AAMRL-TR-88-061

AD-A209 397

INVESTIGATION OF A LINEAR SYSTEMS MODEL FOR HUMAN VISUAL DETECTION AND SPATIAL FREQUENCY DISCRIMINATION (U)

HAROLD S. MERKEL, CAPTAIN, USAF

ARMSTRONG AEROSPACE MEDICAL RESEARCH LABORATORY

DECEMBER 1988

Final Report for October 1986 - December 1988.

Approved for public release; distribution is unlimited.

89 6 27 194

NOTICES

Please do not request copies of this report from Armstrong Aerospace Medical Research Laboratory. Additional copies may be purchased from:

> National Technical Information Service
> 5285 Port Royal Road
> Springfield, Virginia 22161

Federal Government agencies and their contractors registered with Defense Technical Information Center should direct requests for copies of this report to:

> Defense Technical Information Center
> Cameron Station
> Alexandria, Virginia 22314
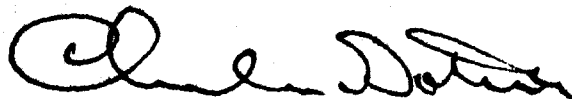
TECHNICAL REVIEW AND APPROVAL

AAMRL-TR-88-061

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Director, Human Engineering Division
Armstrong Aerospace Medical Research Laboratory

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	Approved for public release; distribution is unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S) AAMRL-TR-88- 061	5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION Armstrong Aerospace Medical Research Laboratory, AFSC, HSD	6b. OFFICE SYMBOL (If applicable) AAMRL/HEF	7a. NAME OF MONITORING ORGANIZATION

6c. ADDRESS (City, State, and ZIP Code) Wright-Patterson AFB OH 45433-6573	7b. ADDRESS (City, State, and ZIP Code)

8a. NAME OF FUNDING / SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			

	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO	WORK UNIT ACCESSION NO.
	62202F	7184	18	D2

11. TITLE (Include Security Classification)
Investigation of a Linear Systems Model for Human Visual Detection and Spatial Frequency Discrimination (U)

12. PERSONAL AUTHOR(S)
Merkel, Harold S., Captain, USAF

13a. TYPE OF REPORT Final	13b. TIME COVERED FROM Oct 86 TO Dec 88	14. DATE OF REPORT (Year, Month, Day) 1988 December	15. PAGE COUNT 50

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Linear Systems Analysis Visual Detection
14	02		Contrast Threshold Contrast Sensitivity
			Frequency Discrimination Human Model

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This research investigated the application of a linear systems model to two parameters of vision: the contrast threshold and the spatial frequency difference threshold. The contrast threshold is the contrast of a target when an observer is just able to detect its presence. The spatial frequency difference threshold is the smallest difference in spatial frequency that permits two grating targets to be distinguished.

The model investigated was for observation with a fixed visual field size of one-dimensional, time-invariant sinusoidal grating targets. A mathematical development indicated that a linear model could be used to represent the human spatial frequency difference threshold function.

(CONTINUED)

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL Harold S. Merkel, Capt, USAF	22b. TELEPHONE (Include Area Code) 513-255-8763	22c. OFFICE SYMBOL AAMRL/HEF

DD Form 1473, JUN 86   Previous editions are obsolete.   SECURITY CLASSIFICATION OF THIS PAGE

Abstract continued:

The model was implemented using an electro-optical hardware system which consisted of a charge-coupled device video camera, a frame grabber, and a personal computer. Because of its similarities to the structure of the eye, linear response, and ability to acquire digital image data, a charge coupled device array video camera and frame grabber were used to simulate the eye. The action of the neural pathways and visual cortex was simulated by Fourier transform computations on the camera and frame grabber output. The experimental data from the electro-optical hardware system agreed with the theoretical models for both parameters.

Preface

The research described in this thesis was performed at the Armstrong Aerospace Medical Research Laboratory (AAMRL), Human Engineering Division, Crew Systems Effectiveness Branch. Initial funding was provided by the Laboratory Director's Fund Program and the work was accomplished under project 7184 18D2. This thesis was submitted to the School of Graduate Studies at Wright State University (WSU) in partial fulfillment of the requirements for the degree of Master of Science.

There are several individuals who provided valuable guidance during the course of this research. I would like to thank Dr. H. Lee Task (AAMRL) for his overall direction, tutoring on visual and optical systems, and careful review of the drafts. I would also like to express my appreciation to Dr. Daniel Repperger (AAMRL) for his counsel on statistics and Mr. Charles Goodyear (Systems Research Laboratories) for his assistance with the Statistical Analysis System. Finally, I would like to thank my thesis advisor, Dr. Russell Hannen (WSU), and committee members Dr. Kefu Xue (WSU) and Dr. Richard Bethke (WSU) for their review of and constructive comments on this thesis.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction

For many decades Linear Systems Analysis (LSA) has proved useful in modeling responses of optical and electrical systems. More recently LSA has also been applied to modeling parameters of the human visual system. Since the visual system contains optical components and its neural pathways are similar in many respects to electrical circuits, this seemed like a natural step. However, the visual system also has some fundamental differences from optical and electronics systems.

The primary difference is the visual system is a biological system, whereas optical and electrical systems are strictly physical. Many biological parameters are known to exhibit nonlinearities, vary over time, and are not homogeneous or isotropic. Because linearity, homogeneity, isotropy, and time invariance are requisites for the proper application of LSA, it is uncertain that LSA will yield accurate results for visual parameters. While it is true that virtually all physical systems (even optical and electronic systems) do not absolutely meet these criteria, many are sufficiently linear, homogeneous, isotropic, and time invariant over a specified operating region that LSA does provide good results.

Since it is arguable that the visual system meets the requirements for employing LSA, two fundamental questions are of interest: "Can a relatively simple model based on LSA provide adequate modeling of visual parameters?" and, if so, "Under what conditions and for which visual parameters does the model hold?"

The answer to the first question is a qualified yes; LSA models have been able to predict some responses. For example, several researchers have related contrast thresholds of various

targets using LSA. For many other parameters, however, the utility of LSA is unknown.

The contrast threshold and the spatial frequency difference threshold are two examples of visual parameters that may be modeled using LSA. The contrast threshold is the contrast of a target when an observer is just able to detect its presence. The difference threshold is a measure of ones ability to discriminate between two spatial frequencies. (There are several other parameters that need to be specified to give the contrast threshold meaning and the difference threshold meaning, such as the mean target and background luminances, the type of target pattern, and the location within the field of view.)

The objectives of this thesis are to: 1) investigate the application of a simple LSA model to two specific visual parameters: the contrast threshold and the spatial frequency difference threshold, 2) implement a LSA model using electro-optical hardware, and 3) compare some responses of the hardware implementation to responses of the human visual system, obtained from the empirical data available in the published literature.

1.2   Scope

Constructing a visual model is inherently difficult because there are numerous parameters which affect visual processes and the exact functioning of the human visual system is still partly unknown. For example, a model of visual detection might account for parameters such as luminance of the target, modulation of the target, location of the target within the visual field, target orientation, color contrast, and temporal properties, such as duration of exposure. Due to the large number of variables associated with virtually every visual performance metric, most models are relatively limited in scope or they become extremely complex and require massive computational capability in implementation.

Due to the reasons stated above, the scope of this research must be well defined so the number of parameters will not enlarge the scale of the research to an unmanageable level. This effort will consider a one-dimensional threshold model for two visual parameters and implement this model using a basic hardware system. The research will be limited to the detection and frequency discrimination of black and white sinusoidal grating targets. The model will be for a fixed visual field size with evenly distributed receptors, to simulate the fovea of the eye, where the spatial density of receptors (cones) are fairly uniform. The model will also assume time invariance, so temporal effects are not considered.

2

Chapter 2

Background

2.1 The Human Visual System

Since the objectives of this research involve modeling parameters of the human visual system, it is appropriate to briefly review the basic components and functions of the human eye-brain system.

The human visual system is usually considered to include the two eyes, the neural pathways which connect the eyes to the brain, and the visual cortex portion of the brain (see figure 2.1). For the purposes of this research it is not necessary to consider the neural pathways and the visual cortex except to note they are responsible for transmitting, processing, and making discriminatory judgements on the signals that originate at the eye. In this sense, the visual system will be treated as a "black box." The primary focus in this section is on the basic structure of the eye, specifically its first order optical characteristics and the distribution of the receptors (rods and cones) within the eye.

The eye works functionally as follows: Light rays enter the eye and are transmitted through the cornea, lens, and vitreous humor and brought to a focus on the retina. Optically, most of the refraction of these rays is due to the air/cornea interface, which typically provides about 40 diopters of optical power. Additional refraction is provided by the lens, which has variable optical power and allows one to adjust the focal distance of the system (accommodate) to obtain a clear image of objects at various distances. The amount of light transmitted is controlled by the iris, which expands and contracts to constrict and dilate the pupil.

At the retina the optical image falls upon two types of photoreceptors: rods and cones.
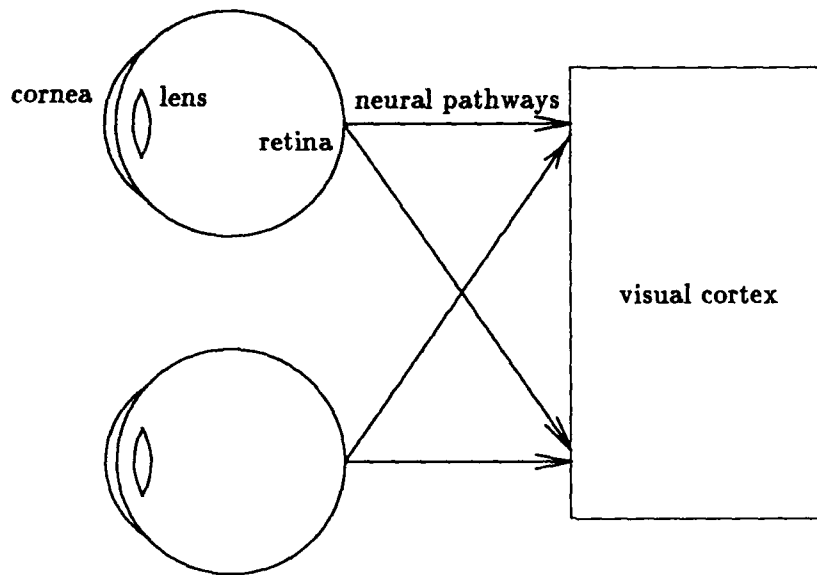
Figure 2.1: The Human Visual System.

These receptors have differing characteristics, such as their spatial distribution on the retina and their sensitivity to light. The cones, which are less light sensitive than the rods, are concentrated in the central portion of the retina. In the central 2°of the retina, called the fovea, the receptors are almost exclusively cones, with the rods being virtually absent. Here, where the distribution of cones is highest and approximately uniform, humans have their best visual acuity. Contrary to the distribution of the cones, the density of the rods increases with distance from the fovea, reaching a maximum density at about 20°off axis. This is the point of best light sensitivity. The distribution of the rods and cones is shown graphically in figure 2.2.

In terms of light sensitivity, there are two important distinctions between rods and cones. First, rods are much more light sensitive than cones, and second, rods are not sensitive to the wavelength, or color of light, whereas cones are differentially sensitive to wavelength.
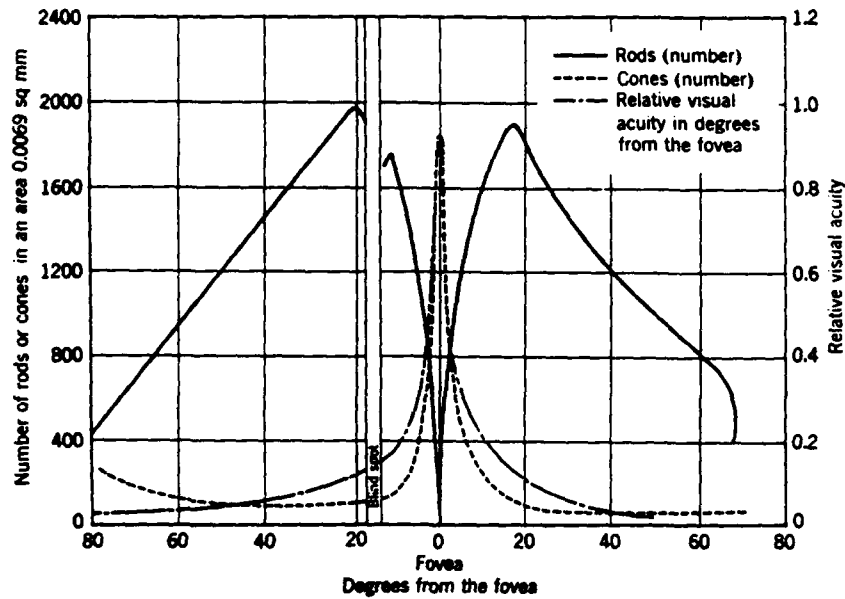
Figure 2.2: Distribution of rods and cones.

2.2 Previous Work by Other Researchers

The general area of interest under which this thesis may be classified is linear systems modeling of the human visual system. The objective of most researchers who have worked in this area has been to develop hypotheses and models of visual responses to explain the results of psychophysical experiments. Some of this work relevant to this thesis is discussed below.

