

EVALUATION AND OPTIMIZATION OF IMAGE DISPLAYS

by

Ting-Lan Ji

A Dissertation Submitted to the Faculty of the
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
WITH A MAJOR IN ELECTRICAL ENGINEERING
In the Graduate College
THE UNIVERSITY OF ARIZONA

1 9 9 3

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9410669

Evaluation and optimization of image displays

Ji, Ting-Lan, Ph.D.

The University of Arizona, 1993

U·M·I

300 N. Zeeb Rd.
Ann Arbor, MI 48106

EVALUATION AND OPTIMIZATION OF IMAGE DISPLAYS

by

Ting-Lan Ji




A Dissertation Submitted to the Faculty of the
DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING
In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
WITH A MAJOR IN ELECTRICAL ENGINEERING
In the Graduate College
THE UNIVERSITY OF ARIZONA

1 9 9 3

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE


As members of the Final Examination Committee, we certify that we have
read the dissertation prepared by Ting-Lan Ji
entitled Evaluation and Optimization of Image Displays

and recommend that it be accepted as fulfilling the dissertation
requirement for the Degree of Doctor of Philosophy

	<u>9/7/93</u>
Malur K. Sundareshan	Date
	<u>9/7/93</u>
Robert A. Schowengerdt	Date
	<u>9/7/93</u>
Robin N. Strickland	Date
_____	Date
_____	Date

Final approval and acceptance of this dissertation is contingent upon
the candidate's submission of the final copy of the dissertation to the
Graduate College.

I hereby certify that I have read this dissertation prepared under my
direction and recommend that it be accepted as fulfilling the dissertation
requirement.

	<u>October 5, 1993</u>
Dissertation Director	Date
Malur K. Sundareshan	

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided the accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the dean of the Graduate College when in his or her judgement the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Englan Si

ACKNOWLEDGEMENTS

My deep appreciation and thanks are expressed to my advisor, Dr. Malur K. Sundareshan, for his insight, guidance, and patience during the research and writing of this dissertation. I am very grateful to Dr. Hans Roehrig, as my minor advisor, for his constant direction, and encouragement over the years I have spent in the Radiology Research Laboratory.

I would like to thank Dr. Hartwig Blume of the Philips Medical Systems for his encouragement and valuable suggestions. I am also indebted to Mr. Marvin (Mike) Arthur for his help during the research with electronics, mechanics, and trouble-shooting, etc.

I am very thankful to my wife, Haixiu and my parents for their continual support and love. They have been a source of enthusiasm and strength throughout my graduate years in the University of Arizona.

This work has been supported by a grant for "Development and Characterization of a Viewing Console" (Grant No. CA 49261) from The National Institute of Health.

TABLE OF CONTENTS

	<u>Page</u>
LIST OF FIGURES	7
LIST OF TABLES	10
ABSTRACT	11
1. INTRODUCTION	13
1.1 Image Display Devices	14
1.2 Evaluation of Image Display Devices	19
1.2.1 Physical evaluation	20
1.2.2 Psychophysical evaluation	21
1.3 Optimization of Display Function	23
1.4 Image Contrast Enhancement	26
1.5 Overview of the Dissertation	29
2. PHYSICAL EVALUATION OF IMAGE DISPLAYS	32
2.1 Introduction	32
2.2 Display Function	32
2.3 Physical Dynamic Range	38
2.4 Internal Light Scatter	39
2.5 Spatial Resolution	40
2.6 Noise Characteristics	46
2.7 Contrast Transfer Factor and Luminance Uniformity	54
2.8 Summary	55
3. PSYCHOPHYSICAL EVALUATION OF IMAGE DISPLAYS	58
3.1 Introduction	58
3.2 Human Visual System	59
3.3 Threshold Contrast and Just-Noticeable-Difference (JND)	61
3.4 Visual Response Function and Perceived Dynamic Range (PDR)	70
3.5 Summary	79

TABLE OF CONTENTS — *Continued*

4.	CORRELATION BETWEEN PHYSICAL AND PSYCHOPHYSICAL EVALUATIONS	80
4.1	Introduction	80
4.2	Effect of Display Device Noise on the Threshold Contrast	85
4.3	Dependence of PDR on Physical Parameters	94
4.4	Summary	97
5.	OPTIMIZATION OF THE DISPLAY FUNCTION OF IMAGE DISPLAYS	99
5.1	Introduction	99
5.2	Some Concepts from Information Theory	101
5.3	Optimizing the Display Function Based on Human Visual Response	104
5.4	Experimental Results	111
5.5	Discussion	116
5.6	Summary	122
6.	ADAPTIVE IMAGE CONTRAST ENHANCEMENT BASED ON HUMAN VISUAL PROPERTIES	124
6.1	Introduction	124
6.2	JND-Guided Adaptive Contrast Enhancement (JGACE)	136
6.2.1	Low-pass filtering	138
6.2.2	Separating detail and smooth regions	140
6.2.3	Determination of local spatial frequency and contrast	145
6.2.4	Determination of local contrast gain	147
6.2.5	The transformation $F[\cdot]$	148
6.3	Results and Comparison with Other Algorithms	149
6.4	Discussion	157
6.5	Summary	162
7.	CONCLUSIONS	164
7.1	Contributions of This Dissertation	164
7.2	Suggestions for Future Investigation	167
	REFERENCES	170

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1.1 Schematic of Cathode Ray Tube (CRT)	17
1.2 A typical H&D curve for conventional film	24
2.1 Display functions of (a) a CRT and (b) a laser film printer	34
2.2 Photograph of the SMPTE pattern	36
2.3 The ideal and the actual (for SMPTE pattern) display functions of a CRT .	37
2.4 Comparison of internal scatter for two CRT monitors	41
2.5 A two-stage model of CRT display system	44
2.6 MTFs of a CRT and a laser film printer	47
2.7 Schematic for the measurement of display device noise	50
2.8 (a) The spatial noise power spectrum of a CRT, (b) the temporal noise power spectrum of the same CRT	52
2.9 (a) Spatial and (b) temporal SNR per pixel for two CRTs	53
3.1 (a) Contrast sensitivity function and (b) its reciprocal at various luminance levels	60
3.2 A typical threshold contrast curve (Weber-Fechner relationship).	63
3.3 Measured threshold contrast for two CRTs as a function of background luminance for detection of a square object of 0.3 degree of arc, compared with the Blackwell's data	64
3.4 Illustration of the visual response function $S = h(L)$	72
3.5 Normalized visual response functions $h_N(L)$ for different values of p with other parameters fixed	73
3.6 Normalized visual response functions $h_N(L)$ for different values of C_0 with other parameters fixed	74

LIST OF FIGURES — *Continued*

3.7	Normalized visual response functions $h_N(L)$ for different values of m with other parameters fixed	75
3.8	Normalized visual response functions $h_N(L)$ for different values of L_0 with other parameters fixed	76
4.1	Probability of detection as a function of displayed signal-to-noise ratio . . .	84
4.2	Threshold contrast as a function of the side length (in pixels) of a square object with background luminance as a parameter, measured for two CRTs	89
4.3	Threshold contrast as a function of SNR_p for two different object sizes . . .	90
4.4	The measured SNR_p of a CRT as a function of the background luminance	91
4.5	The measured threshold contrast of the same CRT in Fig. 4.4 plotted as a function the background luminance for a square object of 15×15 pixels	92
4.6	The perceived dynamic range (PDR) as a function of the physical dynamic range	96
5.1	The same digital image displayed with three different display functions	100
5.2	Display/Human observer system	105
5.3	The original and the optimum display functions of a CRT	110
5.4	Typical receiver operating characteristic (ROC) curves	113
5.5	The original and the optimum display functions of the CRT used in the ROC study described in §5.4	115
5.6	ROC curves for diagnosing thorax radiographs when presented on a CRT operated with either the original or the optimum display function given in Fig. 5.5	117

LIST OF FIGURES — *Continued*

6.1	Illustration of linear contrast stretch	126
6.2	(a) The histogram of a chest image and its clipped version, (b) The input-to-output transformation (look-up table) derived from the original histogram and its clipped version given in (a)	134
6.3	A schematic diagram of JGACE algorithm	139
6.4	The dependence of low-pass filtered background image on the filter's window size	141
6.5(a)	The distribution of regional gradient means of an image	143
6.5(b)	Another example of the distribution of regional gradient means	144
6.6	Comparison of five different contrast enhancement algorithms for a test image	151
6.7	A cross profile along the middle of the test image: the original image (a), the image processed by a global histogram equalization (b), by the modified LRM (c), by the multichannel filtering (d), by CLAHE (e), and by JGACE (f)	152
6.8	Comparison of five different contrast enhancement algorithms for a chest image	154
6.9	A cross profile along the middle of the chest image (through the rectangle): the original image (a), the image processed by a global histogram equalization (b), by the modified LRM (c), by the multi- channel filtering (d), by CLAHE (e), and by JGACE (f)	155
6.10	The increase of spatial activity varies as a function of the original spatial activity of the test image in Fig 6.6(a) for five different algorithms: (a) a global histogram equalization, (b) the multichannel filtering, (c) the modified LRM, (d) CLAHE, (e) JGACE	159

LIST OF TABLES

<u>Table</u>	<u>Page</u>
2.1 Comparison of CRT monitors	57
5.1 Mean adequacy ratings of spatial and contrast resolutions	116
6.1 Comparison of computing time (minutes) for JGACE and CLAHE	156

ABSTRACT

This dissertation presents procedures for systematic and quantitative evaluations of both physical and psychophysical performance of image display devices. A mathematical expression for the visual luminance response function is derived, which permits developing an optimum display function for display devices. Direct quantitative relations between the physical and the psychophysical parameters are established. It is concluded that in the present state of modern CRTs, the spatial noise due to phosphor granularity offers the major limit to the contrast resolution, and that trying to decrease the spatial noise of a CRT is a more effective approach to increase the perceived dynamic range of the CRT among other considerations.

A systematic procedure is developed to optimize the display function such that the contrast information transfer through the display device/human vision system is maximized. The presently derived result indicates that the optimum display function is the inverse of the scaled visual response function determined from the Just-Noticeable-Difference (JND) curve, and is independent of the object size and the noise level (RMS) of the display device. The optimum display function perceptually linearizes the display device in that equal changes in grey level produce changes in luminance that are perceptually equal throughout the entire dynamic range of the display device.

This dissertation also presents a novel adaptive contrast enhancement algorithm, called JND-Guided Adaptive Contrast Enhancement (JGACE), to compensate for the limited contrast capability of display devices and to improve the quality of image display.

Existing methods for image contrast enhancement focus entirely on the properties of the image to be processed without consideration of the human visual characteristics. The presented algorithm quantitatively achieves an adequate amount of contrast enhancement in terms of the human visual JNDs, and effectively eliminates two common drawbacks of many existing contrast enhancement techniques: ringing artifacts around sharp edges and enhancement of background noise. JGACE can be applied to a variety of images and provides a superior performance compared to previously available techniques. In particular, it offers considerable benefits in digital radiography applications where the objective is to increase the diagnostic utility of images.

