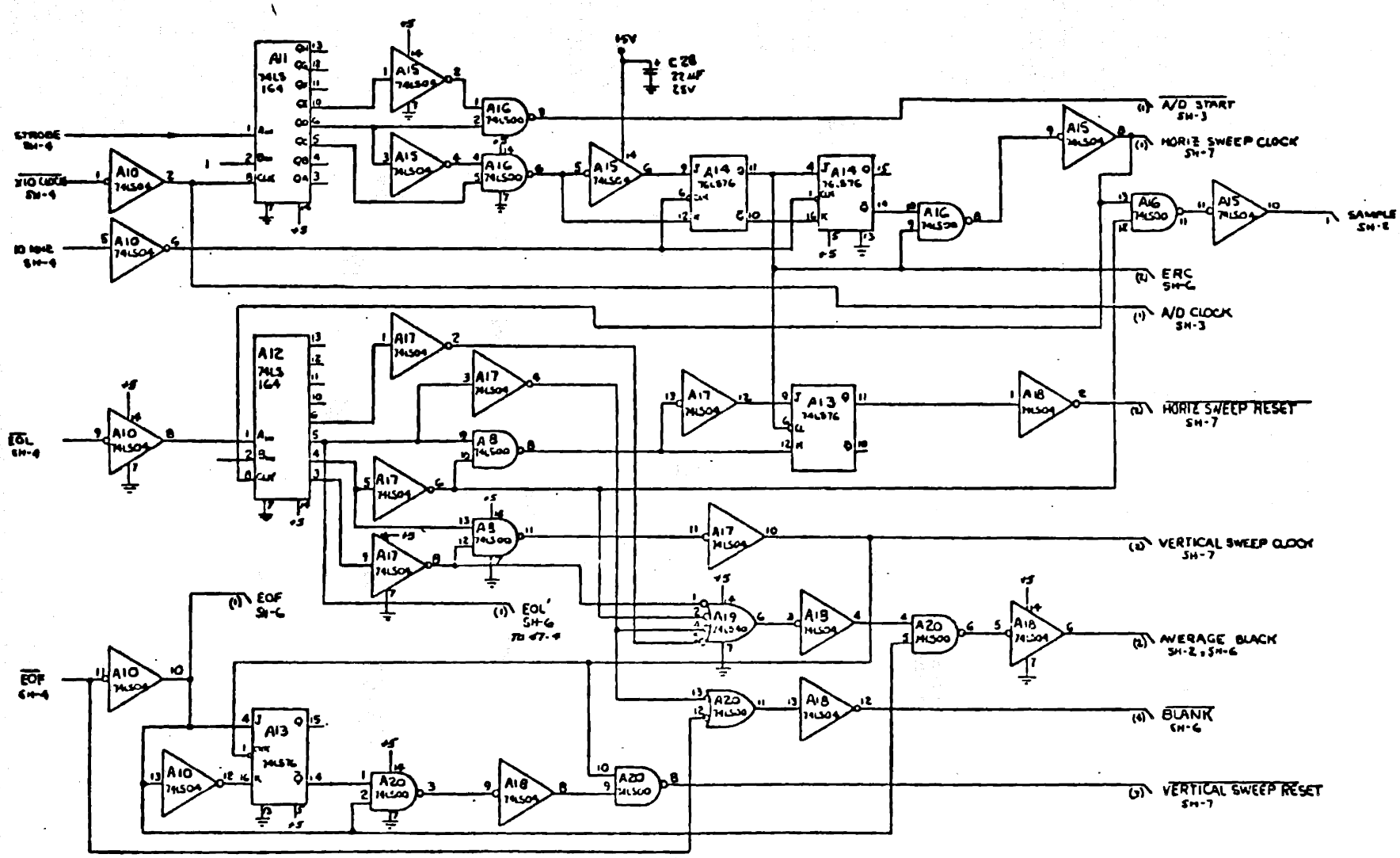


FIGURE B.5: Internal Timing Generator

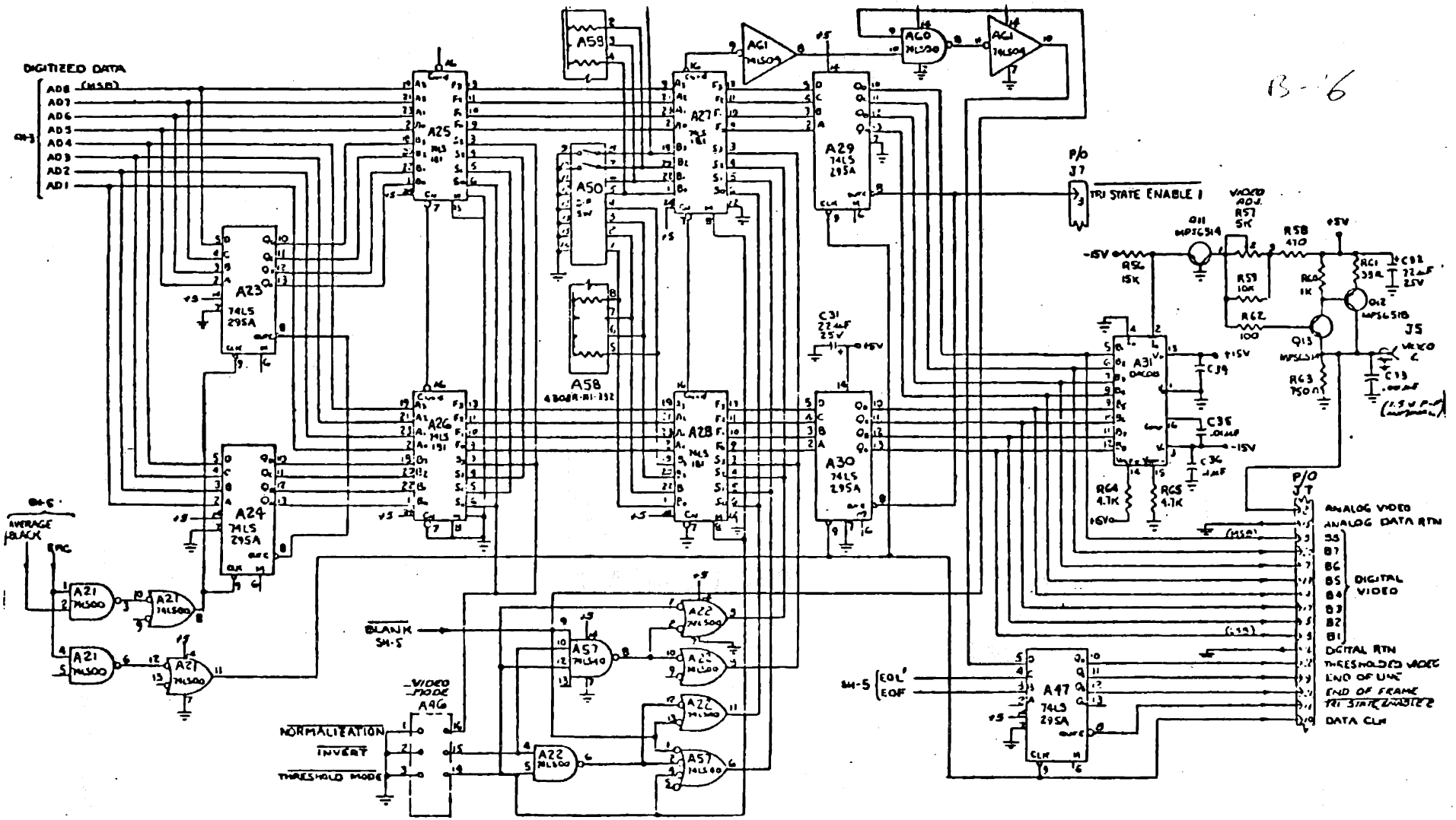


FIGURE B.6: Digital Support Processor