There are a few well known publications which are renowned in the area of LSA modeling of visual parameters for having pioneered a specific theory or model. One of the foremost articles to use LSA to explain visual detection was by Campbell and Robson [5]. They showed the contrast thresholds of various types of gratings could be related using Fourier theory. Furthermore, grating patterns of complex wave forms (square, rectangular, or sawtooth) could not be distinguished from sine wave gratings until their contrast had been raised to a level at which the higher harmonic components reached their respective detection threshold. These results propagated the theory that the visual system could be modeled by LSA.

In 1970 Sachs, Nachmias, and Robson published *Spatial-Frequency Channels in Human Vision* [17]. This was one of the first works to propose a specific mathematical model of visual detection using the concepts of LSA. A block diagram of this multiple-channel model is shown in figure 2.3.

5

The Sachs model uses the concept of channels, or frequency-limited bands, which have individual characteristic thresholds to explain the shape of the contrast sensitivity function (CSF). The CSF is a plot of contrast sensitivity (the reciprocal of contrast threshold) versus spatial frequency. Instead of a continuous contrast threshold function, a set of frequencies bands with a threshold for each band is specified. Since visual detection is a probabilistic event, the model also has a random noise factor which is added to the signal. When the amplitude of the sum of the signal and the noise within a band exceeds the threshold for that band, the signal is detected. Since the model has several channels acting in parallel and an image may contain frequency components across several channels, the image is detected if any one of the frequency components exceeds threshold within its respective channel. Because the noise is random, detection may occur on one occasion and not on another.

This model permits the detection threshold of a complex grating to be predicted by knowledge of the thresholds of the sinusoidal gratings which compose the complex grating. This theory was validated with a simple experiment using human subjects observing two sinusoidal grating patterns and a third complex grating pattern which was a linear combination of the first two.

Another parameter of the visual system which may be related using a LSA model is the spatial frequency difference threshold. The spatial frequency difference threshold, or more succinctly, the difference threshold, is a measure of one's ability to discriminate between two grating patterns which differ in spatial frequency. The difference threshold may be defined as the smallest difference in frequency that allows two grating patterns to be distinguished. Several researchers have studied the human difference threshold in the last twenty years. The first study was reported in 1970 by Campbell *et al.* [4] They measured the just noticeable frequency ratio $\Delta f/f$ and determined that it was approximately constant independent of frequency, with no local irregularities. In 1982 Regan *et al.* [15] reported results similar to those of Campbell. These results are in agreement with Weber's law, which states that the just noticeable difference of a stimulus is proportional to the magnitude of the stimulus. (This is a well known law in psychophysics and has been observed to hold for numerous sensory parameters.)

Other studies have confirmed that the just noticeable frequency ratio is roughly constant independent of frequency, but they have also provided evidence that the $\Delta f/f$ curve is not exactly flat, but contains regular monotonicities which were observed for almost all

6
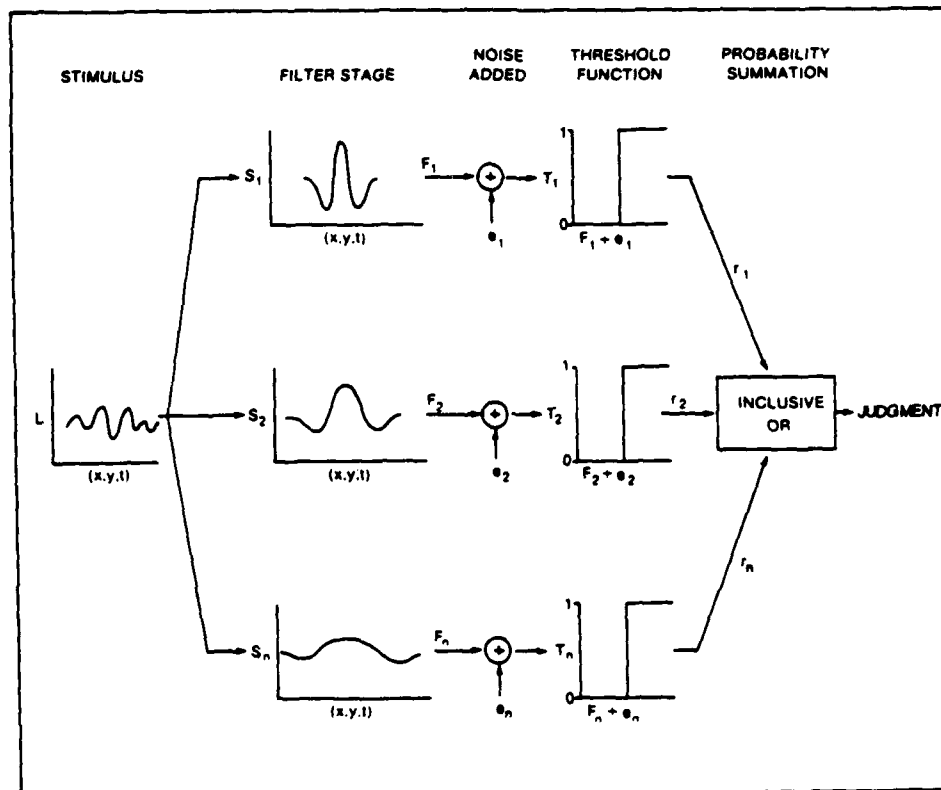
Figure 2.3: Threshold model for detection. The stimulus, specified by the luminance function $L(x, y, t)$, is processed by n pathways that act in parallel. A band pass filter F_i passes frequencies which correspond to a specific channel. Independent noise is added, and the sum is passed through the threshold function T. The output of the threshold device r_i takes the value 0 or 1. Detection occurs is $r_i = 1$ for one or more pathways. Reprinted with permission from Boff, Kaufman, and Thomas [3].

subjects in the experimental group. Hirsch and Hylton [9] and Richter and Yager [16] both reported this behavior in their research. The discovery of these monotonicities was due to their sampling between wider spaced spatial frequencies of the earlier studies. All studies reported, however, that the just noticeable frequency ratio was approximately 0.02–0.05.

There are several factors worth noting that have been reported concerning the behavior of the just noticeable frequency ratio. First, all of the data gathered was for suprathreshold conditions, *i.e.*, the contrast of the grating patterns was significantly above the detection threshold. Secondly, the results did not depend to a large degree on the number of cycles present in the stimulus, as long as there were a minimum number of cycles present. This minimum was usually determined to be two to three cycles. Also the results did not depend upon the exact contrasts of the gratings, as long as they are greater than approximately three times the detection threshold. Finally, for the literature reviewed, no LSA-based models were offered to explain the behavior of the difference threshold.

To summarize, there have been many articles published in the last twenty-five years on LSA models of visual parameters. Many of these publications allude to or speak only qualitatively about the application of Fourier analysis to vision. Those that do present a quantitative model are usually models for the contrast threshold. The predominant models for the contrast threshold treat the visual system as a black box and use the concept of frequency channels to explain the shape of the CSF. Though the spatial frequency difference threshold has been investigated by several researchers, no LSA models were found for this parameter.

2.3 Application of Fourier Theory to Visual Models

Fourier theory is the cornerstone of LSA and has been the basis for several models of visual parameters. This section will present a brief introduction to the Fourier transform and derive the relationships of the Fourier transform to visual models for the contrast threshold and the spatial frequency difference threshold. (There are numerous texts available that offer an indepth treatment of Fourier theory, such as *Signals and Systems* [14], to which the reader may refer.)

A general periodic function may be represented by an infinite sum of sine and cosine functions. More specifically, a periodic function $f(t)$ in the time domain may be expressed

8

as:

$$f(t) = \sum_{n=-\infty}^{\infty} \alpha_n e^{jn\omega_0 t} \tag{2.1}$$

where

$$\alpha_n = \frac{1}{T} \int_{t_0}^{t_0+T} f(t)e^{-jn\omega_0 t}dt \tag{2.2}$$

and $\omega_0 = \frac{2\pi}{T}$. These equations represent the exponential form of the Fourier series. Each α_n is the coefficient of the amplitude of a sine or cosine term of frequency $n\omega_0 t$.

Closely related to the Fourier series is the integral expression of the Fourier series, the Fourier transform. Unlike the Fourier series, the Fourier transform permits frequency domain representation of nonperiodic functions. The Fourier transform of a time function $f(t)$ is defined by:

$$\mathcal{F}\{f(t)\} = F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt \tag{2.3}$$

and the inverse Fourier transform of $F(\omega)$ is:

$$\mathcal{F}^{-1}\{F(\omega)\} = f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{j\omega t}d\omega \tag{2.4}$$

As long as $f(t)$ satisfies a rather loose set of conditions (called the Dirichlet conditions) $F(\omega)$ and f(t) can be uniquely determined from one another.

At this point it is appropriate to look at a couple of relevant examples of Fourier transform pairs, *i.e.*, specific functions of time and frequency which are related by equations 2.3 and 2.4.

The transform of the sinusoid $f(t) = A\cos\omega_0 t$ is a pair of scaled delta functions, as shown in figure 2.4. Figure 2.5 shows the transform of a constant function $f(t) = A$ is a scaled delta function at a frequency of zero, called "dc." These two transform pairs will be of particular interest in explaining the contrast threshold model.
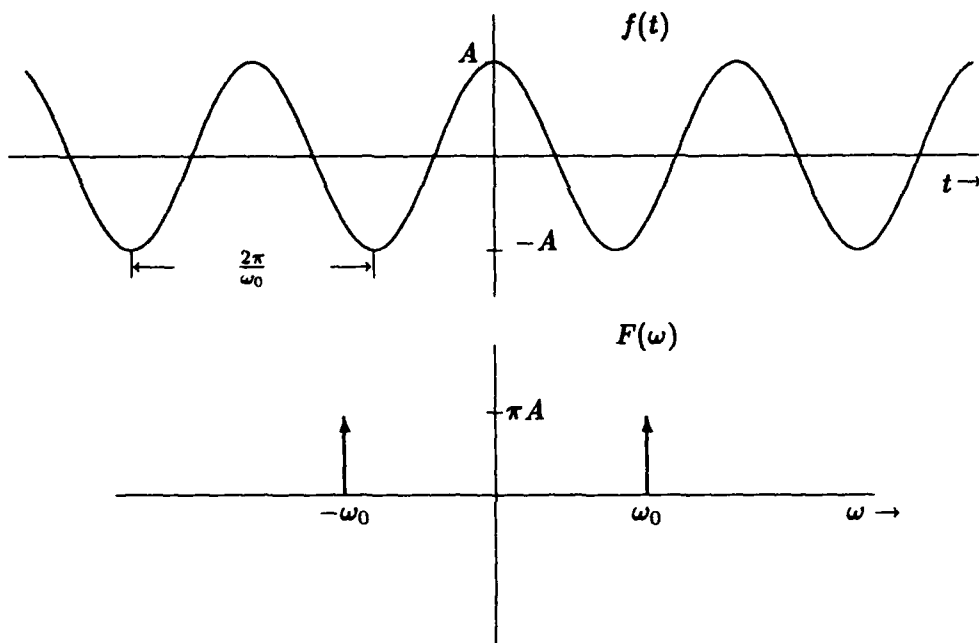
Figure 2.4: Fourier transform pair for $f(t) = A\cos\omega_0 t$.