CHAPTER 1

INTRODUCTION

Advances in modern digital imaging techniques and computer systems have brought about revolutionary changes in diagnostic medicine. Digital radiography, computed tomography (CT), magnetic resonance imaging (MRI) and diagnostic ultrasound etc. have all developed rapidly in the past few decades [1], [2]. Today, medical imaging has become the major method for almost all kinds of diagnoses. Compared with other applications, it is also medical imaging that puts the most challenging demands on imaging tasks, such as image acquisition, image storage, image processing and image display. In chest radiography, for instance, the resolvable pixel size is usually required to be as small as 0.2 mm with a typical image dimension of 14" \times 17", which requires about 1778×2159 resolvable pixels per image, while digital mammography may demand even higher spatial resolution up to 3556×4318 pixels per image for the same image size [3]. The requirement for grey scale resolution is also high. It has been estimated that at least 10 bits per pixel in grey scale are needed for chest radiographs and as high as 12 bits per pixel may be needed for digital mammography to record small differences in image intensity [3].

Compared with other subsystems, the image display system has been considered a bottleneck of medical imaging systems, since the capability of imaging devices in collecting information is now much better than the capability of image display devices in conveying the collected information to the observer. Clearly, any improvement in the

performance of the image display system is one of the key factors for the improvement in the overall performance of the whole medical imaging system. To this end, the investigations reported in this dissertation will be directed to the subject of systematic and quantitative evaluations of image display devices and optimization of image displays, with particular emphasis on image display for medical images.

1.1 Image Display Devices

Image display devices can generally be divided into two categories: "hardcopy" and "softcopy". The former is a film-based display, for example, a laser film printer (writer), which generates a permanent copy of image (film). The latter is an electronic display, such as the cathode-ray tube (CRT) display, the plasma display, the light-emitting diode display (LED) and the liquid-crystal display (LCD). A displayed image can be electronically removed or changed with a softcopy display device. Among all the various electronic displays mentioned above, the CRT is the dominant display device so far for visual information, especially pictorial information, because the CRT is at present superior to other electronic displays in some important aspects: fast response (bandwidth), high resolution, high brightness, high versatility and reliability, low cost and long life [4]. Hence, the CRT has almost become the exclusive softcopy display device used in medical applications. In medical imaging, acquired digital images can be either recorded on the film by a laser film printer and then observed by hanging the film in front of a light-box (known as a viewing box), or directly displayed on a high-

resolution CRT monitor.

The laser film printer utilizes special films that are sensitive to the wavelength of a laser. The digital value (or, grey level) of each pixel in a digital image is converted into the intensity of the laser beam that exposes a single spot on the film. Since the laser beam can be made very narrow and well aligned, the pixel size on the film can be very small. The laser written film is then developed normally and is displayed in front of the viewing box. One such laser printer, which is equipped in the radiology research laboratory at the University of Arizona, is the Ektascan laser printer manufactured by the Eastman Kodak company of Rochester, New York. This laser film printer prints a digital image onto a 14" \times 17" film with a fixed pixel size of 0.08 mm and allows a maximum digital image size of 4084 \times 4987 pixels and 4096 grey levels (12 bits). For smaller digital image sizes, the Ektascan laser printer can replicate or interpolate digital pixels (upon request) to achieve a desired physical dimension on film. The viewing box consists of fluorescent tubes mounted behind a white plastic faceplate. The faceplate is coated to diffuse the light so that the entire panel is uniformly illuminated. The luminance level of the viewing box is typically about 500 fL. Hence, for a film of density 1.0, the luminance transmitted is about 50 fL because of the logarithmic relation between the film density and transmission. It is typical for a x-ray film to have a density range from about 0.2 to 3.2 [2]. Equivalently, the achievable luminance dynamic range (defined as the ratio of the maximum to minimum luminance levels) is 1000 : 1.

A CRT monitor consists of a cathode-ray tube, horizontal and vertical deflection systems and a video amplifier system. The CRT monitors used for displaying medical

images are usually monochrome. A schematic of a typical CRT is shown in Fig. 1.1. The operation of the CRT can be briefly described as follows: The cathode, as the electron source, is heated and emits electrons that are focused by the electron gun to form the electron beam. The electron beam is scanned over the CRT's screen in a raster pattern of parallel lines by means of the deflection yoke when the proper potentials are applied to the CRT. The information to be displayed -- that is, the video signals -- are applied to the electron-beam current-controlling electrode of the electron gun in order to modulate the intensity of the electron beam in synchronization with the deflection signals. The phosphor screen of the CRT is made of glass and has a phosphor material deposited on it that emits visible photons when struck by the electron beam. By this means, the CRT screen provides a two-dimensional visual display which corresponds to the serial electrical information contained in the electron beam. To maintain the illusion of a continuous light output to the human eye, the electron beam has to be scanned repeatedly with a refresh rate fast enough (usually above 70 frames per second) to avoid the sensation of flicker in the human eye.

As a part of the CRT display system, an image memory board, called display buffer, is needed, which is normally dual-ported so that it can be written to by the computer and read by the CRT monitor. The digital image to be displayed is stored in the monitor's display buffer. The dimension of the pixel matrix and the number of grey levels that can be displayed on the CRT's screen are determined by the size of the display buffer. The grey level of each pixel is converted by a digital-to-analog (D/A) converter into the analog video signal to drive the electron beam. Presently, all

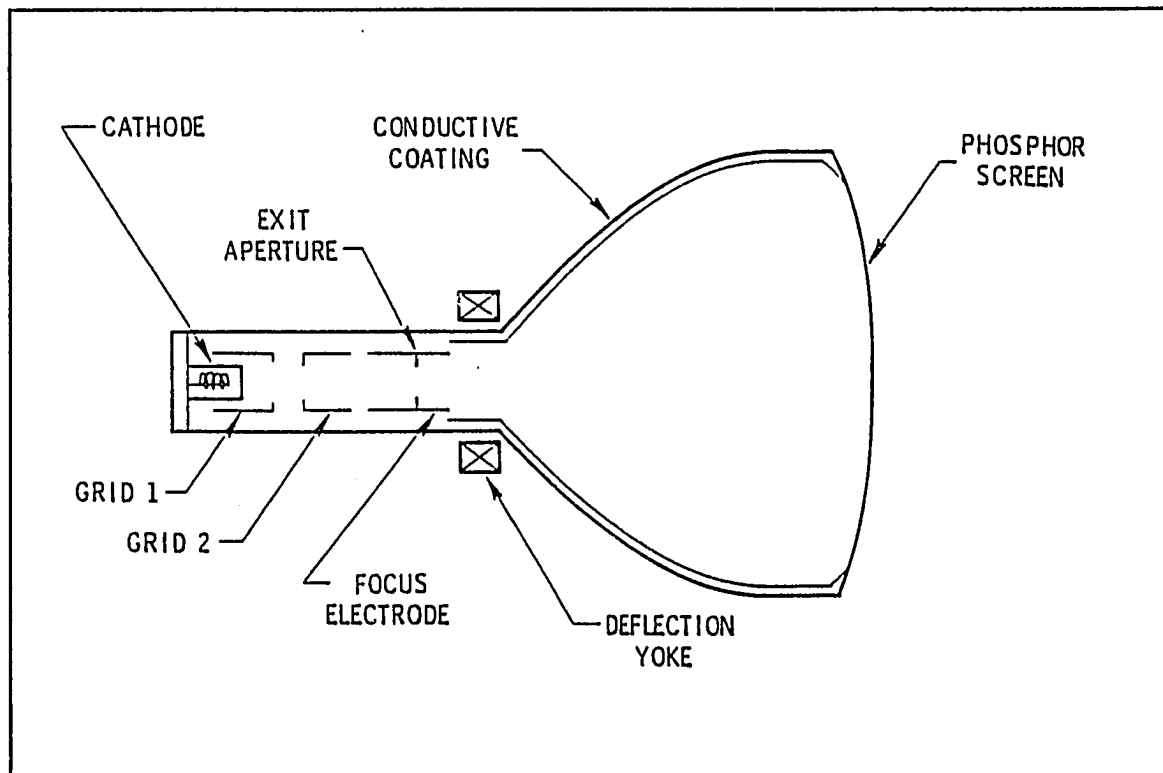


Fig. 1.1 *Schematic of Cathode Ray Tube (CRT).*

commercially available display buffers use D/A converters with quantization of 8 bits ($2^8 = 256$ grey levels).

The relation between the electron beam current i_b and the voltage of the video signal V can often be approximated by a power law [5]:

$$i_b \approx c_1 \cdot V^\gamma \quad (1.1)$$

where c_1 and γ are constants. Typical values of γ are about 2 to 3. Since the luminance of the phosphor is linearly proportional to the electron beam current, the relation between the output luminance and the input grey level also follows the same power function. The number of pixels addressable in a CRT monitor is primarily limited by the electronic bandwidth Δf of the video amplifier, which can be described by [6]

$$\Delta f = 1.8 \cdot N_p \cdot f_f \quad (1.2)$$

where N_p is the total number of pixels, and f_f is the frame rate (the number of image frames per second). The state-of-the-art CRT display systems have already achieved a matrix size of 2560×2048 pixels with a frame rate of 72 Hz, which requires a very high bandwidth of about 680 MHz according to (1.2).

The CRT monitor usually has two controls: contrast and brightness. The contrast control actually varies the gain of the video amplifier, while the brightness control adjusts an offset in electron beam intensity. These controls can be used to adjust the grey scale rendition of the CRT.

The hardcopy displays have advantages over the softcopy displays in higher spatial resolution, finer grey scale quantization, larger dynamic range and higher

brightness (with a luminance level of about 500 fL for a typical viewing box). The biggest disadvantage of the hardcopy display devices is the lack of flexibility as compared to the softcopy displays. With the softcopy display, the digital images can be stored on computer disks, transmitted across large distances, and can be easily manipulated (processed) and then redisplayed. Hence, the softcopy display is particularly useful in modern teleradiology and Picture Archiving and Communication Systems (PACS).

1.2 Evaluation of Image Display Devices

Everybody knows a "good" display device is better than a "bad" one. But what does "good" mean? There must be some parameters that describe the quality of performance of a display device so that we can evaluate them. It would be ideal if one single number, a figure of merit, could be found that completely characterized the performance of the display device. Such a number could then be used to compare the performance of different display devices. Unfortunately, this is not the case. The display device is a complex system and there are so many factors influencing the performance that no one single number can completely characterize any one display device. Researchers have tried to evaluate image display devices with many parameters. Although cost, volume, reliability, etc. are also important factors for the display device, for performance evaluation attention will be focused on those parameters that are more related to the displayed image quality, such as display function, dynamic range, spatial and contrast resolutions, and noise characteristics, etc.