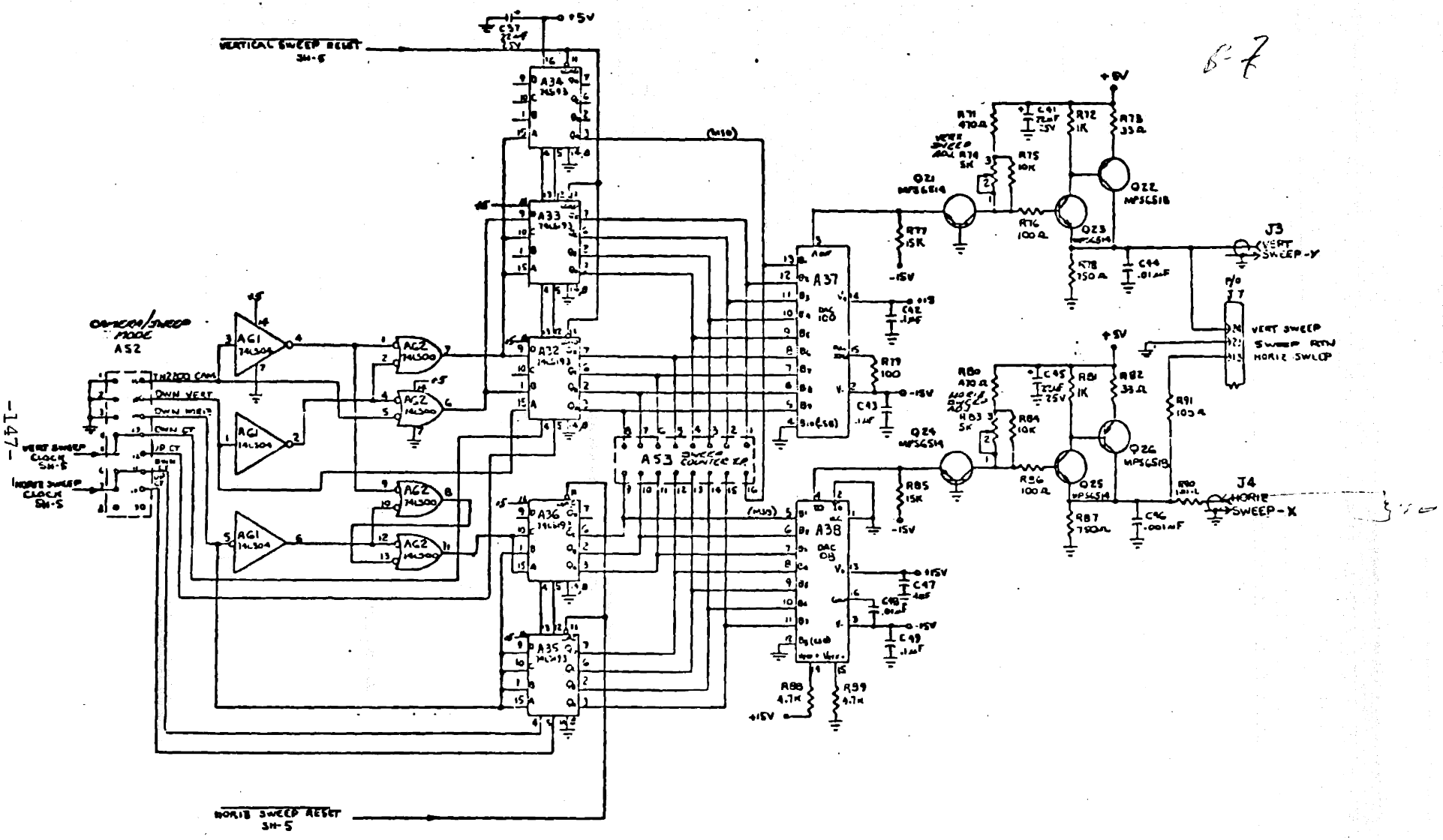


FIGURE B.7: **Horizontal & Vertical Sweep Generator**

APPENDIX C: SPOX - SPECIAL PURPOSE BOX

This box had to be specially designed to allow the kind of processing and display that SYSTEM GREAF required. A conventional microprocessor could not be used because of its speed limitations.