Though the previous examples were functions of time, the Fourier transform also applies to spatial signals. Consider a spatially varying sine wave with a maximum luminance A and a minimum luminance B and contrast C defined as:

$$C = \frac{A - B}{A + B} \tag{2.5}$$

This function can be expressed as the sum of a cosine of amplitude $\frac{A-B}{2}$ centered around a luminance of zero and a signal of constant luminance of $\frac{A+B}{2}$. Because of linearity, the
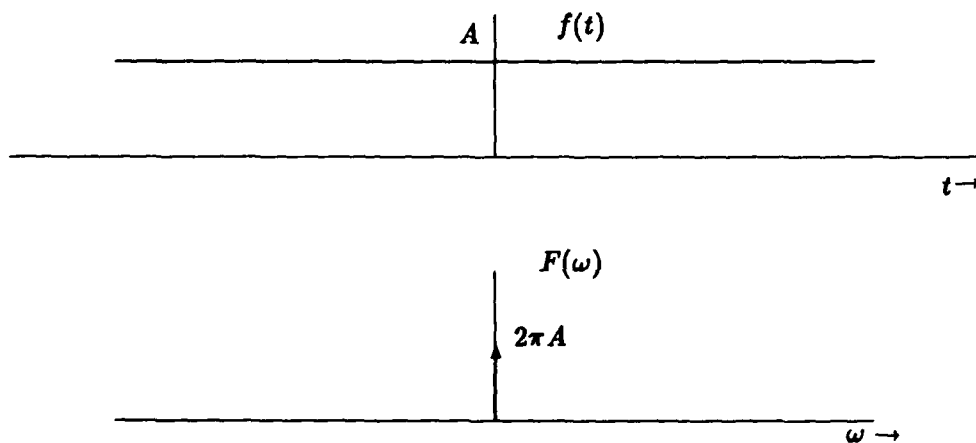
Figure 2.5: Fourier transform pair for $f(t) = A$.
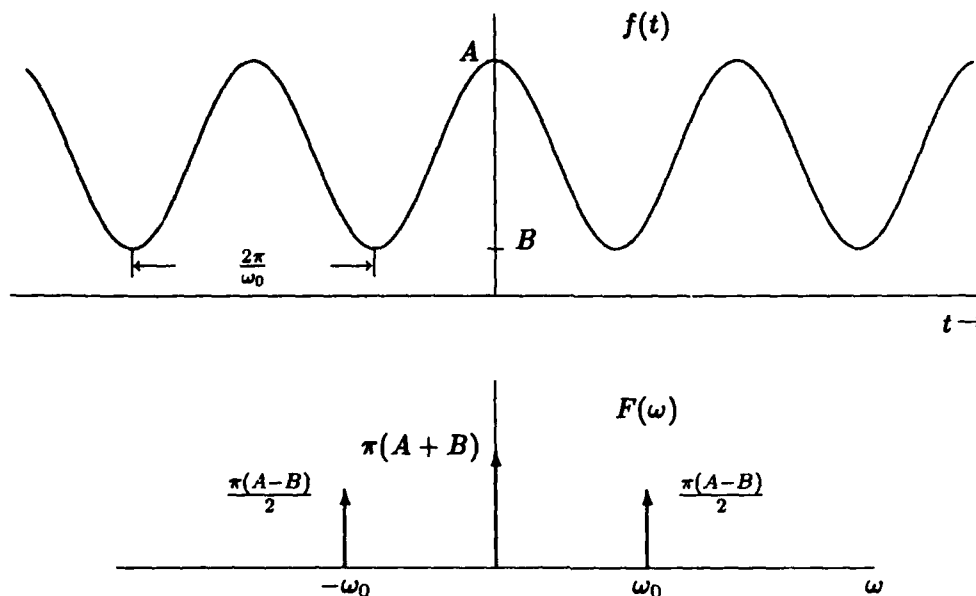
Figure 2.6: Fourier transform pair for a sinusoid of maximum luminance A and minimum luminance B.

transform of the sum of the signals is equal to the sum of the transforms. Thus the frequency domain representation of this signal consists of a delta function at dc of magnitude $\pi(A+B)$ and a pair of delta functions of magnitude $\frac{\pi(A-B)}{2}$ at frequencies of $\pm\omega_0$ (see figure 2.6). The ratio of the magnitude of the signal to the magnitude at dc in the frequency domain is $\frac{A-B}{2(A+B)} = \frac{1}{2}C$. For a signal with a constant mean luminance $\frac{A+B}{2}$, the magnitude of the signal at ω_0 is proportional to the contrast in the spatial domain.

A simple threshold model for the detection of a sinusoidal pattern of constant mean luminance considers the amplitude of a signal in the frequency domain. As derived above, the amplitude of the signal at a frequency of ω_0 is proportional to the contrast of the pattern in the spatial domain. If the contrast of the pattern is decreased a point will be reached at which the pattern is no longer detected by the system, whether it be a human visual system or a video camera. This minimum detectable contrast level defines a threshold in frequency space, for which the pattern is detected if the magnitude of the signal exceeds threshold and the pattern is undetected if the magnitude is below threshold. This is the concept upon which LSA models for visual detection are based. The utility of such models rests upon the superposition principle of the Fourier transform. Because sinusoids constitute an orthogonal set of basis functions for the Fourier transform, it is possible to predict the detection of more complex patterns based on the system's detection thresholds for sinusoidal patterns.

11
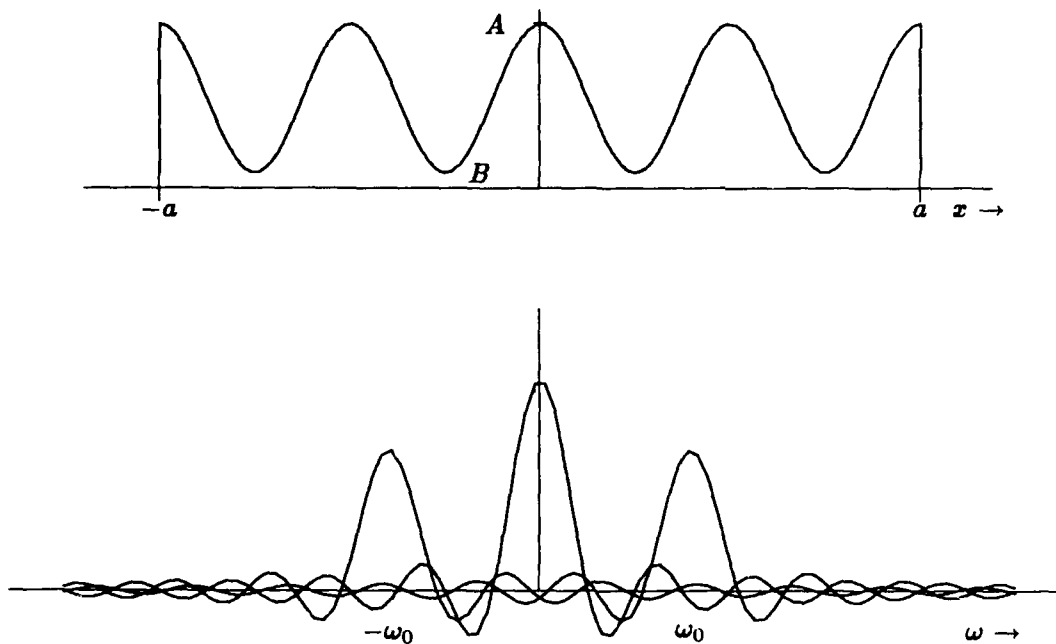
Figure 2.7: Transform pair of a finite extent sinusoid.

For the purpose of illustration, the above derivation used signals of infinite extent. However, these signals do not exist in the real world. A derivation using finite extent signals can also be made, leading to an identical result. The fundamental difference is that the impulse functions in frequency space will be replaced by sinc functions ($\frac{\sin x}{x}$). A straightforward approach for obtaining this result is to use the convolution theorem. A finite extent sinusoid is the product of an infinite extent sinusoid and a pulse, or rectangle function. The Fourier transform of this result is the convolution of the individual transforms. The transform of the rectangle is a sinc, and the convolution of a sinc and a set of delta functions is a set of sinc functions located at the frequencies of the delta functions.

Consider such a finite extent grating that extends spatially from $-a$ to a, as shown in figure 2.7. Its Fourier transform is given by

$$F(\omega) = (A - B) \left[\frac{\sin a(\omega - \omega_0)}{\omega - \omega_0} + \frac{\sin a(\omega + \omega_0)}{\omega + \omega_0} \right] + (A + B)\frac{\sin \omega}{\omega} \qquad (2.6)$$

Since the transform is always symmetric about zero frequency, the remainder of this discussion will deal only with positive frequencies, with the assumption that the magnitude of the signal at the corresponding negative frequency is identical.

In a real system, a finite extent sinusoid grating produces a sinc function at dc and a sinc

12

Figure 2.8: Two point resolution of two intensity distributions.

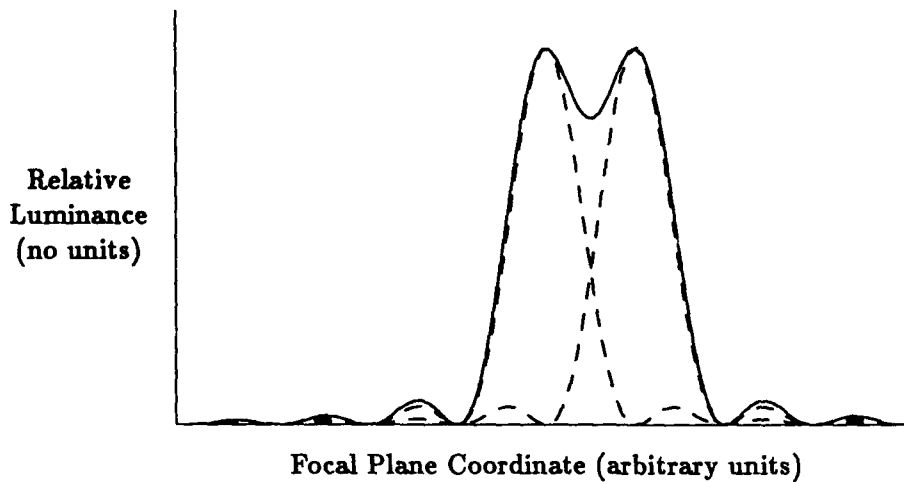function at a frequency of ω_0. The location of the signal due to the sinusoid is determined only by the fundamental frequency ω_0; the height and width depend upon both the contrast and the spatial extent of the signal. Thus for a given contrast, the width, or distribution of the signal in the frequency domain is determined only by the spatial extent, *i.e.* the width of the receptive field, of the signal. This implies the distribution of the frequency domain signals is independent of the spatial frequency ω_0 of the grating pattern.

Because the width of the frequency domain signal is independent of the frequency of the grating pattern, and the difference threshold is a function of the signal width, one would not expect the difference threshold to change as a function of spatial frequency for the model proposed.

Consider two grating patterns of identical contrast and spatial width which differ only in spatial frequency. To define the spatial frequency difference threshold of these patterns, one could borrow a concept from optics used to define the resolution between two points of light. The imaging of two distant point sources by a lens will result in two intensity distributions at the back focal plane of the lens. When the two point sources are separated by a sufficiently small angle, their focal plane distributions will overlap such that the two points are indistinguishable. The amount of overlap at which the two point sources become indistinguishable may be specified by the Rayleigh resolution limit. This limit is derived from the diffraction equations, and equates to a 20% dip from the maximum of the sum of the intensity distributions of the point sources, as illustrated in figure 2.8.

The difference threshold of the system might also be defined in other ways. For example,

one might use the statistics of the signal distributions to determine when the two signals are distinguishable. Another alternative would be to use the width of the signals at half height to determine the frequency resolution.

Regardless of the method, it is clear that the frequency resolution, *i.e.*, the difference threshold, of the signals depends only upon the distribution of the signals. Since the signal distribution is independent of frequency, the difference threshold should also be independent of frequency.

Chapter 3

Implementation and Hardware

3.1   The Hardware System

To implement the threshold model a suitable hardware system was selected that could emulate the basic components of the human visual system. Because of its similarities to the structure of the eye, linear response, and ability to acquire digital image data, a charge coupled device (CCD) array video camera and frame grabber were used to simulate the eye. The action of the neural pathways and visual cortex was simulated by computations performed on this data. These computations were for the most part the application of Fortran programs written specifically for this research.

The hardware system consisted of a *Sierra Scientific* CCD array video camera, a *Zenith* model 248 personal computer, and an *Imaging Technology* frame grabber. Additional computations and analyses were done on a *Digital Equipment Corporation* Vax 8650 computer.

The CCD video camera was a solid state MS-4000 series with a monolithic MOS charge-coupled frame transfer array. The imaging area on the array was 6.0mm horizontal by 4.5mm vertical, although the entire array was slightly larger. There were 610 pixels horizontally and 492 pixels vertically in National Television System Committee standard format. The CCD video camera produced an analog RS-170/330 video signal which was output to the frame grabber.

The frame grabber was a single board that plugged in to an expansion slot of the Zenith 248. It digitized the incoming analog signal from the video camera to an accuracy of eight bits at 30 frames per second and stored the resulting pixels in a frame memory. Each pixel had one of 256 possible gray shades. The frame grabber also had display logic that converted
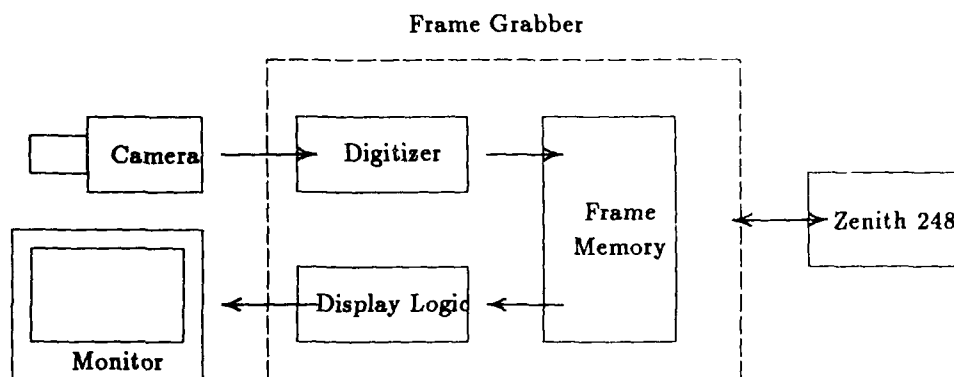
Frame Grabber

Figure 3.1: Components of the Hardware system.

the digital image in frame memory back to analog format for display on a monitor.

The equipment was connected as specified in the *IMAGEACTIONplus* users manual. A block diagram of the hardware components is shown in figure 3.1.

A software package called *IMAGEACTIONplus* was used to save images from frame memory to the computer's disk drives. *IMAGEACTIONplus* also provided many other functions for image manipulation and processing, although few of them were necessary for this application.