1.2.1 Physical evaluation

Brightness, contrast and spatial resolution are probably the most often mentioned parameters for a display device. However, these terms are often imprecise and may result in some confusion in interpretation. For example, when people say "brightness", most likely they actually mean "luminance". Brightness is a subjective impression of the human eye, which is not a measurable quantity, while luminance is a measurable quantity which is defined as the luminous intensity per area projected in a given direction. Its SI unit is candela per square meter (cd/m^2). Another commonly used unit is footlambert (fL). Luminance most closely corresponds to brightness and is measured with a photometer. Contrast of a display device is often used to mean the displayed maximum-to-minimum luminance ratio, which is actually the luminance dynamic range. The bit depth (the number of bits) of the grey scale is sometimes quoted as the contrast resolution. But this number only tells us how fine the quantization of the driving signal is. The real contrast resolution must take into account the effect of noise of the display device, because it is the noise that really determines the smallest contrast the observer can see. Spatial resolution in some context is used to mean the number of pixels per unit display dimension [4]. Sometimes, the limiting resolution is used which is simply the spatial frequency (in the unit of line pairs per unit length) at which the observer can just barely discriminate the black and white bars of the image [4], [7]. In the TV industry, the TV lines per picture width (or height) is often used to describe the spatial resolution [7]. In recent years, the Modulation Transfer Function (MTF) has gained a strong popularity in describing spatial resolution, as it completely characterizes the spatial

frequency response of a display system and has a good correlation with the quality of displayed images [8]-[12].

A literature survey has revealed that very few investigations have been conducted on complete, systematic and quantitative evaluations of the physical parameters of modern, high-performance display systems. On the other hand, the applications of high-performance display systems, especially in medical imaging area, make it necessary to attempt maximizing the potential capabilities of a given display system. It is impossible to accomplish this objective without a systematic and quantitative characterization and evaluation of the physical parameters of the display system.

1.2.2 Psychophysical evaluation

The purpose of displaying an image is for observers to extract the visual information from the image. It is therefore reasonable to regard the display device and the human visual system as one combined system. We not only need to know the physical parameters of the display system as they are more objective, but also have to figure out how the observer's perception of the displayed image is affected by these physical parameters and their changes.

Threshold contrast is probably the most important psychophysical quantity that is closely related to both the physical characteristics of the display device and the human visual response. Any complex image can be decomposed into many simple objects of various contrasts, such as squares, rectangles, circular disks, and bar patterns, along with a spatially slowly-varying background. Many efforts have been directed to investigate

how the human visual system responds to these simple stimuli, particularly in the threshold detection situations. Almost a half century ago, Blackwell [13] performed a classic experiment for determining the threshold contrast in detection of a small circular disk surrounded by a uniform background as a function of the size of the disk and the background luminance. His experimental results verify that for a fixed object size, the threshold contrast decreases with background luminance, following $1/L^{0.5}$ relation, at low background luminance levels, and then remains constant when background luminance becomes sufficiently high. This threshold contrast vs. background luminance behavior is referred to as the Weber-Fechner relationship [14]. The threshold detection of periodic signals, such as sine-wave or square-wave patterns, is characterized by contrast sensitivity which is the reciprocal of the threshold contrast as a function of the spatial frequency of the grating pattern. This threshold contrast is also dependent on the background luminance. The concept of the Just-Noticeable-Difference (JND) is directly related to the threshold contrast (C_T) by $C_T = JND/L_b$, where L_b is the background luminance. Thus, JND is the smallest luminance difference the human observer can detect when an object of a certain size is displayed at a certain background luminance level. JND is a more important parameter for determination of the true dynamic range of a display device. Pizer and Chan [15] have proposed that the perceived dynamic range is given by the number of JNDs within the luminance range of a display device.

The detection of a small contrast was widely considered to be dependent on the displayed signal-to-noise ratio. de Vries [16] and Rose [17] first suggested that the photon noise was the limiting factor which determines the detectability of signals in low

light levels. The model proposed by Rose predicts that the threshold contrast is proportional to the square-root of the photon fluence and inversely proportional to the square-root of the object size. This prediction agrees with Blackwell's results over a limited range. At high luminance levels, however, the Rose model is no longer valid because the limiting noise source is changed from photon noise to the internal neural noise of the human visual system. It is of great interest to investigate if the Rose model can be modified or extended to predict the threshold contrast in the case of displaying objects on a CRT where the limiting noise is not the photon noise but the phosphor granularity.

1.3 Optimization of Display Function

It is well known to those who are familiar with photography that the quality of picture is greatly dependent on the film's gradation curve, often called the Hurter-Driffield (H&D) curve [18]. Typically, the H&D curve has a sigmoidal shape, as shown in Fig. 1.2. In the region where the curve is nearly a straight line, the relation between the optical density D and the relative exposure E is approximately given by

$$D = \gamma \log_{10} \frac{E}{E_0} \quad (1.3)$$

where γ is the slope of the curve and controls the contrast of the picture, and E_0 is a reference exposure. The grey scale of the picture (here we only discuss monochrome film) is completely determined by the H&D curve. This input-to-output relationship of

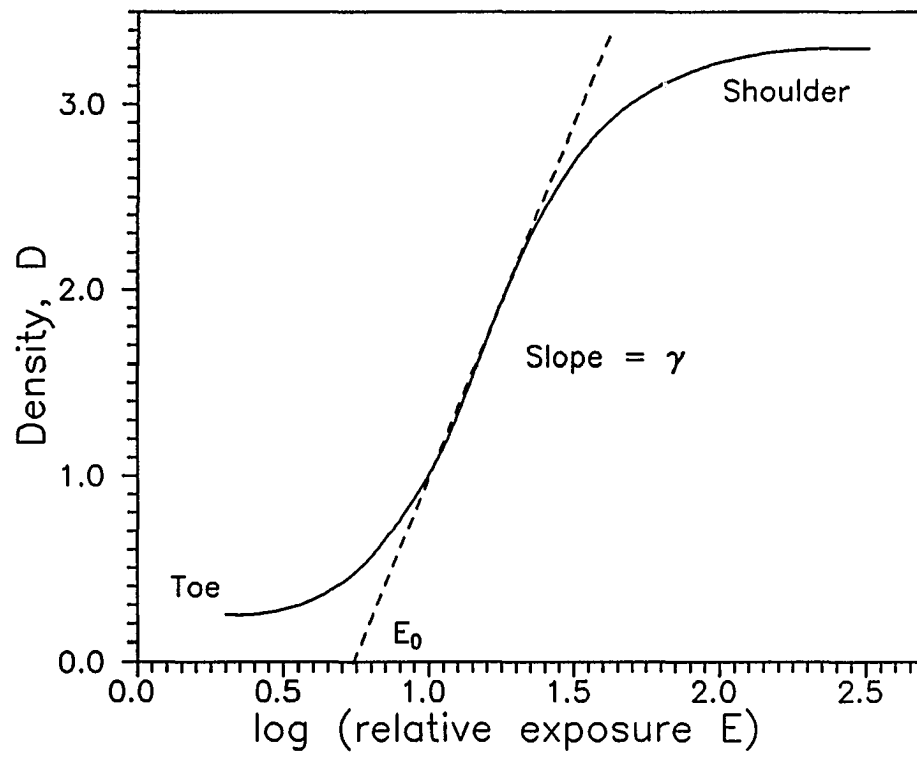


Fig. 1.2 *A typical H&D curve for conventional film*

the film is referred to as the characteristic curve or display function. In photography, the adjustment of the display function is most likely to make the picture more pleasant to view for aesthetic or artistic purposes. Whereas in medical imaging, it is desired to have the best display function for a display device in order to provide the observer with the maximum amount of information out of the displayed image. For softcopy display systems, it is very easy to manipulate the display function interactively with commonly available contrast and brightness controls, or windowing and leveling software facilities. It seems that such approaches of interactive examination of the image can allow the observer to extract all the relevant information and therefore the optimization of display function is unnecessary. While it is probably true that the interactive manipulation of the image display is quite effective, it does depend strongly on how efficiently the observer uses the display facilities and it may be a time-consuming process. In a clinical environment, it is highly likely that the above manual procedure will be inefficient and non-repeatable. There are other problems with this kind of arbitrary manipulation of the display function. When an image is displayed on a CRT, the grey scale, and hence the appearance of the image, may be changed from one observer to another and from time to time due to the inappropriate interactive manipulation. Since the display function varies from monitor to monitor, the grey scale of the same image will look different when displayed on different monitors. Hence, when transferring an image with grey scale satisfactorily presented on the CRT display of a work-station to another softcopy display, the rendition of grey scale on the second monitor may not be satisfactory. In addition, radiologists have long been used to viewing conventional films to make their

primary diagnosis. In digital radiology, when an image is displayed on a CRT monitor, even with a film-like gradation curve, the grey scale rendition is not automatically the same as seen on film. These problems have contributed to reservations held by many radiologists about working with softcopy displays. These problems exist largely because display systems (CRTs and laser film printers) do not have a mathematically defined standard display function. Such a standard must be established with consideration of the human visual response to the displayed luminance. The concept of perceptually linearizing display devices has been proposed by many researchers in the past [19]-[22]. Perceptual linearization produces a display function such that any equal changes in driving signal (grey levels) result in changes in luminance that are perceptually equal. Perceptual linearization has also been utilized either directly or indirectly for defining quantization coefficients in lossy data compression [23] and for determining the minimum digitization resolution without generating visible digitization contour artifacts [14], [24]. It is of great interest to investigate whether the perceptually linearized display function is the optimum display function for contrast information transfer for display devices and to determine the nature of a mathematical expression for the optimum display function.

1.4 Image Contrast Enhancement

Even with the optimum display function, the contrast capability of softcopy displays may still be insufficient to present some low-contrast subtle details in the image. However, one of the advantages with the softcopy displays is that image processing can

be applied to the original image before displaying it. Among the techniques used in image processing, contrast enhancement improves the appearance of the image, allowing observers to see more information from that image. This is particularly important in digital radiography applications where abnormal structures on radiographs are often presented by small luminance differences caused by small differences in the x-ray path attenuation and in object thickness. In addition, current softcopy display systems are incapable of displaying as many different discernible levels of luminance as can be recorded in a digital image. The contrast dynamic range of CRTs is also not as adequate as in laser printed films. These factors are very likely to cause some subtle, but probably crucial, contrast information loss when a digital image is displayed on a CRT monitor. The diagnostic accuracy can be reduced due to the poor contrast in the displayed image. Contrast enhancement is clearly one way to compensate for the decrease of contrast in the displayed image.

Numerous contrast enhancement methods exist, ranging from the simplest linear contrast stretch (sometimes called windowing and leveling) to some very sophisticated adaptive enhancement techniques [25], [26]. They can be generally classified into two categories: global techniques and adaptive techniques. In general, for global methods one transformation is applied to all the pixels of the input image, while adaptive contrast enhancement involves a mapping of the form

$$I_O(x,y) = f[I(x,y), O_\Omega(x,y)] \quad (1.4)$$

where $I(x,y)$ is the original image, $I_O(x,y)$ is the output image and $O_\Omega(x,y)$ represents some local characteristics within the neighborhood Ω around the location (x,y) . Hence,

the mapping f changes adaptively with the local characteristics of the image. Global methods may work well for some images. However, in the real world there are more complex situations where an image may have sufficient global contrast with considerable low-contrast local details, or the contrast is poor in some parts of the image but adequate in other parts of the image. For instance, chest images typically show a very large contrast between different anatomical regions such as the heart and lungs, whereas the contrast of detail structures is de-emphasized. In these cases, adaptive techniques usually provide a better performance than the global ones. Evidently, adaptive contrast enhancement algorithms are usually more complex and need more computations.