SPOX interfaced with the automation interface and the display. It's hardware allows the adjustments of framerate, resolution and grayscale.

Framerate is adjusted so that the sampling rate is changed while keeping the display rate a constant. This is accomplished by storing each frame in RAM and displaying it a number of times. If the sampling rate is 28 f/s then each frame is displayed after being sampled. In this event it is not stored in RAM. For slower frame rates, however, each frame is stored and displayed over to keep the display rate at 28 f/s. Thus, if the sampling rate is 14 f/s, each frame will be displayed twice. The system clock always runs at 10 MHz, so that at lower frame rates, some frames are merely ignored. This method is found preferable to lowering the clock speed for two reasons:

Figures B-1 through B-7 are various diagrams of relevance. For more details refer to [4].

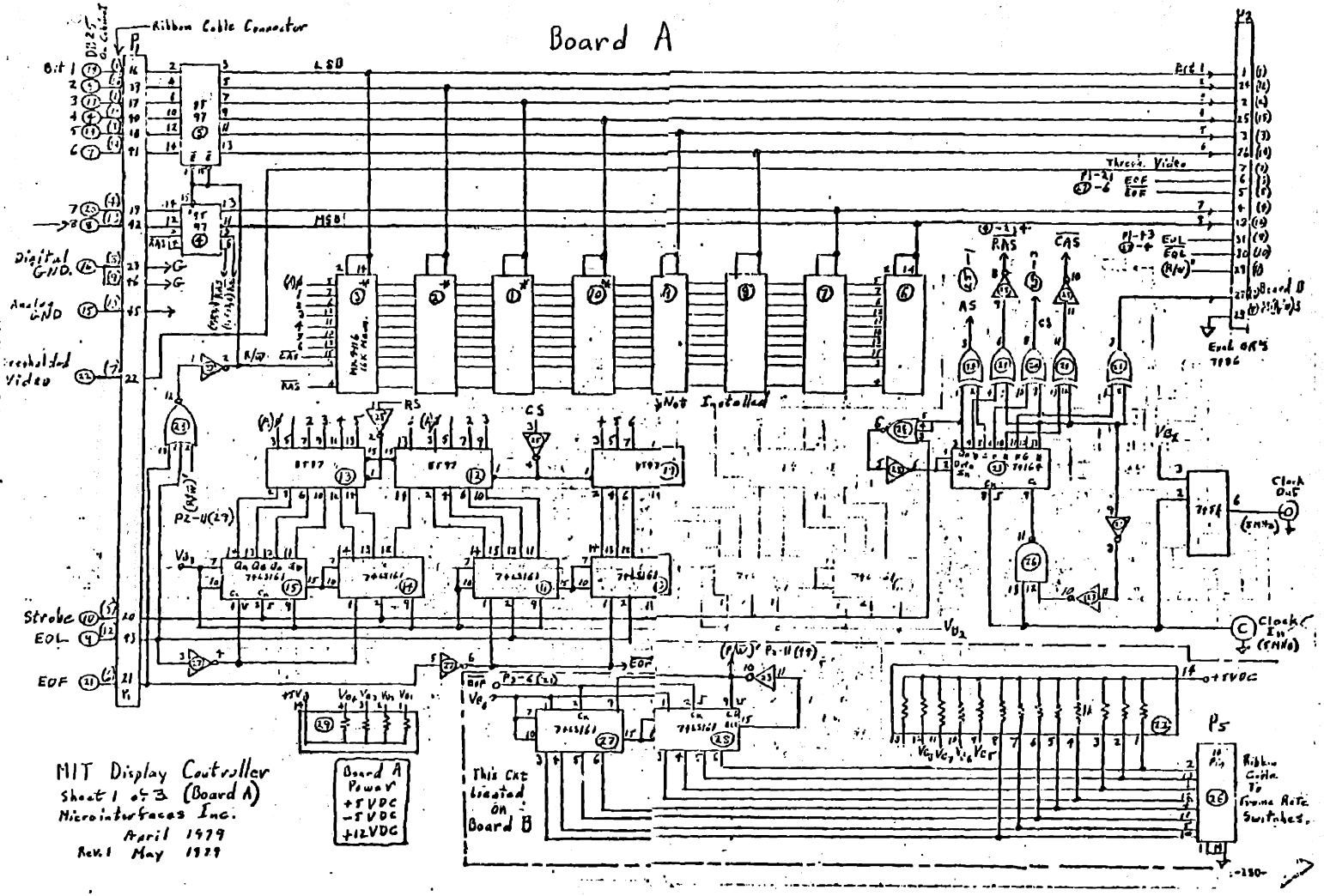
(1) The clock is used for display purposes and changing its frequency would affect the display.