3.2 Implementation

The above system was used to implement a LSA model of vision for the contrast threshold and the difference threshold. It was not possible to implement the models exactly as they were described in chapter two. There were several differences between the ideal case, described by mathematical models, and the practical situation, which used electro-optical hardware and digital computers.

First, the implementation employed sampling, which resulted in discrete image data. This differed from the mathematical model, which assumed the spatial signals were continuous. However, it was similar to the human eye which has discrete receptors (rods and cones) to sample image data. The hardware was arranged so its spatial sampling frequency corresponded to the spatial sampling frequency of the eye. A fast Fourier transform (FFT) was implemented to obtain the spatial frequency domain distribution of the target patterns.

Additionally, the implementation was one-dimensional and time invariant. Instead of the target luminance being a function of two spatial dimensions and time $L(x,y,t)$, the

target luminance was only a function of one spatial dimension, $L(x)$. The rationale behind this was twofold. First, if the model was expanded to two dimensions, the complexity of the model both in number of relevant variables and in computational requirements would have increased dramatically. Secondly, a one-dimensional model was adequate to study the concepts of interest and presented fewer implementation problems.

A third difference was in the treatment of noise. The contrast threshold model as presented by Sachs *et al.* is a mathematical model. To account for noise in the system it had to be introduced directly using a random noise function. Since the implementation hardware contained noise of a magnitude comparable to the signal amplitude at low contrasts, it was not necessary to introduce additional random noise.

Chapter 4

Study I: Measuring Contrast Thresholds of the System

4.1 Introduction

The objective of this study was to physically measure the contrast sensitivity function of the hardware system and compare the results with the human contrast sensitivity function and theoretical predictions. Some of the topics discussed include: selection of stimuli, derivation of the apparatus set-up parameters, procedure for data collection, results, and a discussion of the results.

4.2 Producing Stimuli

One of the unanticipated difficulties of this study was producing the test stimuli to use with the hardware system. It was desired to use sine wave gratings of controlled contrast, spatial frequency, and mean luminance level to measure the frequency response of the visual system. This would be a trivial problem if the sine waves were voltage or current signals in the time domain. This was not the case in the spatial domain, where the luminance of the target was to vary sinusoidally in one dimension at a specific spatial frequency and mean contrast level. Producing these stimuli in the laboratory would have been a difficult effort involving a significant expenditure of time and resources, and the commercially available choices of such sine wave grating patterns were limited. Two commercial sources of sinusoidal grating patterns were identified: the Nicolet Optronix CS 2000 Contrast Sensitivity Testing System

18

and the Vistech Contrast Sensitivity Test System (VCTS). Both of these were tested to determine which was better suited for this research.

The Optronix system is a microprocessor controlled black and white monitor used to measure the detection thresholds of human observers. It is set up as a two alternative forced choice device, specifically designed for human testing. The system consists of a control unit, a response unit, and a monitor. The experimenter programs the control unit to measure the desired frequencies. Once the test is started, the subject responds using switches on the response unit indicating whether or not the sine wave pattern presented by the monitor is detected. The computer then adjusts the contrast of the pattern and iteratively queries the observer until a threshold is determined.

The Optronix system was tested with the CCD hardware system, but was not selected for use due to several factors:

- It exhibited temporal aliasing.

- Because it was designed for human testing (two alternative forced choice method) and not for hardware measurements, it had a set program that could not be modified. This program did not lend itself well to hardware measurements.

- The Optronix monitor was subject to luminance variations across its screen and required frequent calibration to ensure good results.

The Vistech VCTS 6500 is a photographically produced chart that presents a series of sine wave gratings at calibrated levels of contrast (see figure 4.1). At the prescribed viewing distance of 10 ft, gratings of 1.5, 3, 6, 12, and 18 cpd are tested. The test patterns are three inch diameter circles containing a sine wave pattern of a specific contrast level and spatial frequency. There are five rows of patterns, each row consisting of one spatial frequency, and nine columns of varying contrast levels.

The Vistech chart was selected for use in this research. Though this system, as well as the Optronix system, did not allow measurements of the upper spatial frequencies that humans can detect, it had certain advantages. Because its target patterns were mounted on a fixed chart, it was convenient to use. It also required no calibration, other than adjusting the light source for uniform illumination, and did not present any temporal aliasing problems. The drawbacks were that it presented a 1.4°target rather than the desired 2°visual field
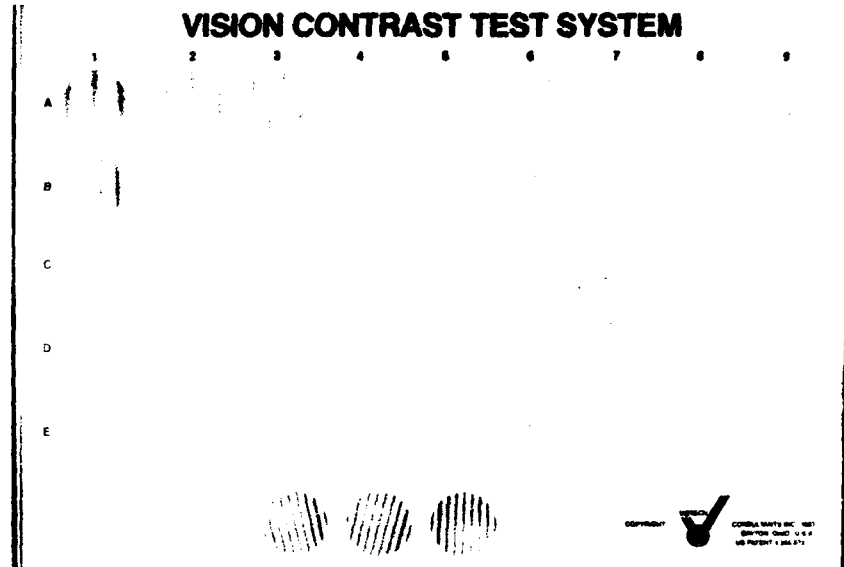
Figure 4.1: Vistech VCTS 6500.

and there was considerable variance in the luminance and contrast profiles of the grating patterns. However, the Vistech charts are apparently widely used and provided a convenient stimulus set for this research.

4.3 Calculation of Set-up Parameters

To simulate the detection scenario of the human vision system with respect to the number of receptors and the size of the stimulus field, it was necessary to calculate some set-up parameters of the hardware system. The objective was to duplicate, with respect to the number of receptors and the field size of the stimulus, a human observing a VCTS 6500 grating pattern at the recommended ten foot viewing distance.

The visual angle Θ that each test pattern subtends can be calculated by the equation:

$$\Theta = 2\arctan(1.5''/10') = 1.43° \tag{4.1}$$

The number of receptors (cones) that this pattern covers on the retina can be calculated knowing the width of receptors in the fovea and the equivalent first order optics of the eye. The width of receptors in the fovea has been physically measured to be about 2.5 μm and a typical human eye lens has an equivalent focal length of about 17mm. Thus the width of the image on the retina of the test pattern is 17mm $\times$ 1.43°= 17mm $\times$ 0.025 rads =

20

0.425mm. The number of receptors across this area is then 0.425mm/2.5micron = 170. This calculation agrees with the known spatial frequency limit of the eye of about 60 cpd. Assuming the eye is efficient and the receptors are spaced at the Nyquist limit, this implies that there are about 120 receptors per degree of visual angle in the fovea. Since the VCTS pattern subtends 1.43°, the number of receptors covered by the pattern would be:

$$120 \text{ receptors/degree} \times 1.43° = 171.6 \text{ receptors} \tag{4.2}$$

To emulate the human visual situation, the hardware must be arranged such that the VCTS test pattern covers 170 detector elements, or pixels, on the CCD array. From the manufacturer's information on the array, the horizontal array width is given as 6.95mm, and the number of pixels across the array as 604. Therefore, the separation between the center of adjacent pixels is 6.95/604=11.51 μm. Thus, 170 pixels covers a width of $170 \times 11.5 \mu$m = 1.96mm. The optical system must be designed to achieve this image size.

4.4 Optical Design

Since the object size is 76.2mm (3 inches) and the desired image size is 1.96mm, a transverse magnification of $1.96/76.2 = 0.0257 \approx 1/39$ is required. The resulting equation to be satisfied is:

$$\frac{1}{i} + \frac{1}{o} = \frac{1}{f.l.} \tag{4.3}$$

where i is the image distance, o is the object distance, and $f.l.$ is the focal length of the lens. For the desired magnification:

$$m = 1/39 = i/o \tag{4.4}$$

Thus $o = 39i$ and equation 4.3 becomes:

$$\frac{1}{i} + \frac{1}{39i} = \frac{1}{f.l.} \tag{4.5}$$

After surveying the lenses available for the camera, a 12.5mm lens was chosen; this lens provided the desired magnification while keeping working distances reasonable. Substituting

this value into the above equation yields:

$$\frac{1}{i} + \frac{1}{39i} = \frac{1}{12.5} \tag{4.6}$$

Solving for the image and object distances:

$$i = 12.82\text{mm} \tag{4.7}$$

$$o = 500\text{mm} \tag{4.8}$$

The previous calculations ensure the image of the VCTS pattern covers the appropriate number of pixels on the CCD array. These calculations, however, do not account for one important factor, which is aliasing. Aliasing may occur if frequencies higher than one half the sampling frequency are passed to the detectors. For the human eye, the lens is generally of low quality and its modulation transfer function (MTF) is bandlimited at about 60 cpd. This matches well the limit of the receptors on the retina, which are spaced about 120 receptors per degree, so no aliasing occurs. However, the MTF of commercial grade lenses typically extends well beyond 60 cpd, so aliasing may be encountered in the current hardware set-up.

To prevent aliasing the optical system must not be allowed to transmit frequencies higher than 120 cpd. This was accomplished by adjusting the aperture of the camera lens so it was diffraction limited at the desired frequency. It was assumed the lens was of excellent quality and approximately diffraction limited. This is usually a good assumption for professional quality television camera lenses, such as the one selected. The resolution of a diffraction limited lens is determined by the width of the effective aperture. The angular resolution α is approximated by:

$$\alpha = \frac{5.5}{w} \text{ seconds of arc} \tag{4.9}$$

where w is the diameter of the lens in inches. Equation 4.9 is derived from the diffraction equations for a circular aperture optical system [18]. To pass only frequencies less that 60 cpd, the desired resolution is 1/60 degree = 60 arc seconds. Substituting this value into equation 4.9 yields:

$$60 = \frac{5.5}{w} \tag{4.10}$$

$$w = 0.0917 \text{ inches} = 2.33 \text{mm} \tag{4.11}$$

For a 12.5mm lens, this yields an $f/\#$ of $12.5/2.33 = 5.4$. The nearest available $f/\#$ on the lens aperture control was 5.6, which gave an aperture width of 2.23mm and a corresponding resolution of 62.6 arc seconds = 57.5 cpd. Thus, the transmission of the lens was sufficiently bandlimited to avoid aliasing the information gathered at the CCD array.

After mounting the hardware using the specifications derived above, the set up was verified using the "line measure" command in the *IMAGEACTIONplus* software. The number of pixels across a VCTS test pattern was measured several times, with results ranging from 169-171. This verified the desired image size was achieved by the optical set-up.

4.5 Data Collection

The VCTS chart was mounted vertically on an optical table. The camera was mounted on an optical rail 500mm from and parallel to the Vistech chart. This allowed for the camera to be moved horizontally across a row of the Vistech chart while maintaining the correct object distance. Height adjustments were made by moving the camera up or down in the rail-mounted base.

The camera was positioned so the desired target pattern was located in the vertical center of the field and centered in the left half of the horizontal field (x pixels 1-256). Then the cursor was positioned in the center of the target to determine its x, y coordinate. The line save command of the *IMAGEACTIONplus* software was used so that a horizontal line passing through the center of the pattern was saved and written to the computer's hard disk. If the grating pattern was tilted at a 15°angle, then the camera was also tilted at the same angle such that the grating patterns always remained vertical with respect to the camera array. Flood lamps were used to illuminate the test chart. The lamps were adjusted to yield a uniform luminance of 100 foot-lamberts on the area of the chart surrounding

the subject grating pattern. The luminance of the chart was measured using a hand held luminance meter.

4.6   Processing the Image Data

After taking a series of data files on all the target patterns, these files were uploaded to a Digital Equipment Corporation VAX 8650 for further processing and analysis.

Fortran programs were written to perform the desired calculations on the raw data. The first program to operate on the files read in the files created by the *IMAGEACTIONplus* software and extracted the desired image data. The program then performed a one-dimensional 256 element Fast Fourier Transform (FFT) on the extracted line data. The magnitude of the FFT and a stripped down version of the image line data were then written to disk.

Recall that there were only 170 samples of the grating pattern that were obtained from the CCD camera and frame grabber. To perform a FFT, 2^n samples are required. The nearest numbers meeting this criteria were 128 and 256. Possible alternatives to convert the 170 samples to a number of samples suitable for a FFT were truncating the 170 samples down to 128 or copying part of the 170 samples and adding them to the original samples to obtain 256 samples. Truncating to 128 samples would correspond to observing a 1.1°FOV, since $128 \times 1.43°/170 = 1.1°$. Alternatively, copying out to 258 samples would correspond to a 2.2°FOV, since $256 \times 1.43°/170 = 2.2°$. Since the angular width of the fovea is about 2°, and the intent of this research was to simulate foveal detection, the latter alternative was selected. In implementation, the center 128 of the 170 samples were copied and added to the original 128, to obtain 256 samples with which to perform the FFT.