Examination of several existing techniques reveals some common problems in contrast enhancement. First, it is not clear how much enhancement is adequate at each location of the image. It is likely for a processed image to have some regions where the local contrast is already too much, but have other regions where the local contrast is still not enough for observers to see the details. Secondly, the over-enhancement of noise is distracting, especially in relatively homogeneous regions of the image, which is partly related to the first problem because it is difficult to know a priori where the contrast needs to be enhanced and the amount of enhancement. Finally, ringing artifacts commonly appear around the sharp transitions of the image. Since all these problems are related to the perceptual effects of the displayed image, it may be beneficial to design an adaptive contrast enhancement algorithm based on some pertinent human visual properties, such as JNDs. Besides, such an adaptive contrast enhancement may be combined with utilizing the optimum display functions of the display device to allow the

observer to detect as much contrast information as possible.

1.5 Overview of the dissertation

This dissertation consists of seven chapters. The first chapter is an introduction to the whole dissertation, which contains an outline of the problems of interest, a brief review of the investigations conducted in the discussed field, and the organization and contributions of the dissertation.

Chapter 2 presents a systematic and quantitative evaluation of the physical performance of image display devices. The emphasis is placed on those physical parameters that greatly affect the quality of image displayed by the display device. Each parameter is carefully defined, and its importance is discussed in great detail. The techniques for measurement of these parameters are described, and some measured results are presented to give a quantitative comparison of different display devices.

Since the human observer is the final receiver of the displayed information contained in an image, a complete evaluation of image display system must consider the response of the human vision system. Chapter 3, beginning with a brief review of the human visual system, describes psychophysical evaluation of image displays. The important perceptual characteristics associated with display devices, which include the threshold contrast, just-noticeable-difference (JND) and perceived dynamic range (PDR), are introduced. A mathematical curve fitting to the threshold contrast is proposed, and the important visual response function, which represents the relation between the

perceived brightness and the displayed luminance, is then derived from the proposed fitting function of the threshold contrast.

In Chapter 4, an analytical relationship between the physical noise characteristics and the psychophysical performance parameter -- threshold contrast -- is established by extension of the Rose model. Experiments that have been conducted to confirm this relationship are described. Other correlations between the physical and the psychophysical performances of the display device are also discussed.

Chapter 5 describes how to optimize the display function of image display devices in order to maximize the contrast information transfer through the display device/human vision system. The optimum display function is derived from the visual response function, and will perceptually linearize the display device. Psychophysical experiments are carried out to show that the optimum display function is generally superior to the original display function for displaying images. The effects of several physical parameters on the optimum display function are investigated as well.

Because of a large compression in dynamic range from the stored digital image to the displayed image by a CRT monitor, some subtle, but probably crucial, contrast information will be lost if nothing is done before displaying the image. It is also likely that the contrast of some abnormalities in the original image is already below the visual threshold due to the very small differences between the objects and their surroundings and/or poor imaging conditions. In both cases, contrast enhancement will be needed to compensate for the decrease of contrast in the displayed image. In Chapter 6, a novel adaptive contrast enhancement algorithm, called JND-Guided Adaptive Contrast

Enhancement (JGACE), is developed based on the basic human visual properties. The new algorithm is designed to provide a perceptually adequate contrast to all local details of the image when the original image is processed by JGACE and then displayed with the optimum display function.

Finally, Chapter 7 summarizes the results and conclusions of the investigations reported in this dissertation and also provides some suggestions for possible further investigations.

The major contributions of this dissertation are the following:

- (1) Procedures for systematic and quantitative evaluations of both physical and psychophysical performance of image display devices are presented.
- (2) A systematic procedure is developed for optimization of the display function for image display devices in order to maximize the contrast information transfer through display device/human vision systems.
- (3) A novel adaptive image contrast enhancement technique which is based on the basic human visual properties is developed to compensate for the limited contrast displaying capability of display devices and to improve the quality of image display.

CHAPTER 2

PHYSICAL EVALUATION OF IMAGE DISPLAYS

2.1 Introduction

A systematic and quantitative evaluation of the physical parameters of an image display device is clearly important for users to judge the image quality of the display device, compare different display devices objectively, and be able to extract the optimum performance from the specific display device. Some of the important physical parameters which determine the overall image quality of a display device are the display function, absolute luminance, physical dynamic range, internal light scatter, spatial resolution, noise characteristics, contrast transfer factor, and luminance uniformity. In the following sections, these parameters will be defined and analyzed in detail and how to evaluate them will also be described. The detailed procedures and methods for the measurement of these characteristics will not be presented here because they are too long to contain in this dissertation. They can, however, be found in other technical reports and publications [27]-[34].

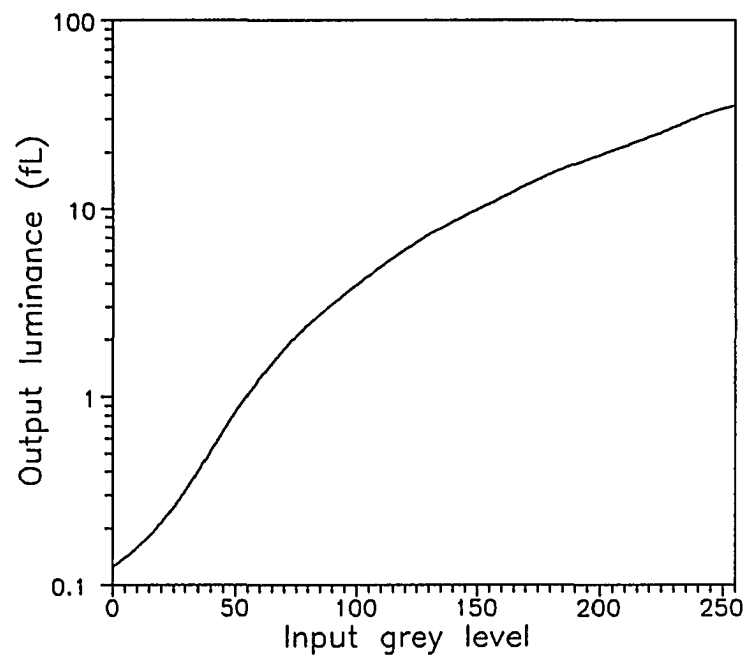
2.2 Display Function

The display function is also called the characteristic function which represents the input-to-output relationship of an image display system. For CRT monitors, it is defined

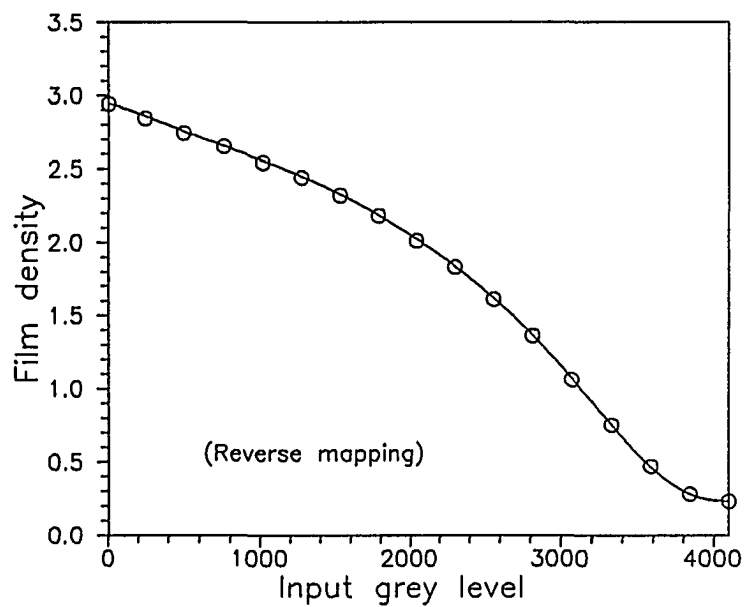
as the relation of the input grey level to the output luminance. If the image display system is a laser film printer, the display function is defined as the relation between the input grey level and the optical density on the printed film. We denote the display function by $L = f(GL)$ for CRTs or $D = f(GL)$ for laser film printers. The typical display function $f(\cdot)$ is non-linear and monotonic. As examples, Fig. 2.1 illustrates the display functions of a CRT and a laser film printer.

The display function is one of the fundamental characteristics of an image display system. As we will see later, many other physical parameters, such as absolute luminance, physical dynamic range and contrast transfer factor, can be determined from the display function. The quality of the image displayed by a display system is critically influenced by its display function. The best display function is the one that matches human perceptual requirements, which will be the central topic discussed in Chapter 5.

The display function of an image display system is not completely fixed after the image display system is made. Quite often, there are contrast/brightness knobs (sometimes only one combined knob) available with a CRT monitor. Roughly speaking, the contrast knob controls the steepness of the display function (i.e., the gain of the CRT's video amplifier), and the brightness knob shifts the curve up or down (i.e., the offset of the brightness). Before any physical evaluation, the contrast/brightness knobs must be set properly. An appropriate contrast/brightness setting should produce a smooth display function which is neither saturated at the high signal end (white) nor cut off at the low signal end (black), and provide a large dynamic range as well as a high output luminance. A standard test pattern, SMPTE (Society of Motion Picture and Television



(a)



(b)

Fig. 2.1 Display functions of (a) a CRT, (b) a laser film printer.

Engineers) pattern, as shown in Fig. 2.2, is usually used to make the contrast/brightness setting such that (1) the 5% and the 95% contrast tablets of the SMPTE pattern can be seen equally well, (2) the maximum luminance measured from the 100% signal tablet is made as high as possible without a noticeable degradation of sharpness for the detail contents of the SMPTE pattern, and (3) the boundary between the black border of the SMPTE pattern and the inactive border area of the CRT is barely discernible. This contrast/brightness setting is then fixed and considered as a reference setting for the subsequent physical evaluation. It should be noted that the display function achieved with the reference setting is not necessarily the optimum display function as analyzed later in Chapter 5.