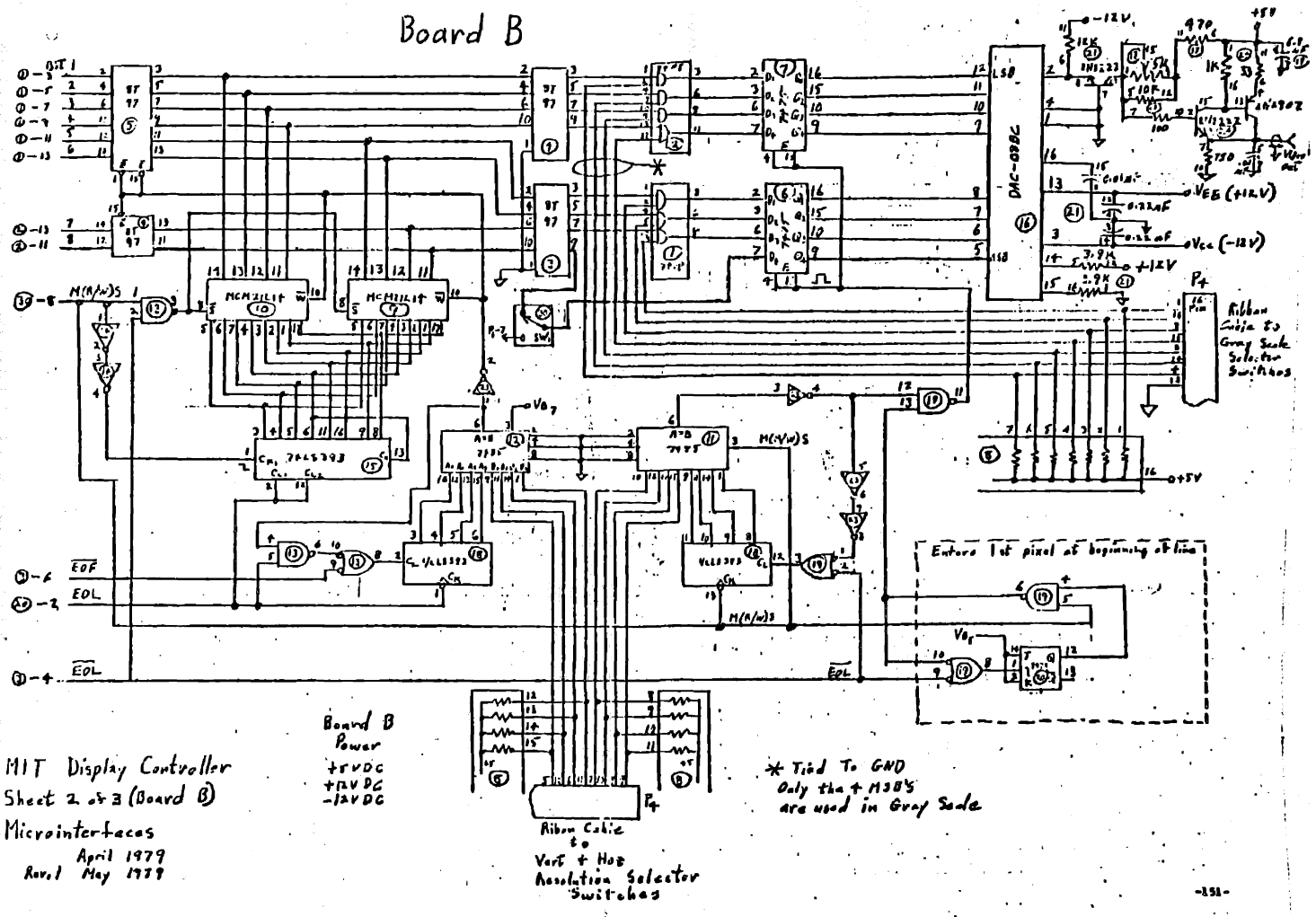
(2) The camera is a charge coupled device. At low frequencies the charges integrate noise so as to adversely affect the signal to noise ratio.