Several plots of data files are shown in figure 4.2. Note how the magnitude of the signal due to the grating pattern decreases as the contrast of the pattern decreases.
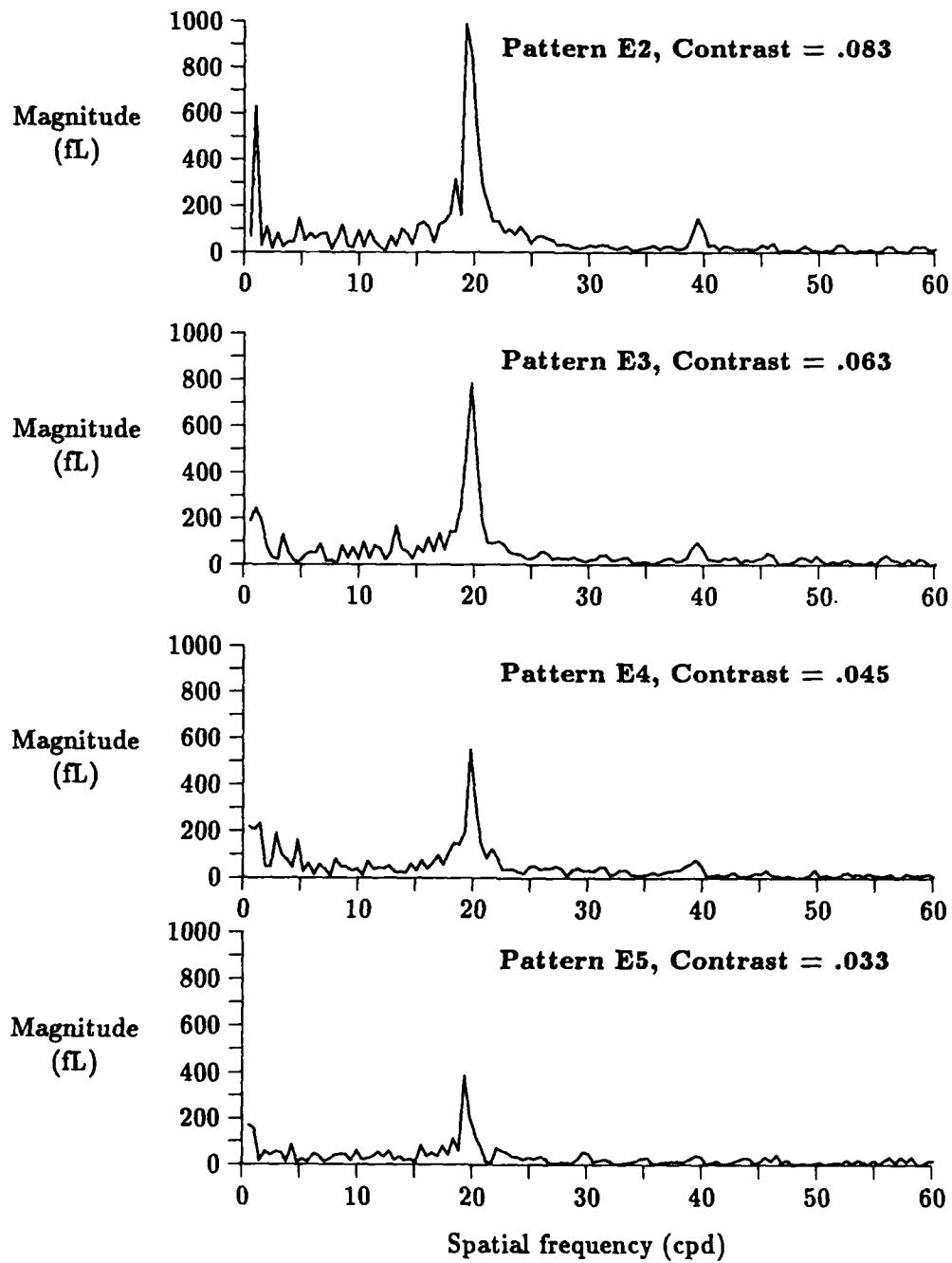
Figure 4.2: Frequency plots of several grating patterns.

4.7 Determining the Contrast Thresholds

The task in determining the contrast thresholds for each of the different frequencies was to decide when the spike-like signal due to the sine wave, which decreased in magnitude as the contrast of the grating decreased, was no longer distinguishable from the noise in the frequency domain. Thus the contrast threshold was not only dependent upon the signal magnitude but also the system noise. There are several possible approaches to defining the contrast thresholds; these mainly arise from the different ways in which the noise may be treated.

If the grating patterns were perfect sine waves and if the hardware system contained no noise, our data would be statistically uninteresting. The transforms of the sine waves would be precise and there would be no noise in the data files, so the signal would always be distinguishable down to a contrast of zero.

However, since the data were collected in real world conditions, it had an abundance of noise due to two primary sources. First, the grating patterns were not sinusoidal in luminance. Though each line was approximately sinusoidal, there were variations due to the grain of the film and nonuniformity of the master printing pattern that caused the luminance profile as measured on one horizontal line of the target to differ from the luminance profile as measured on another horizontal line.

Another source of noise in the data was electrical noise. This was primarily radiated and conducted noise that could be introduced almost anywhere in the system, from the CCD array to the video cable or the digitizing hardware.

The total noise in the system resulted in a variance of these data and in uncertainty in the determination of contrast thresholds. It was desirable to reduce the effect of this variance to obtain more repeatable results. One technique for accomplishing this is to average over a number of runs.

Like any sampled data set, the magnitude of the frequency domain noise may be characterized by its mean $\bar{X}$ and standard deviation σ.

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad (4.12)$$

and the standard deviation σ is:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i^2 - \bar{X}^2)} \qquad (4.13)$$

From equation 4.13 it is apparent that σ is proportional to $\frac{1}{\sqrt{n}}$. As the number of samples increase, the standard deviation (and variance) of the noise decreases. By using 10 data samples as opposed to one, the standard deviation of the noise is reduced by a factor of $\frac{1}{\sqrt{10}}$, which equals a $20\log(\frac{1}{\sqrt{10}})$, or a -10 dB reduction. With the variance of the samples decreased, a more repeatable determination of the contrast threshold can be made.

To illustrate this concept, compare the transforms of figure 4.3. Figure A is the transform of a single line of data from target pattern C7, and figure B is the transform of the average of 10 data lines from the same target. In both figures, the signal is present at approximately six cpd. This is readily apparent in figure B whereas in figure A the signal is not distinguishable from the noise. The mean value of the noise in both figures is about the same, but the standard deviation of the noise in figure B is 10 dB less than that of figure A. Because of the reduced standard deviation, the signal in figure B is distinguishable from the noise with a higher probability than in figure A.

Note also the signal that occurs at approximately 45 cpd. This unknown artifact may be a harmonic or aliased temporal frequency from the frame grabber or perhaps originate from one of the oscillators in the Zenith computer. Whatever the source of this signal, it was consistent throughout the data collection and did not pose a problem in the data analysis.

Averaging 10 runs of data to reduce the variance of the noise was accomplished to increase the probability that a repeatable contrast threshold was determined. However, we still have yet to define the criterion for distinguishing the signal from the noise.

A nonparametric approach was used to establish the contrast thresholds for the system. This approach relies on the basic statistics of the data and does not require fitting the data to a known distribution function. In the nonparametric approach, a histogram of the noise is first plotted. Then the histogram is used to determine the percentage of the noise that is less than or equal to each noise magnitude. This percentage, sometimes called the empirical distribution function, $F_n(x)$, is plotted as a function of the magnitude, as shown in figure 4.4. The use of the empirical distribution to determine the contrast threshold is best explained using the following example.

Figure 4.4 shows the noise distribution for target pattern C8. On the ordinate is the
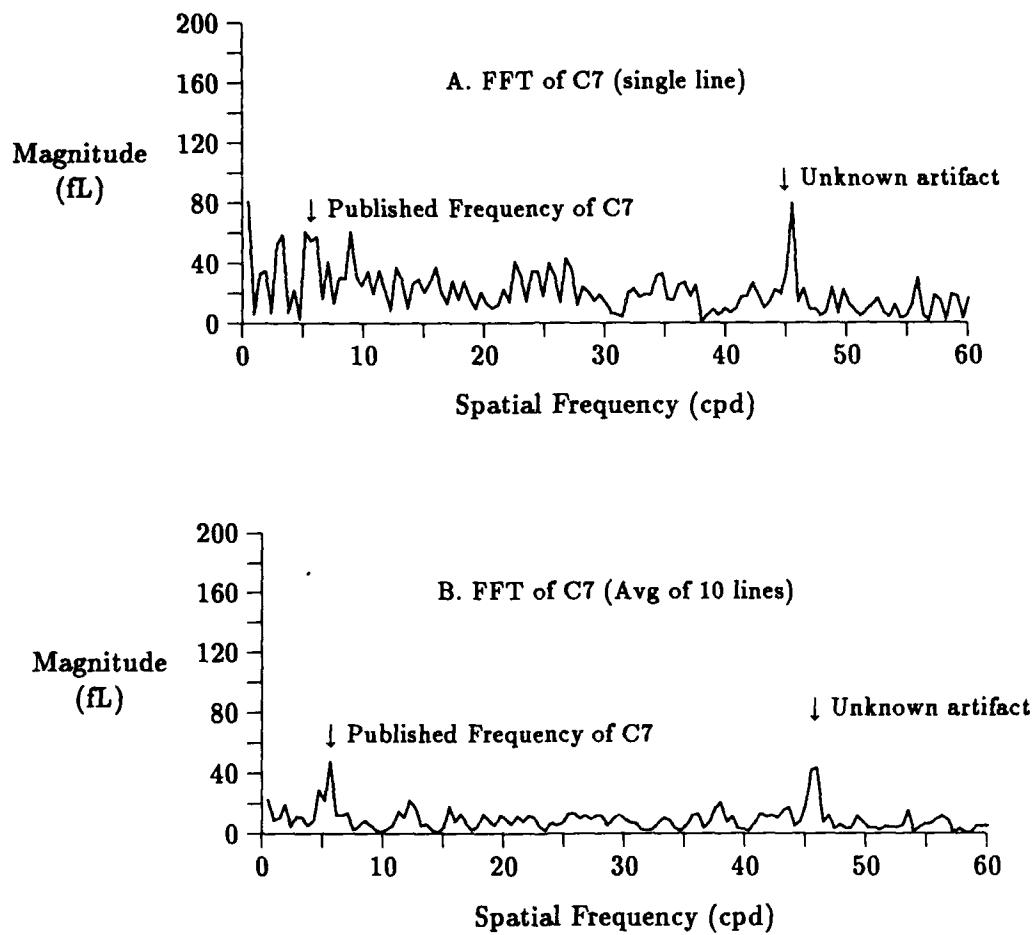
27

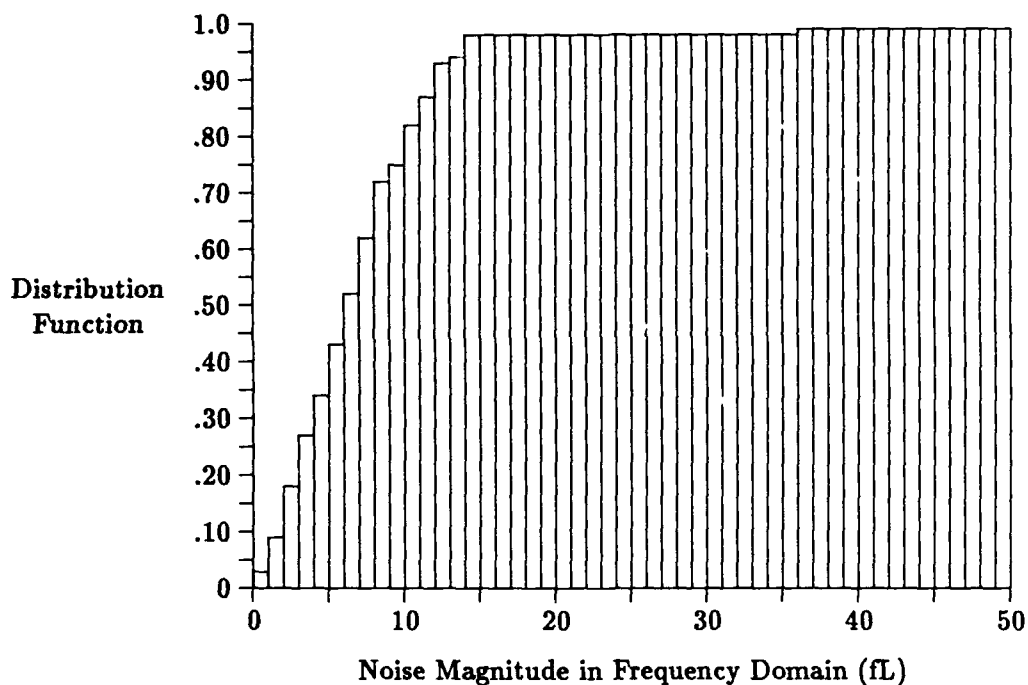Figure 4.3: Fast Fourier Transforms of grating pattern C7.