For a CRT, the measurement of the display function is made by using a calibrated photometer and a video signal generator. Each specific grey level is applied by the video signal generator to the entire active area of the CRT and the corresponding output luminance at the center of the CRT is recorded by the photometer. The display function measured in this way is referred to as the ideal display function. When a real image is displayed the shape of the display function will be somewhat different from the ideal one due to the internal scatter which is caused by multiple reflection of light between the phosphor and the faceplate of the CRT. The SMPTE pattern is usually utilized to define an actual display function. As shown in Fig. 2.2, the SMPTE pattern has 13 square tablets which are addressed to 0%, 5%, 10%, 20%, ... , 90%, 95%, and 100% input signal levels. The luminance is measured at the center of these tablets. Fig. 2.3 shows an example of both the ideal and the actual display functions of a CRT. As can be seen

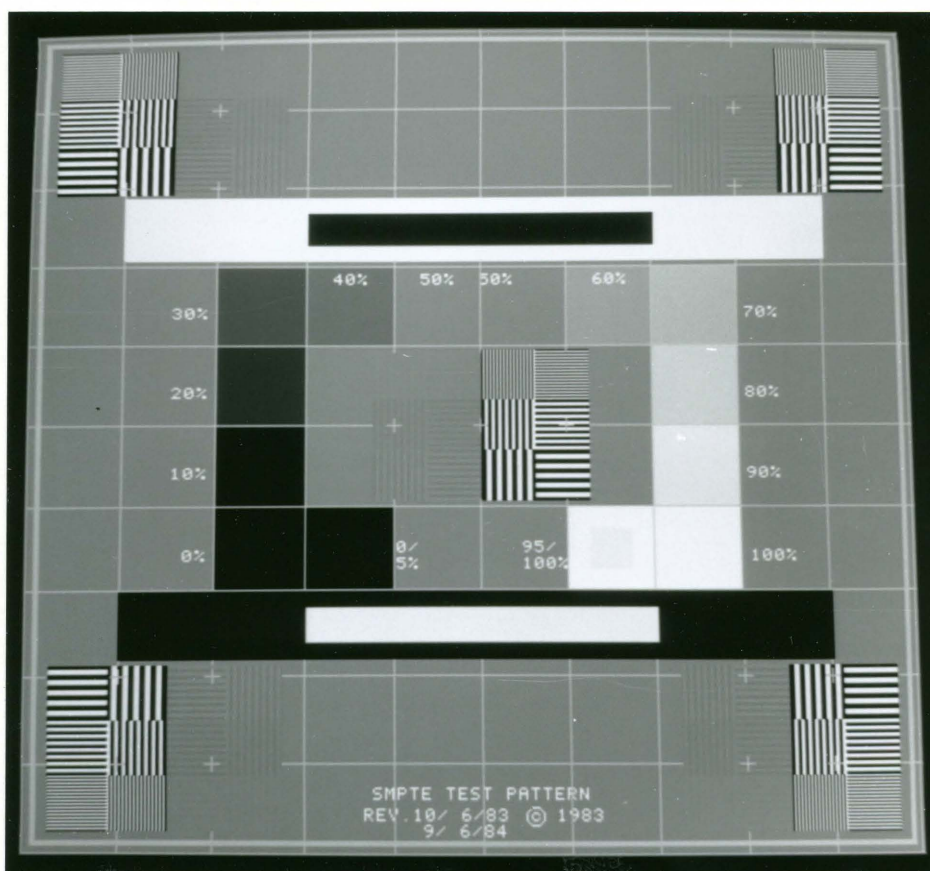


Fig. 2.2 *Photograph of SMPTE pattern.*

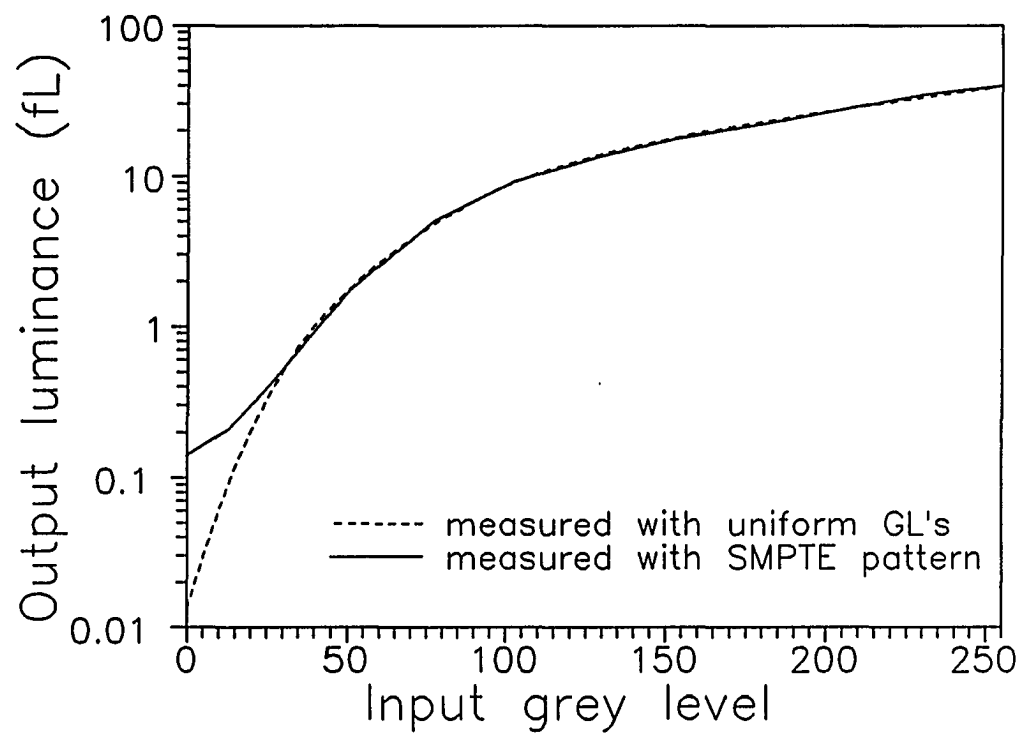


Fig. 2.3 *The ideal and the actual (for SMPTE pattern) display functions of a CRT.*

from the figure, the internal light scatter has almost no effect on the display function at mid to high grey levels but dramatically changes the shape of the display function at low grey levels.

The measurement of the display function for a laser film printer is simpler because of almost no internal scatter problem. The ideal and actual display functions are actually identical. One can use a densitometer to measure the density of the printed film written to specific grey levels by the laser film printer under test.

2.3 Physical Dynamic Range

The achievable luminance range directly affects the shades of grey for a displayed image, which represents the contrast resolution of the image display system. Although several definitions of the dynamic range exist, it appears reasonable to define the luminance dynamic range as the maximum-to-minimum luminance ratio which can be obtained directly from the measured display function. It is well known that the brightness perceived by human observers is approximately a logarithmic function of the displayed luminance for quite a large range of the luminance [35]. If we assume that the highest and lowest brightness correspond to the maximum and minimum luminance, L_{\max} and L_{\min} , respectively, the perceived brightness range is proportional to $(\log L_{\max} - \log L_{\min}) = \log(L_{\max}/L_{\min})$. Hence, there is a nice log-relationship between the luminance dynamic range and the perceptual range of brightness (the definition of the so called "perceived dynamic range" will be given in Chapter 3). For laser film printers, L_{\max} and

L_{\min} are not fixed and depend on the luminance level of the viewing box in use. However, since $\Delta D = \log(L_{\max}/L_{\min})$ where ΔD is the difference between the maximum and minimum film densities, the ratio L_{\max}/L_{\min} is independent of the brightness of the light box and is given by $10^{\Delta D}$. To distinguish it from the later defined perceived dynamic range, we shall call L_{\max}/L_{\min} physical dynamic range.

2.4 Internal Light Scatter

As mentioned before, the internal light scatter (also called veiling glare) is a discriminating phenomenon for such softcopy displays as CRTs. The light emitted from the CRT's phosphor is partly reflected by the glass envelope to the phosphor, and is then partly reflected again. This process repeats causing the luminance of one area, especially a dark area, to be increased by the luminance of the surrounding bright areas. The internal scatter significantly reduces the actual physical dynamic range of a CRT when an image is displayed, because of a large increase of the luminance level in "black" areas. For example, one CRT we evaluated has an ideal physical dynamic range of 3000:1. But when the SMPTE pattern is displayed on the CRT, the physical dynamic range is decreased to 276:1! This one order of magnitude reduction in physical dynamic range is typical for CRTs, which is one of the major reasons why the contrast capability of CRTs is inferior to that of the laser film printer (the dynamic range for laser printed films is typically 1000:1, see §1.1). We have already seen in Fig. 2.3 the influence of the internal scatter on the display function. The internal scatter also degrades the spatial

resolution because it reduces especially mid and high-frequency modulation transfer.

To evaluate the internal scatter we will introduce a measure called Veiling Glare Index (VGI). The VGI is defined and measured as the ratio of the luminance in a black area surrounded by a full screen (except the centered black area) white background to the maximum luminance of the CRT. Of course, VGI is actually a function of the fraction size of the black area over the full size of the active CRT screen. Fig. 2.4 shows VGI curves for two CRTs. It is clear from the VGI curves that CRT 1 has better internal scatter performance than CRT 2. To reduce the internal scatter in the CRT, one can use an anti-reflection coating on the glass faceplate. Unfortunately, it will also decrease the absolute luminance, which is the maximum output luminance corresponding to the maximum input signal. In general, a high absolute luminance for a CRT is desired because it allows users to work with it in the environment of nearly room ambient light. But other factors must be considered as they are closely related. A high absolute luminance achieved by a large beam current usually causes a large spot size and consequently degrades the spatial resolution. The actual physical dynamic range may not be large with a very high absolute luminance because of the corresponding high internal scatter. Some trade-off has to be made among these parameters.

2.5 Spatial Resolution

The spatial resolution is probably the parameter that people are most interested in. In the area of digital imaging systems, manufacturers usually provide the addressable