Grayscale adjustment is relatively easy since lowering the number of bits of grayscale merely requires ignoring the most significant bits. The grayscale lines are used directly from the automation interface.

Resolution can be changed by making use of a 4-bit counter in SPOX. For resolutions lower than 128x128, the first pixel is stored and displayed over again to extend one horizontal line. For the purpose of reducing the vertical resolution, this pixel must be stored until a new line is displayed since it has then to be displayed again. Figure C-1 is a schematic of the whole system.

The SPOX was designed and built to M.I.T. specifications by A. Woodin of Microinterfaces, Inc.


Board A
Ribbon Cable Connector
P1
P2
LSB
MSB
Not Installed
RAS
CAS
RS
CS
Strobe
EOL
EOF
Digital GND
Analog GND
Thresh. Video
Clock Out
Clock In
Board A Power
+5VDC
-5VDC
+12VDC
This Ckt located on Board B
P5
Ribbon Cable To Frame Rate Switches
MIT Display Controller
Sheet 1 of 3 (Board A)
Microinterfaces Inc.
April 1979
Rev 1 May 1979



Board B
Power
+5VDC
+12V DC
-12V DC

* Tied To GND
Only the + MSB's
are used in Gray Scale

MIT Display Controller
Sheet 2 of 3 (Board B)
Microinterfaces
April 1979
Rev.1 May 1979

REFERENCES

1. Sheridan and Verplank, Human and Computer Control of Undersea Teleoperators.

2. Ocean Engineering Report: Naval Undersea Center, San Diego, CA 92132, January 1975.

3. TN2200 Manual [General Electric].

4. PN2110A Manual [General Electric].

5. Rosenfield & Kak, Digital Picture Processing.

6. Pratt, Digital Image Processing.

7. Brooks, S.M. Thesis, M.I.T., June 1979.

8. Meyers, J.J., Holm, C.H., and McAllister, R.F., Handbook of Ocean and Underwater Engineering, McGraw Hill, 1969.

9. Rechritzer, A.B. and Sutter, W., "Naval Applications of Remote Manipulation", Remotely Manned Systems Exploration and Operation in Space, California Institute of Technology, 1973, pp. 493-502.

10. Schneider, M.R., "Task Analysis for Undersea Manipulators", S.M. Thesis, M.I.T., March 1977.

11. Winget, C.L., "Hand Tools and Mechanical Accessories for a Deep Submersible", Remotely Manned Systems Exploration and Operation in Space, CIT, 1973, pp. 502-523.