Figure 4.4: Empirical distribution function of the noise for pattern C8.

percentage of the noise that falls at or below the corresponding abscissa value. Approximately 26% of the noise falls at or below a magnitude of 4, and 98% of the noise data is less than or equal to 20. Though this graph actually shows the percentage of noise samples that are less than or equal to each magnitude, it also allows one to specify a contrast threshold with a known probability that the data which exceed the threshold are signal and not noise. The relative frequency of noise samples that are less than or equal to a given magnitude is equal to the probability that a sample of the same magnitude is signal and not noise. Using the same example as above, if the contrast threshold is defined as 4, then there is only a 26% probability that this value is signal and not noise. If the contrast threshold is defined as 19, then we can specify with a 98% probability that any data with a magnitude of 19 are signal and not noise. Conversely, one could first select the desired probability for the threshold determination, and then use this graph to find the corresponding contrast threshold. The latter method was used to determine the contrast thresholds of the hardware system.

4.8   Results

Using the previously described method, a probability of 99% was used to specify the signal threshold. According to this criterion, the last detectable signals occurred in grating pat-
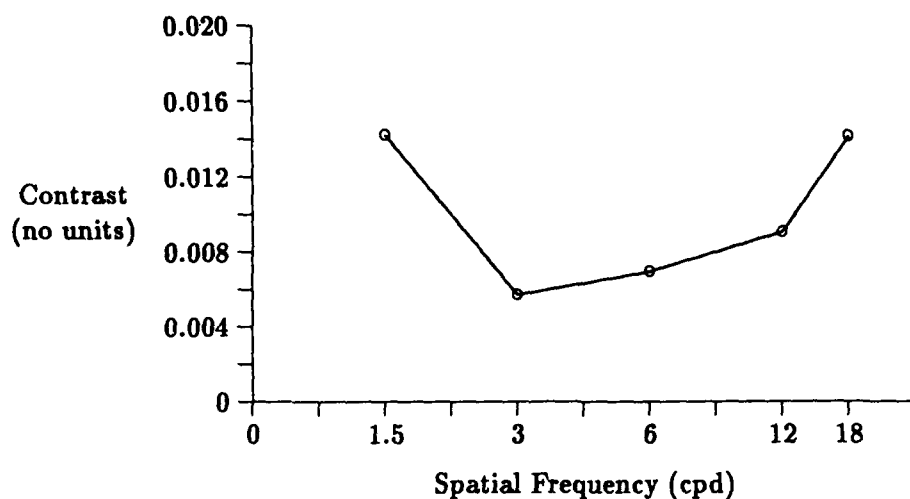
29

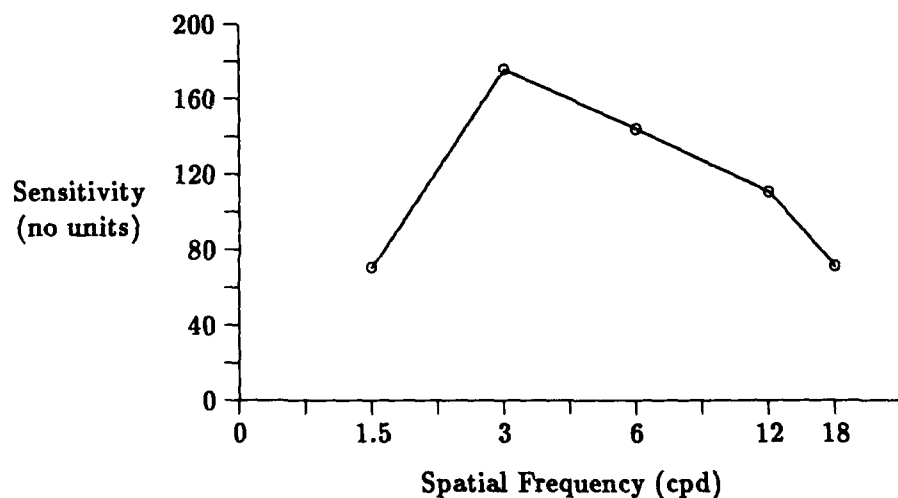Figure 4.5: *Contrast threshold function for the hardware system.*

Figure 4.6: *Contrast sensitivity function for the hardware system.*

terns A6, B8, C7, D7, and E7, which correspond to spatial frequencies of 1.5, 3, 6, 12, and 18 cpd, respectively.

The contrast of each of these five patterns was calculated. Because of the variance of the gratings, every cycle in a pattern did not have the same contrast. Therefore, contrast for each cycle within a pattern was determined, and the average of the maximum and minimum contrasts found within the grating pattern was used as the contrast threshold. These contrast thresholds for the five spatial frequencies, which ranged from 0.5–1.5%, were then plotted as a function of spatial frequency as shown in Fig 4.5.

These data were also plotted as contrast sensitivity (the reciprocal of the contrast threshold) versus spatial frequency to obtain the hardware's CSF, shown in figure 4.6. When the
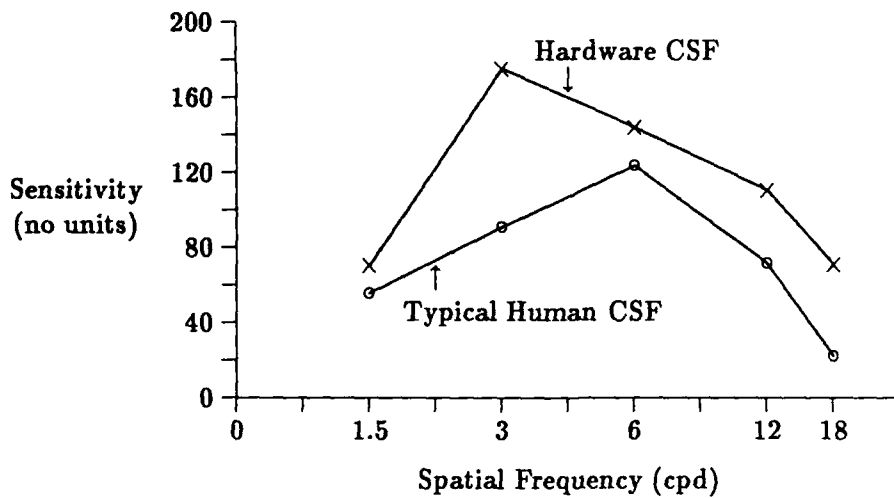
30

Figure 4.7: Contrast Sensitivity Function of the hardware system compared with that of a typical human observer.

CSF of the hardware system is compared to a typical human CSF (see figure 4.7), several factors are noticeable. First, the shape of the characteristic curves is very similar. Intuitively this might be expected because of the similarities between the two systems. Secondly, the contrast sensitivity is generally higher for the hardware system, indicating that it can detect lower differences in contrast than the average human. Finally, the hardware system is most sensitive to signals about 3 cpd, near the frequencies to which humans are most sensitive.

4.9   Discussion

As observed above, the CSF of the hardware model is similar to that of the human population. Though the model has slightly better sensitivity than the human data show, it preserves the same characteristic curve shape and the same relative sensitivity with respect to frequency. Since this curve is observed for both the human and the hardware system, it is desirable to know if Fourier theory as applied in the threshold model can explain the shape of this curve.

The CSF of the human or the hardware represents the frequency response of the entire system, which is made up of several components. According to the Fourier theory, the frequency transfer function of a system that is comprised of several components in cascade is the product of the frequency transfer functions of the individual components. Thus the observed CSFs of these systems are a product of the transfer functions of their respective
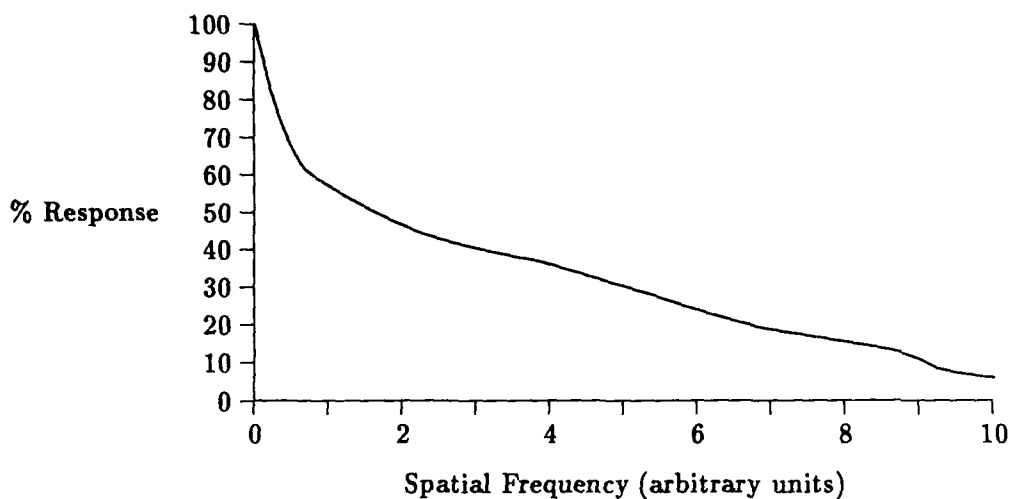
31

Figure 4.8: Typical Modulation Transfer Function of a lens.

components.

It is convenient to consider the hardware system as having two basic subsystems: optical and electrical. The optical subsystem is comprised of a single optical element, the camera lens. The frequency transfer function for a lens is called the modulation transfer function, or MTF. The MTF indicates how well the lens transmits the contrast of an image as a function of spatial frequency. A typical MTF is shown in figure 4.8. In general, the MTF of a lens will monotonically decrease as frequency increases. Though the exact MTF of the lens used in this implementation was not measured, it is unlikely that the shape of its MTF is different from that shown in figure 4.8.

The electrical subsystem includes those components which detect, transmit, and make judgements on the image data. These components are the CCD array camera, the frame grabber, and the application of the threshold criteria.

First consider the transfer function of the detector. The CCD array samples a finite portion of the image, acting as a transducer that converts light energy to an electrical current. In terms of basic mathematical functions, the CCD array may be represented by the convolution of a pulse and a comb function. The pulse represents the constant gain conversion of a finite part of the image's light energy distribution to an electrical signal, and the comb represents the sampling of the image by the pixels. In the frequency domain the convolution of these two signals is equivalent to the product of a sinc and a comb function. However, since the sampling within the pulse is high (every $11.51\mu m$) relative to its width (1.96mm), the spacing of the sinc functions in the frequency domain is very large relative
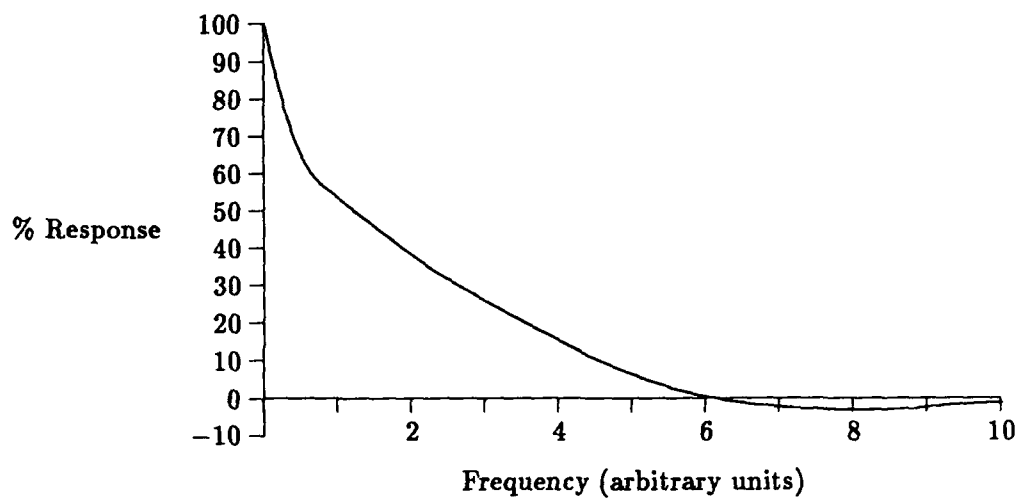
32

Figure 4.9: Theoretical Hardware System Transfer Function.

to the frequency content of the sinc function, and the contributions of the sinc functions at higher frequencies may be ignored. Thus the frequency transfer function of the CCD array is simply a sinc function located at dc.

The frame grabber is the means for acquiring data from the CCD camera. It digitizes the image signal from the camera and stores the digital data in the frame memory. Hence, it might be thought of as a one-to-one mapping of the image information from the CCD to memory, with a small amount of quantization error introduced. The mathematical function which reproduces the original function through convolution is the impulse, or delta function. Since the Fourier transform of the delta function is a constant in the frequency domain, the transfer function of the frame grabber may be approximated as a constant-gain all-pass filter.

When the MTF of the lens is multiplied by a sinc function (transfer function of the CCD) and a constant function (transfer function of the frame grabber) the result is a system transfer function as shown in figure 4.9. (Note that this is a qualitative analysis and the frequency units are purely arbitrary.) It is instructive to compare this system transfer function with the CSF of the hardware system, plotted with a linear frequency scale for direct comparison (figure 4.10.