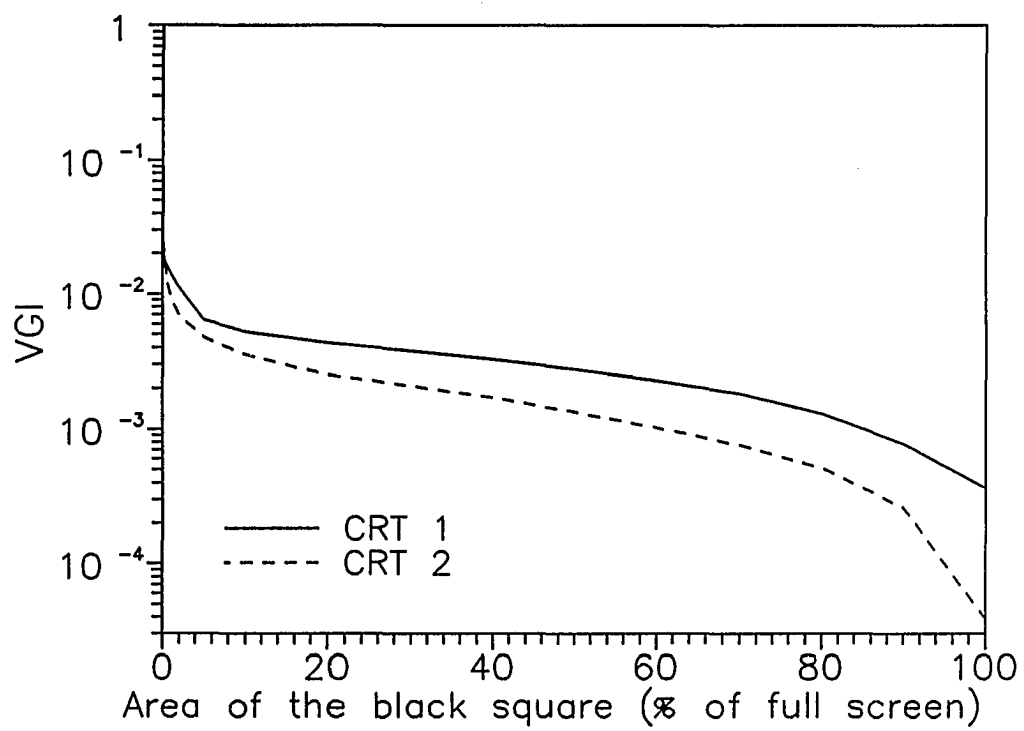


Fig. 2.4 *Comparison of internal scatter for two CRT monitors.*

pixel matrix for a CRT display system. Given the size of the CRT, one can calculate the theoretical pixel size (or nominal pixel size). For instance, assume that the display buffer of the CRT has an addressable pixel matrix of 1536×2048 and an active screen size of $26.7 \text{ cm} \times 35.6 \text{ cm}$. The nominal pixel size of the CRT would be 0.174 mm . However, the use of these values only have a commercial meaning because they do not tell the user how good or bad the CRT presents the contrast information for image details of various spatial frequencies, and we cannot tell from these numbers which one is better in resolution if we have two CRT display systems with the same pixel matrix and the same active screen size. A more useful measure of the spatial resolution is the resolvable pixel size which is usually larger than the nominal pixel size. To describe the spatial resolution more precisely, we use the Modulation Transfer Function (MTF) as a measure of the spatial resolution of an image display. MTF is the ratio of the output modulation to the input modulation of a sinusoidal wave image as a function of the spatial frequency of the sine-wave [36]. The modulation is commonly defined as

$$M = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (2.1)$$

The MTF objectively predicts the ability of an image display system to preserve the input signal modulation for various spatial frequencies. Since the output modulation can never be larger than the input modulation, the maximum value of MTF is 1. For a good spatial resolution we certainly want a large MTF. According to linear system theory, the MTF is the modulus of the Fourier transform of the point spread function which is the output response of a two dimensional system to an input impulse [36], [37]. For

image display systems, the image of a single "on" pixel can be considered as a practical point spread function. The actual size of this light spot is determined by the beam current which usually has a Gaussian-like intensity profile. Conventionally, the Full Width at Half Maximum intensity (FWHM) is used as the measure of the resolvable pixel size.

There is a problem in the direct use of the MTF concept for image display systems. The input-to-output relationship of image display systems is non-linear while the MTF is theoretically used in linear systems. For a linear system, the MTF is independent of the absolute value of the input signal modulation. But for image display systems, the MTF varies with the input signal modulation. If, for a CRT, we have a MTF value of 0.5 for an input modulation of 2%, then we can only predict that the output modulation is 1% for this input modulation. We cannot say that the output modulation will be 20% if the input modulation is 40% for the same system. In other words, the concept of MTF for the image display system is valid only for differentially small input and output levels of modulation. However, one can deal with this problem by introducing the concept of "effective input" which is determined only by the non-linearity of the image display. The display device is modeled as a two stage system whose first stage is the non-linearity $f(GL)$ (the display function) while the second stage is a linear system characterized by its MTF, as shown in Fig. 2.5. The output of the first stage is called the "effective input". The MTF of the display device is actually defined by the MTF of the second stage. Any input modulation can be converted to the modulation of the "effective input". The MTF is then defined by the ratio of the

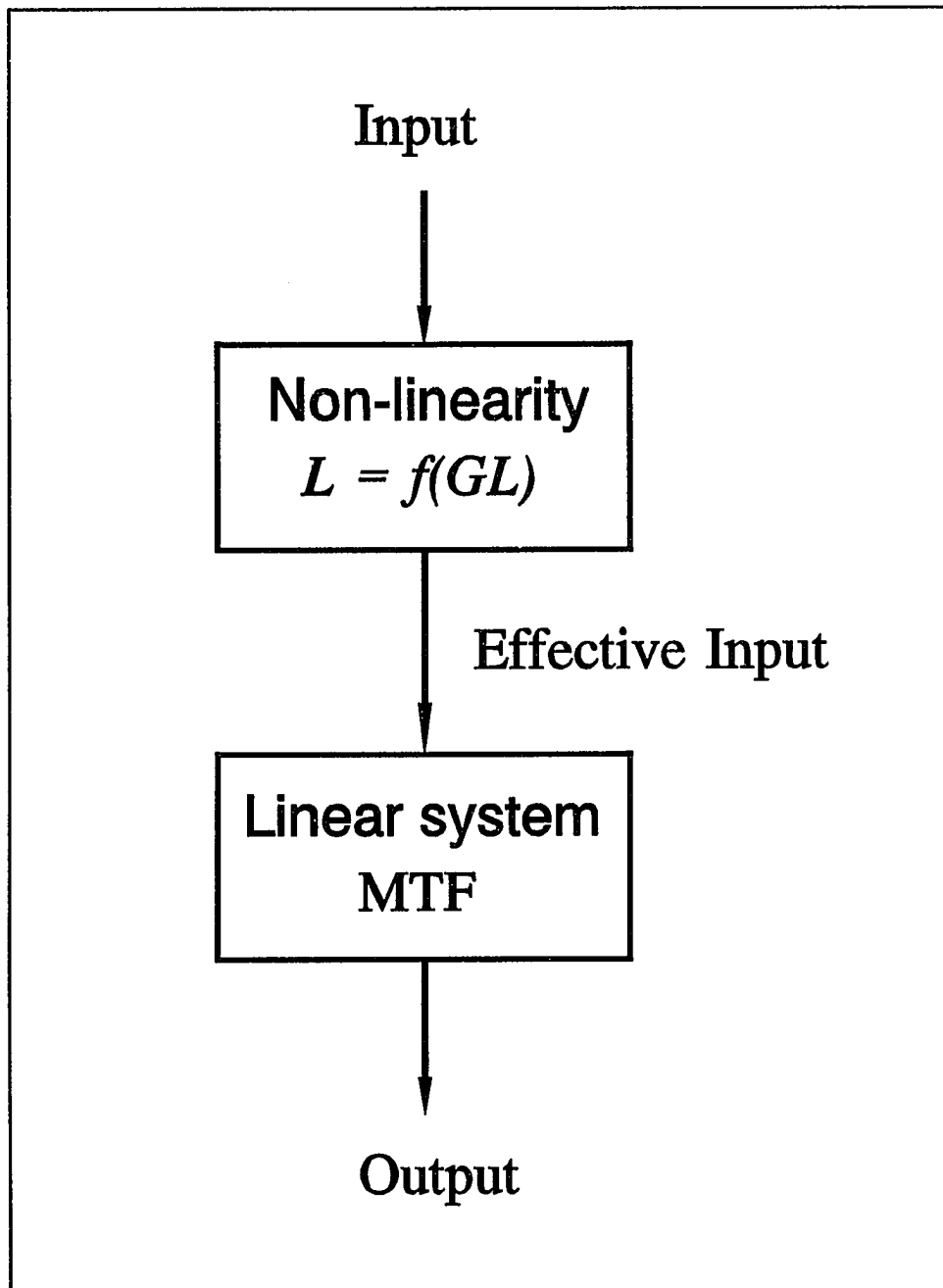


Fig. 2.5 *A two-stage model of CRT display system.*

measured output modulation to the "effective input" modulation. Thus, the MTF does not depend on the magnitude of the "effective input" modulation and uniquely represents the modulation transfer capability of the display device. The measurement of MTF for an image display system can be made several ways. One method is to measure the two dimensional luminance response of a single pixel, as the point spread function (because most image displays are anisotropic, the MTFs in horizontal and vertical directions are usually different), and then calculate its Fourier transform. In the second method, an approximate line spread function can be measured by turning on a row or a column of pixels and scanning across them with a narrow slit. Again, the modulus of the Fourier transform of the line spread function in horizontal and vertical directions gives the desired MTFs. For both methods, a magnifying lens and a CCD camera are used to take the image of either a single "on" pixel or a line of "on" pixels in a black background. The magnification of the lens should be set to ensure that there are enough samples (i.e., CCD pixels) taken within the measured profile. The third method is to measure the square-wave response of the image display and derive the MTF from it. Horizontal or vertical square-wave patterns with various frequencies and unity modulation (i.e., black and white bars) are generated by computer and displayed in the center of the CRT screen (or written on the film). The surrounding background can be set to different luminance levels to represent different effects of veiling glare on the output modulation. A white background will be the worst case of veiling glare, while a mid-luminance background will be a representative of most real images. For CRT displays, the displayed luminance of the patterns are sampled through a long, narrow slit with the

width of about a tenth of the pixel size. For a written film, the density modulation is obtained from a film digitizer. The MTF can be derived from the measured square-wave response $R(f_0)$ by the following formula [38]:

$$MTF(f_0) = \frac{\pi}{4} \left[R(f_0) + \frac{R(3f_0)}{3} - \frac{R(5f_0)}{5} + \frac{R(7f_0)}{7} + \frac{R(11f_0)}{11} - \frac{R(13f_0)}{13} - \frac{R(15f_0)}{15} + \dots \right] \quad (2.2)$$

For CRTs, the MTF obtained with the third method is more realistic because it is closer to the situation of a real image where the veiling glare exists. Typically, the veiling glare will cause the MTF to decrease greatly at mid to high spatial frequencies. Fig. 2.6 shows the MTFs of a CRT and a laser film printer.

2.6 Noise Characteristics

The noise of an image display device is a measure of the fluctuation of the output luminance when the input grey level is constant. The luminance fluctuations can be divided into two types: temporal noise and spatial noise. The temporal noise includes the fluctuation of the photon flux, called photon noise, the electron beam shot noise and video amplifier noise (for CRTs). The spatial noise stems from the phosphor granularity (or film grains). At the current state of display technology, the noise of the image display device is the limiting factor to the detection of the small contrast in the displayed image. In the next chapter, we will see how the threshold contrast of the human observer depends on the noise of the image display device.