The system transfer function only accounts for the transfer functions of the hardware components, whereas the CSF accounts for the transfer functions of the hardware components and the application of the detection criteria. Note the system transfer function is similar to the hardware CSF, and also the human CSF, in its decreasing response at higher
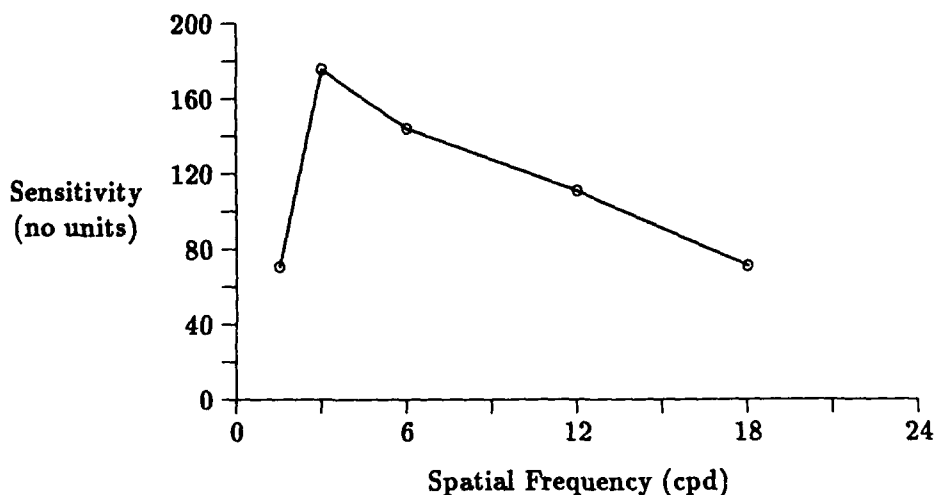
33

Figure 4.10: Contrast Sensitivity Function of the Hardware System plotted with a linear frequency scale.

frequencies, but differs in that it does not contain the characteristic low frequency roll-off that the human and hardware CSFs exhibit. One might conclude that this difference, the low frequency roll-off is due to the application of the detection criteria.

The low frequency roll-off of the CSF indicates the effective noise of the system is somehow increased for these very low frequencies. This increase in effective noise might be due to the way in which the threshold function is applied. Because of the finite size of the receptive field, both the signals due to the dc level of the target and the sinusoid are sinc functions in the frequency domain. When the target is a high frequency sinusoid, these *sinc functions* are separated in frequency and there is little contribution to one from the other. Thus, the signal due to the sinusoid must only be discerned from the electrical noise of the system. As the frequency of the target is lowered, the sinc due to the sinusoid moves toward the dc sinc. When the frequency of the sinusoid is sufficiently low, parts of the two signals occupy the same frequencies, and the sinc due to the sinusoid is harder to discern. In essence, the dc signal is considered part of the system noise.

This concept is illustrated in figure 4.11. Figure 4.11 shows a sinc function due to the mean luminance of the grating along with a random noise function which represents the electrical noise in the system. At high frequencies, the magnitude of the sinc function is very small, so any signal that is present must only exceed the noise to be detected. At frequencies nearer to dc, however, the sinc function increases in magnitude. For a signal to be detected in this region, it must have a larger magnitude than the dc sinc. Thus, as
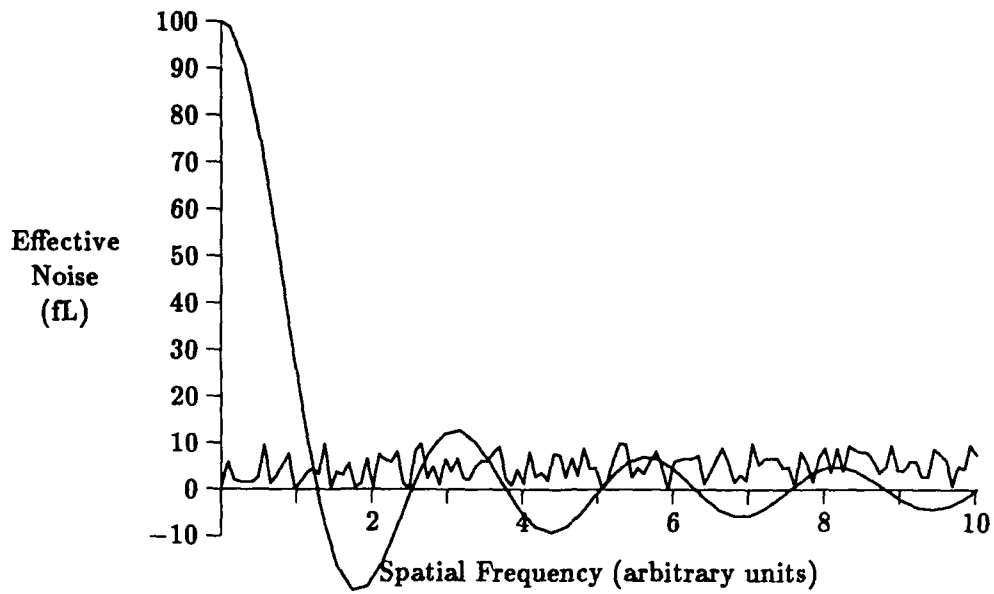
Figure 4.11: Illustration of Effective Noise.

a sinusoid is lowered in frequency, the effective noise in the frequency domain is greater, resulting in a higher contrast threshold, or a lower contrast sensitivity. As a sinusoid is increased in frequency, the attenuation due to the MTF of the lens takes effect, lowering the contrast sensitivity for high frequencies as well.

Chapter 5

Study II: Characterizing the Spatial Frequency Difference Thresholds of the System

5.1 Introduction

The objective of this study was to determine the characteristics of the hardware system's spatial frequency difference threshold function according to the fixed field LSA model presented in chapter 2. From the analysis of the Fourier theory for difference thresholds in section 2.3, this may be accomplished by examining the distributions of the targets' Fourier transforms.

Figure 5.1 shows the spatial frequency distributions of two finite extent sinusoids of slightly different frequency. The ability of an observer to distinguish between the two sinusoidal grating targets is inversely proportional to the amount of overlap of the transforms of the two sinusoids. For two targets whose fundamental frequency components do not overlap, one can discern that the two targets occur at a different spatial frequency. When the difference in spatial frequency between the gratings is small such that their transforms overlap a sufficient amount, then one cannot discern a difference in frequency. Several possible criteria could be used to establish the amount of overlap which corresponds to the difference threshold. For this research, it really didn't matter which of these criteria was chosen because the objective was to determine how the difference threshold behaved and

Frequency (arbitrary units)
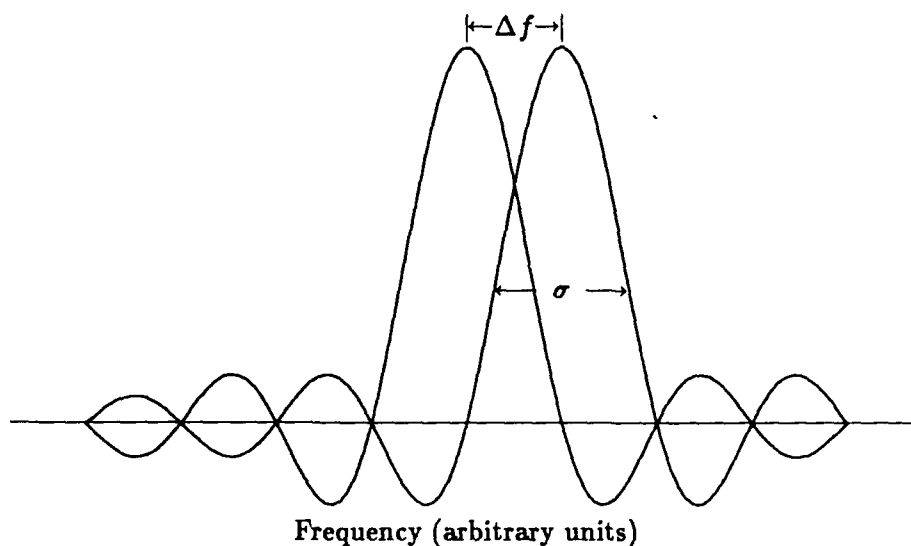
not to determine the absolute difference threshold function. Because of its common use as a measure of distribution width and its ease of calculation, the standard deviation σ of the frequency of the signal was selected as a measure of the signal distribution for this study.

An example of a specific criterion for the difference threshold is that two gratings are distinguishable if the centers of the distributions of their transforms in the spatial frequency domain are separated by at least 2σ. This criterion is proposed as an example; one could also have specified $1.5\sigma, 2.3\sigma,$ or 3σ separation. As long as the criterion is a linear function of the signal's distribution, the exact definition of the difference threshold is relatively unimportant.

Given that a linear function of the signal's distribution is used as the distinguishing criterion, the difference threshold Δf may be defined:

$$\Delta f = m\sigma + b \tag{5.1}$$

where m and b are constants to define the difference threshold criterion. Since Δf is a function of σ, determining the relationship between σ and f will indicate the nature of the relationship between Δf and f.

5.2   Stimulus Material

The stimulus material used to investigate the difference threshold was the Vistech VCTS 6500, the same as used in determining the contrast sensitivity function in Study I. In study I all of the sinusoidal patterns of the VCTS 6500 were used. In study II, however, only one of the grating patterns was necessary.

In the difference threshold studies by Hirsch and Hylton and Richter and Yager, suprathreshold targets were used. To duplicate these conditions as closely as practical, grating E2 was selected as the target for this study. E2 was selected because of its relatively high spatial frequency and high contrast. The E row of the VCTS provided the highest frequency targets, and the lower frequencies could be generated by increasing the magnification of the optical system. Target two had a contrast of 8%.

5.3   Hardware set-up

The hardware was set up with the VCTS mounted vertically on an optical table and the camera mounted on an optical rail one meter from the chart. The camera was equipped with a 12.5–75 mm zoom lens. The zoom lens permitted the magnification, and hence the observed spatial frequency of a grating pattern to be increased or decreased without changing the position of the apparatus. Thus grating targets of the same contrast but different spatial frequencies were generated using a single pattern on the VCTS 6500. This ensured the contrasts of the targets remained constant and allowed targets with small differences in frequency to be easily produced.

5.4   Data Collection

The first step in the data collection was to adjust the camera lens for the correct back focal distance so the image of the grating pattern would remain in focus throughout the zoom range of the lens. If the back focal distance was not correctly set the focal plane of the lens would shift as the magnification was changed. After the back focal distance was adjusted, flood lights were positioned to provide uniform lighting of the grating target such that the luminance of the white area of the VCTS surrounding the E2 target patterns was 100 fL ± 5%. The focal length of the lens was set to 25 mm so the size of the grating pattern

38

covered approximately 170 horizontal pixels. At this setting the camera was positioned to place grating E2 in the center of the left half of the FOV. Then ten consecutive line images from the center of the grating were acquired using the *line save* command. This procedure was repeated five more times, at focal lengths of 29, 33, 40, 50, and 70 mm. Thus sixty images were acquired, ten at each of six frequencies ranging from approximately 7–20 cpd.

5.5 Data Analysis

A FFT was performed on each of the sixty image lines using the same technique as in chapter 4. To extract the signal due to the sine wave a threshold equal to the highest magnitude noise component was applied to the transform data. The parameters of these data points above threshold (lens setting, repetition, frequency, and magnitude) were then placed in file for statistical analysis.

The statistical analysis was performed on the VAX 8650 using the Statistical Analysis System software. For each lens setting, the standard deviation of the frequency variable was calculated. These data were plotted as a function of the mean frequency and a Pearson correlation of the standard deviation and mean frequency was performed.

5.6 Results

The data showing the relationship between the standard deviation of the frequency domain signal due to the sinusoid and the mean frequency of the sinusoid is shown in figure 5.6. There was no apparent trend in the plot; the data appears random with a large variance. A correlation analysis was performed to determine if there was a significant relationship between standard deviation and mean frequency. The Pearson test revealed a correlation coefficient of -0.411. The significance of the test was $p = 0.418$.