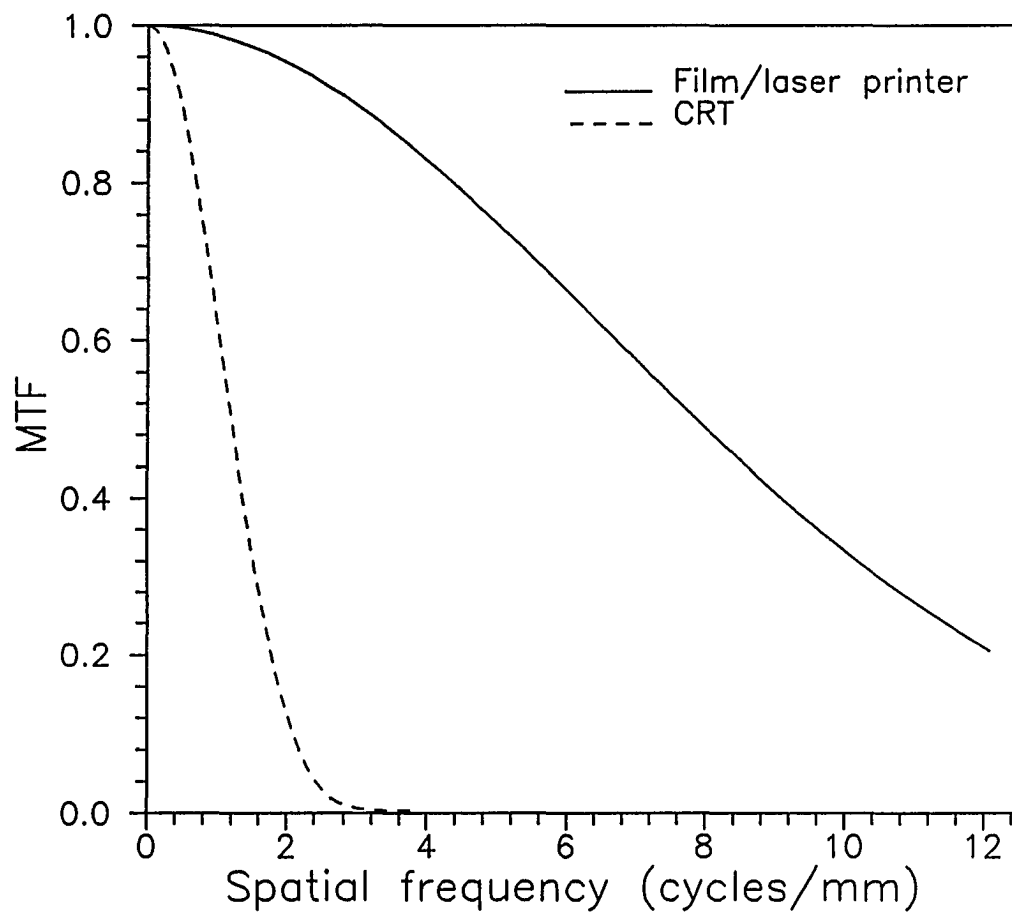


Fig. 2.6 The MTFs of a CRT and a laser film printer (Note: the pixel sizes of the CRT and the laser film printer are 0.215 mm and 0.08 mm, respectively).

The interesting characteristics of the display device noise are primarily the standard deviation of the noise (or noise RMS) and the noise power spectrum. Let $L(i)$ ($i=1,2,\dots,n$) denote the sampled luminance values. The standard deviation of the display noise can then be estimated by

$$\sigma_L = \sqrt{\frac{1}{n} \sum_{i=1}^n [L(i) - \bar{L}]^2} \quad (2.3)$$

where $\bar{L} = \frac{1}{n} \sum_{i=1}^n L(i)$ is the mean luminance. One of the common ways to characterize the noise performance of a system is using signal-to-noise ratio. As far as display devices are concerned, we define a luminance signal-to-noise ratio (SNR) as the ratio of the mean luminance to the standard deviation of the luminance fluctuation:

$$SNR = \bar{L} / \sigma_L \quad (2.4)$$

The noise power spectrum is the measure of the frequency content distribution of the noise by which we can know whether the noise is white (i.e., its spectrum is independent of the frequency) or there are some noticeable frequency components which may correspond to certain fixed noise patterns presented by the display device. The noise power spectrum is estimated from the sampled data by employing a direct Fourier transform method, called the periodogram averaging method, which is summarized in the following steps [39]:

1) Section the sampled luminance data $\{L(i)\}$ (assume that the mean has already been removed), $i=1,2,\dots,n$, into K sections each of length M , where $K = n/M$, i.e.

$$L_k(i) = L(i + M(k-1)), \quad k = 1, 2, \dots, K, \quad i = 1, 2, \dots, M .$$

2) Calculate K periodograms $P_k(\Omega)$ using Discrete Fourier Transform (DFT) as

$$P_k(\Omega) = \frac{1}{M} |DFT [L_k(i)]|^2 , \quad k = 1, 2, \dots, K . \quad (2.5)$$

3) Average the K periodograms obtained above to get the estimate of the power spectrum

$$S(\Omega) = \frac{1}{K} \sum_{k=1}^K P_k(\Omega) . \quad (2.6)$$

A schematic arrangement for the measurement of display device noise (which is developed at the Radiology Research Laboratory of the University of Arizona) is shown in Fig. 2.7. Noise is measured by sampling the luminance output through a slit. The width of the slit is set equal to the nominal pixel size of the display device and the slit is long enough to cover many displayed lines. So the number of pixels sampled by the slit is equal to the number of the displayed lines. The reason for taking the sample of many pixels is to avoid too weak input light to the photomultiplier tube (PMT). The displayed lines are set to a given grey level while the rest of the display device is set to $GL = 0$. For temporal noise measurement, the slit remains stationary. For recording the spatial noise (including the temporal noise as well), samples are taken while the slit scans along a section of the displayed lines. The sampling interval is set equal to the nominal pixel width. Two Butterworth filters with a bandwidth of 8 Hz are used in the noise measurement to provide a corresponding time period of 0.0625 seconds, which is fairly close to the integration time of the human eye. The filtered signal is sampled and

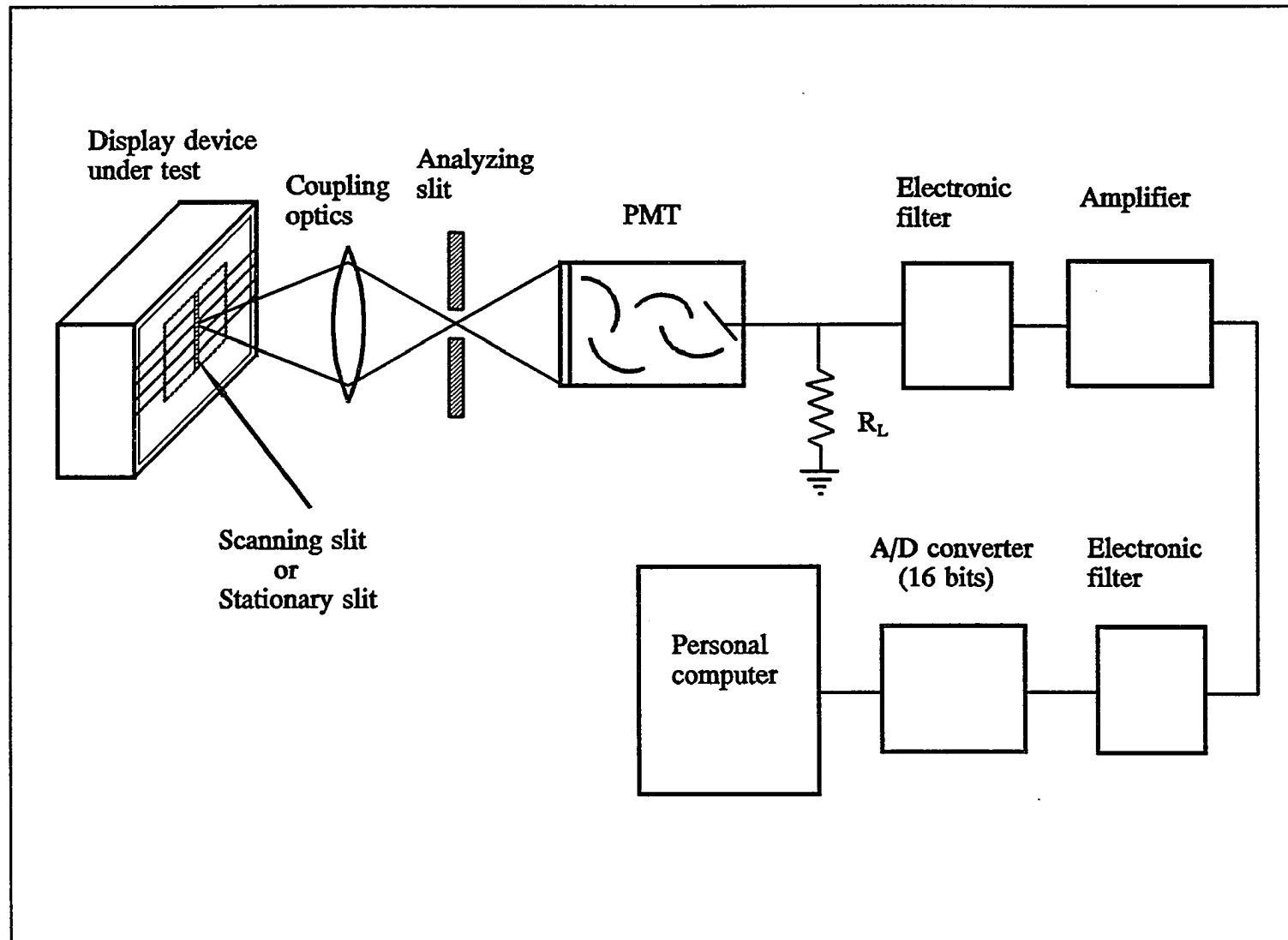


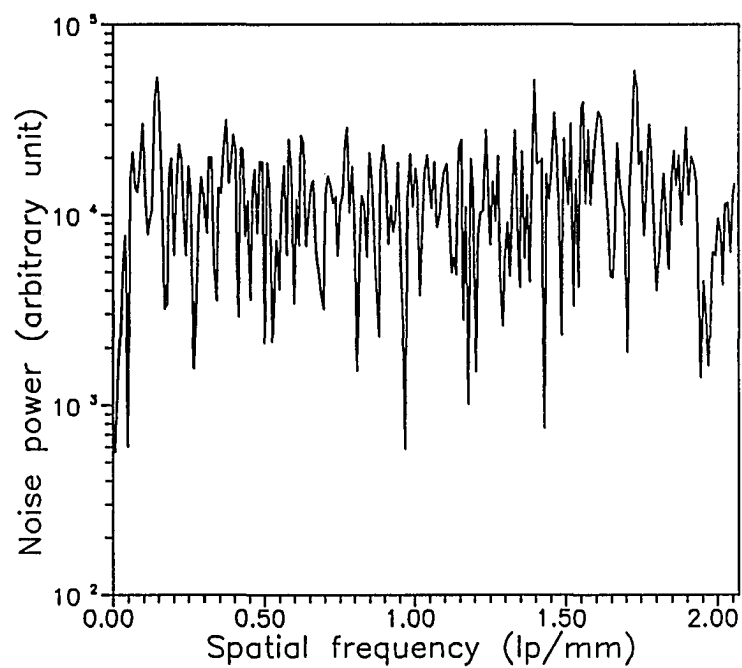
Fig. 2.7 Schematic for the measurement of display device noise.

digitized by a 16-bit A/D converter. The recorded raw data are first pre-processed to remove a slowly-varying mean (a "trend") from the raw data. The "trend"-removed data are then used to calculate luminance SNR and noise power spectra. It is assumed that the temporal noise of all pixels follow the same statistics and is independent of the spatial noise. Consequently, the variance of the total noise measured from the scanned data is the sum of the variances of the spatial and the temporal noises.