5.7 Discussion

Since the Pearson test was not significant there was no correlation between the distribution of the frequency domain signal and the mean frequency of the sinusoid. This was in agreement with the theoretical analysis which showed the distribution of the frequency domain signal should only depend upon the width of the receptive field. Since the width of
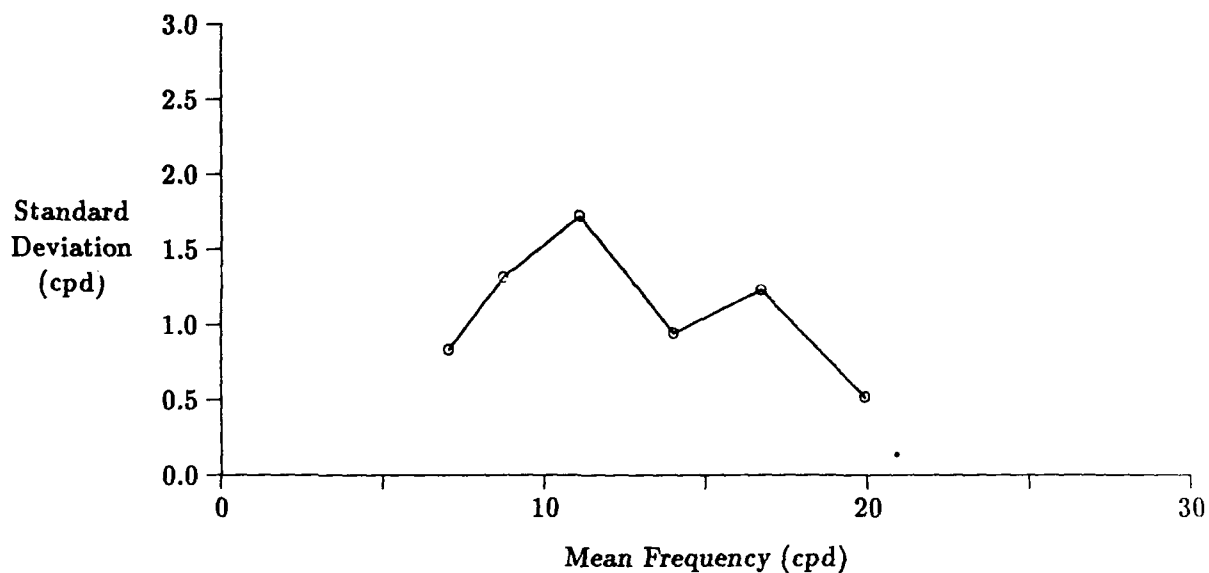
Figure 5.1: Standard Deviation of the frequency domain signal as a function of the mean frequency.
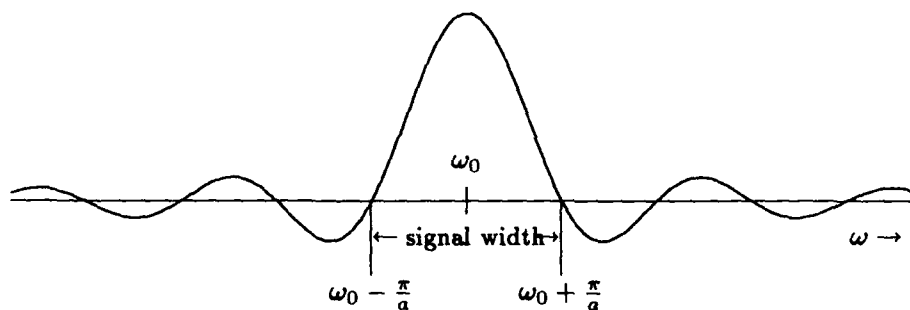

Figure 5.2: Width of a frequency signal.

the receptive field was constant throughout the data collection, there was no discrepancy between Fourier theory and the experimental data. As predicted, the model did not accurately represent the difference threshold of the human visual system. Thus the fixed visual field LSA model for the difference threshold agreed with the hardware system data, but not with human data.

Since the proposed model did not account for the difference threshold function of the human visual system, how might this model be altered so that it does agree with the human data? Recall from chapter two that the distribution of the Fourier transform of a sinusoid of constant contrast and mean luminance depends only upon the extent of the signal, or the width of the receptive field. Consider the finite sinusoid of figure 2.7. If the width of the distribution is chosen as the distance between the zero crossings of the sinc function, as shown in figure 5.2, then this width is $\frac{2\pi}{a}$. Thus the width is inversely proportional to

40

the field size $2a$. To obtain a model where the difference threshold increases linearly with frequency, as the human data show, the width of the signal in frequency would have to increase linearly with frequency. This would require the width of the receptive field, W_f, to decrease linearly with increasing spatial frequency. Thus: $W_f \propto \frac{1}{f}$. One means of obtaining this relationship is to specify that the receptive field observes a fixed number of cycles of the sinusoid. If, for example, the receptive field viewed five cycles of a pattern, then the field width would be $\frac{5}{6}$°for a 6 cpd pattern and $\frac{5}{18}$°for an 18 cpd pattern. Because the field size for the higher frequency pattern is smaller, its Fourier transform will be spread out farther in the frequency domain, resulting in a larger difference threshold. The inverse relationship between the extent of a signal and its transform is illustrated using the pulse in figure 5.7.

There is a physiological basis for the model where the receptive field varies inversely with frequency. Figure 2.2 showed the distribution of rods and cones in the human eye. Note that density of cones decreases with distance from the fovea. To obtain proper sampling of an image the receptors must be spaced at or above the Nyquist limit. Thus the area of the retina that properly samples the highest frequencies is the fovea, where the density of cones is greatest. Lower frequencies, which require a lower density of receptors, could be properly sampled over a larger retinal area, allowing the width of the receptive field to increase for these frequencies. This suggests that the receptive field size for a given frequency be defined by the area of the retina for which the receptors are spaced at or above the appropriate Nyquist limit.

In summary, the proposed Fourier model for the difference threshold was validated by the hardware system, but did not account for the human difference threshold function. It was not necessary to measure the difference threshold of the hardware system to obtain this result; only the relationship between the width of the signal in the frequency domain and the mean spatial frequency had to be determined.

A modified Fourier model where the width of the receptive field varied inversely with the frequency of the target was proposed to account for the human difference threshold. According to theory, this model should exhibit a difference threshold that increases linearly with frequency.
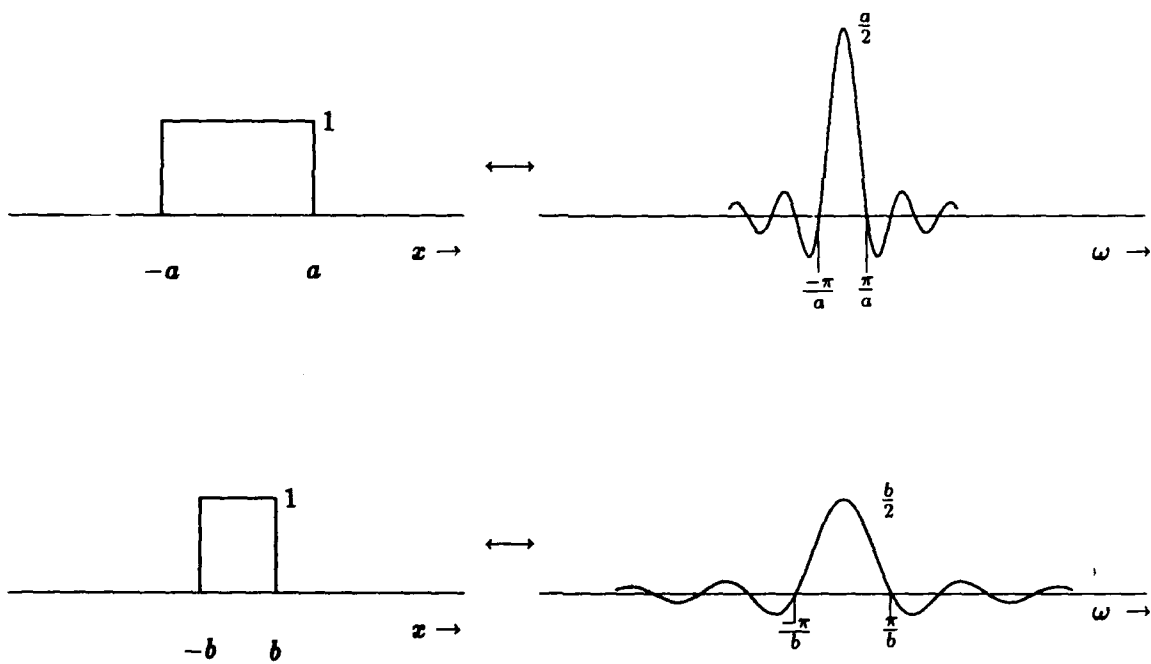
Figure 5.3: Fourier transform of two pulses of different widths.

Chapter 6

Conclusions

6.1 Summary of Results

This research investigated the application of a LSA model to two visual parameters: the contrast threshold and the spatial frequency difference threshold. A one-dimensional model with a fixed visual field size was developed and implemented with hardware.

For the contrast threshold, the results of the implementation agreed with the human data and the Fourier model. An explanation was offered to account for the shape of the CSFs of both systems.

For the difference threshold, the results of the implementation did not agree with the human data, but did correspond to the theoretical model. This implied that the difference threshold of the human visual system could not be simulated with the one-dimensional fixed visual field model. A modified model incorporating a visual field that varies inversely with the spatial frequency of the target pattern was proposed as a way to account for the human difference threshold data.

6.2 Recommendations

This research has revealed a number of areas which may be further investigated. One possible study might verify the explanation of the CSF's low frequency roll-off that was given in chapter four. According to the explanation given, the sinc function at dc is considered as system noise for purposes of detecting a sinusoidal grating pattern. Typically the magnitude of the sinc at dc is much larger than the mean noise level. This accounts for the increased

contrast threshold or lower contrast sensitivity at very low frequencies. As the mean noise level of the system is increased one would expect the CSF to be lowered and the amount of roll-off to decrease. An experiment employing a hardware system similar to that used in this research could determine the effects of random noise levels on the CSF.

Another possible experiment, similar to the previous one, would investigate the effect of mean luminance levels on the CSF according to the fixed field linear system model. For lower mean luminance targets the magnitude of the transform at dc would be less, and thus one would expect the CSF should exhibit less roll-off at low frequencies.

Other studies could be performed to replicate various human studies that have been done. One candidate study is Campbell and Robson's investigation of the detection and discrimination of simple grating patterns [5]. This experiment could be performed using the hardware system instead of the human observer and applying the contrast thresholds obtained from this thesis. One would expect that the various grating patterns (square, triangle, sawtooth, etc.) would not be discernible from a sinusoidal pattern until the higher harmonics in their Fourier transforms had exceeded threshold.

A logical follow-on to Study II of this research would be to implement the modified difference threshold model that was proposed. This follow-on could use the same hardware and types of analyses used in this thesis. The results of this experiment would indicate how well the modified difference threshold model represents the human difference threshold.

Finally, one could develop two dimensional models for the contrast threshold and spatial frequency difference threshold. This research might build upon the models presented in this thesis and expand them to account for directional sensitivity (target orientation) and other factors that are introduced by the expansion into two dimensions.

Bibliography

[1] *Sierra Scientific Solid State Video Camera MS-4000 Series operation and Maintenance Manual.* May 1987.

[2] *VCTS application Manual.* 1987.

[3] Kenneth R. Boff, Lloyd Kaufman, and James P. Thomas, editors. *Handbook of Perception and Human Performance.* John Wiley and Sons, New York, 1986.

[4] F.W. Campbell, Jacob Nachmias, and John Jukes. Spatial frequency discrimination in human vision. *Journal of the Optical Society of America,* 555–566, 1970.

[5] F.W. Campbell and J.G. Robson. Application of fourier analysis to the visibility of gratings. *J. Physiol. (London),* 197,551, 1968.

[6] Stanley Coren et al. *Sensation & Perception.* Academic Press, Inc., New York, 1984.

[7] Tom N. Cornsweet. *Visual Perception.* Academic Press, New York, 1970.

[8] Clarence H. Graham, editor. *Vision and Visual Perception.* John Wiley & Sons, Inc., New York, 1965.

[9] Joy Hirsch and Ron Hylton. Limits of spatial frequency discrimination as evidence of neural interpolation. *Journal of the Optical Society of America,* 72(10), October 1982.

[10] Robert V. Hogg and Allen T. Craig. *Introduction to Mathematical Statistics.* Macmillan Publishing Co.,Inc, New York, 1978.

[11] Rudolf Kingslake, editor. *Applied Optics and Optical Engineering.* Volume 1, Academic Press, New York, 1967.

[12] Rudolf Kingslake, editor. *Applied Optics and Optical Engineering.* Volume 4, Academic Press, New York, 1967.

[13] Leslie Lamport. LaTeX *Users Guide & Reference Manual.* Addison-Wesley Publishing Company, Reading, Massachusetts, 1986.

[14] Alexander D. Poularikas and Samuel Seely. *Signals and Systems.* PWS Publishers, Boston, 1985.

[15] D. Regan et al. Spatial frequency discrimination in normal vision and in patients with multiple sclerosis. *Brain*, 105:735–754, 1982.

[16] Ellen S. Richter and Dean Yager. Spatial-frequency difference thresholds for central and peripheral viewing. *Journal of the Optical Society of America/A*, 1(11), November 1984.

[17] Murray B. Sachs, Jacob Nachmias, and John G. Robson. Spatial frequency channels in human vision. *Journal of the Optical Society of America*, 61(9), September 1971.

[18] Warren J. Smith. *Modern Optical Engineering.* McGraw-Hill Book Company, New York, 1966.

[19] Lawrence E. Tannas, Jr. *Flat Panel Displays and CRTs.* Van Nostrand Reinhold Company, Inc., New York, 1985.

[20] Michael J Wichura. *The PICTEX Manual.* M. Pfeffer & Co., Lawrence, NY, 1987.

[21] B.J. Winer. *Statistical Principles in Experimental Design.* McGraw-Hill Book Company, New York, 1971.