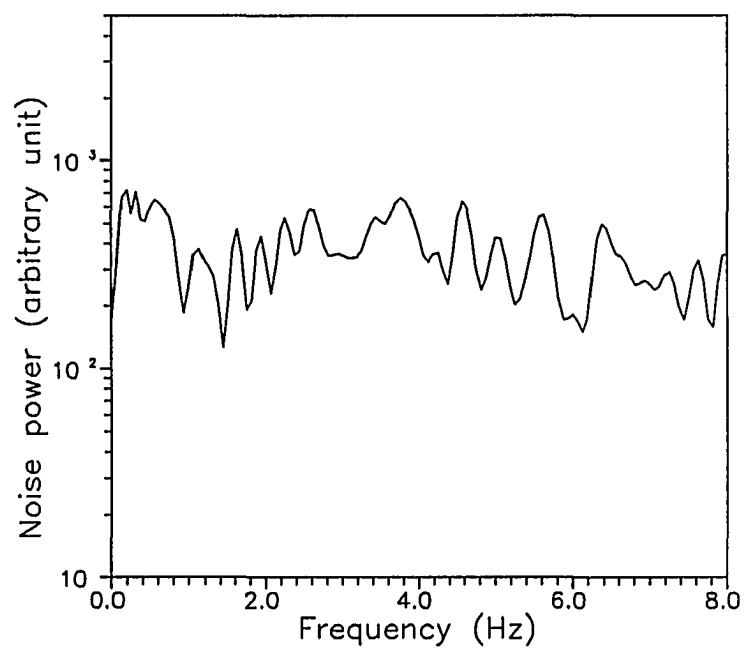
The results of the noise measurement show that the spatial noise is the dominant noise component over almost the entire dynamic range of the display device, except at very low grey levels where the temporal noise is comparable to the spatial noise. Both the spatial and the temporal noises are apparently white noise processes, suggesting that the luminance fluctuations from one pixel to another are uncorrelated. Based on this, it is convenient to compute the luminance signal-to-noise ratio per pixel, SNR_p , from the signal-to-noise ratio for n pixels as described above by

$$SNR_p = SNR/n^{1/2} \quad . \quad (2.7)$$

As examples, Fig. 2.8 shows the spatial and the temporal noise power spectra of a CRT. It can be seen from the figure that they are basically white noise processes. Fig. 2.9 shows the spatial and the temporal SNR_p as a function of displayed luminance for two CRTs. It is clear from the figure that the spatial SNR is much smaller than the temporal SNR (i.e., the spatial noise is the dominant noise component for CRTs). It can also be observed from Fig. 2.9 that the spatial SNR is independent of the luminance level while temporal SNR is dependent on the luminance level. This is what we would expect because the spatial noise comes from the phosphor granularity (or film granularity),

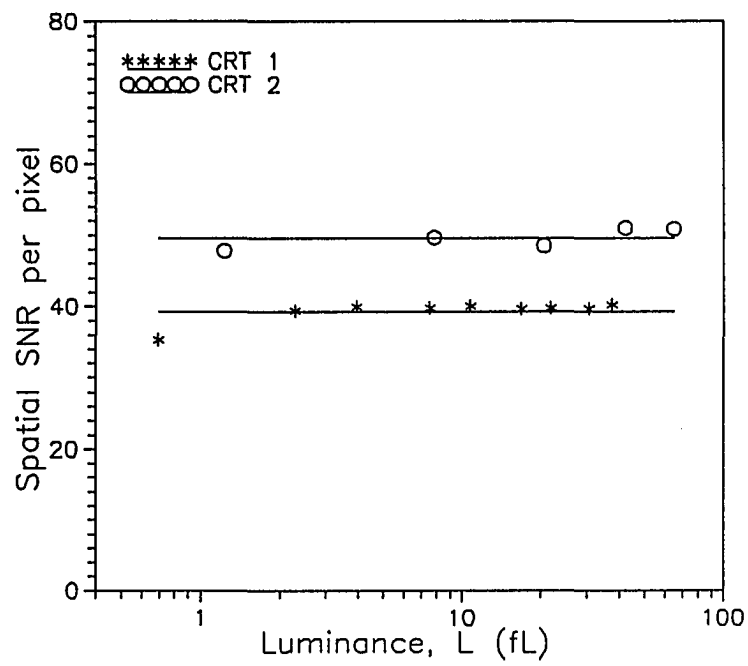


(a)

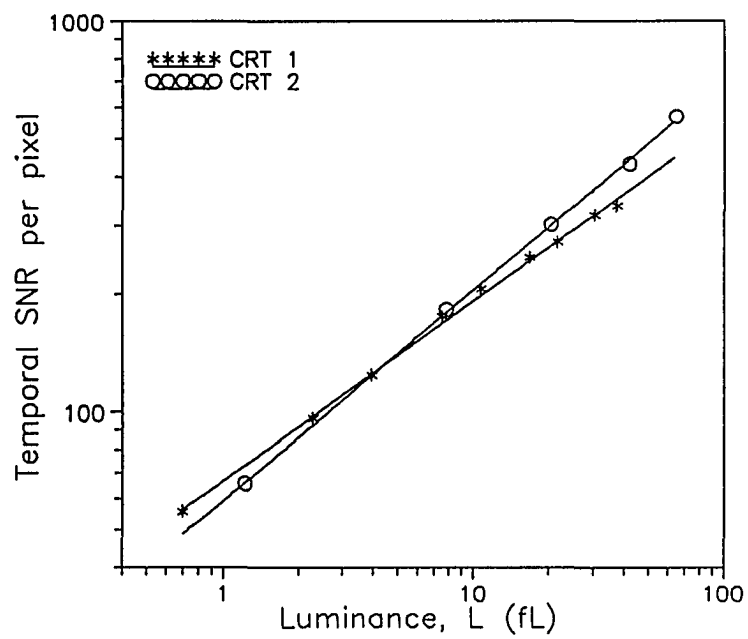


(b)

Fig. 2.8 (a) The spatial noise power spectrum of a CRT, (b) The temporal noise power spectrum of the same CRT.



(a)



(b)

Fig. 2.9 (a) Spatial and (b) temporal SNR per pixel for two CRTs.

which is determined by the fixed distribution of the phosphor particles.

2.7 Contrast Transfer Factor and Luminance Uniformity

Given the intensity of a signal (object), I_O , and the intensity of a background, I_B , the contrast is defined as

$$C = \frac{|I_O - I_B|}{\text{MAX}(I_O, I_B)} \quad (2.8)$$

This definition guarantees $0 \leq C \leq 1$. If the contrast of interest is small, then there is little difference between choosing I_O or I_B as the denominator in the above equation. As the intensities of the object and of the background can be different, one can define a variety of contrasts, such as grey level contrast or luminance contrast. For an N -bit digital display buffer, the smallest grey level contrast is $1/(2^N-1)$. There is an intrinsic small contrast transfer relation between the input and the output due to the non-linear property of the display function. This contrast transfer factor (CTF) is defined as the ratio of the output luminance contrast to a small input grey level contrast. Let $C_{in} = \Delta GL/GL$ be the input contrast and $C_{out} = \Delta L/L$, the output contrast. CTF can then be easily derived from the display function $f(GL)$ as follows.

If ΔGL is relatively small, we have

$$\Delta L \approx f'(GL) \cdot \Delta GL$$

where $f'(GL)$ is the first derivative of $f(GL)$. Dividing both sides of the above equation

by L , we get

$$\frac{\Delta L}{L} \approx \frac{f'(GL) \cdot \Delta GL}{L} = \frac{f'(GL) \cdot GL}{f(GL)} \cdot \frac{\Delta GL}{GL} .$$

Hence, the transfer factor of the small contrast is

$$CTF = \frac{C_{out}}{C_{in}} \approx \frac{f'(GL) \cdot GL}{f(GL)} . \quad (2.9)$$

Most of the CRTs have approximately a power-law display function: $L = a \cdot GL^\gamma$.

Hence, from (2.9), the CTF is simply equal to γ .

Although spatial noise characterizes the mid to high spatial frequency fluctuation of luminance of a display device, the large-scale spatial luminance variation across the entire display screen is evaluated by the luminance uniformity, which is the luminance measured over a relatively large area (e.g., 1 cm^2) as a function of coarsely sampled location on the display screen when all pixels are addressed with the same grey level. The peak-to-peak variation of the luminance relative to the average luminance (in percentage) is often quoted as a measure of luminance uniformity. Obviously, the smaller the peak-to-peak luminance variation, the better is the luminance uniformity.

2.8 Summary

From the analysis given in this chapter, it can be seen that the physical quality measures of an image display device can be generally classified into two types: contrast

resolution and spatial resolution. The discrimination of high contrast, high spatial frequency details in a given image is basically limited by the spot size (or MTF) of the single pixel of the image display device. On the other hand, the ability to distinguish a relatively large object with very small luminance difference from its background is mainly determined by such characteristics as physical dynamic range, display function and luminance noise. However, contrast resolution and spatial resolution are also related to each other because some characteristics have effect on both quality measures. For example, increasing the maximum intensity of beam current may enhance the physical dynamic range of a CRT, but it may as well reduce the CRT's MTF, since a large beam current usually causes a large spot size. For CRTs, the problem of internal scatter decreases both contrast and spatial resolutions. Table 2.1 compares a part of the measured results of the physical parameters for several CRT monitors evaluated in our research laboratory.

Since the output of an image display device is directly presented to the human observer, it is of course interesting to ask how these physical characteristics of the image display system affects the perceptual response of the human observer and how to design or adjust these physical parameters to best meet the needs of the human visual system in order to extract as much useful information as possible out of the image presented by the display device. To answer these questions, it is necessary to investigate the human visual system and psychophysically evaluate the image display system, which is the topic of the next chapter.

TABLE 2.1 Comparison of CRT monitors

	MegaScan	Tektronix	US Pixel	Sony
General				
Active screen size (H × V)	13.5"×10.8"	10.5"×14"	8.5"×13"	19.5"×19.5"
Addressable pixel matrix (H × V)	2560×2048	1536×2048	1024×1536	2048×2048
Frame rate (Hz)	72	60	40	60
Scan mode (Non interlaced/Interlaced)	N	N	I	N
Luminance and dynamic range				
Maximum luminance (ft-L)	65	39.3	87.5	39.1
Absolute dynamic range	889:1	2977:1	7353:1	20820:1
Useful dynamic range (with SMPTE pattern)	138:1	276:1	571:1	285:1
Luminance non-uniformity (peak-to-peak variation at peak luminance)	7%	7%	23%	11%
Spatial resolution				
Nominal pixel size (mm)	0.134	0.174	0.215	0.242
Actual pixel size at center of screen with peak luminance, FWHM (mm)				
Horizontal	0.188	0.191	0.296	0.312
Vertical	0.132	0.282	0.299	0.297
MTF at Nyquist frequency, center of screen, peak luminance				
Horizontal	0.035	0.37	~0	0.301
Vertical	0.068	0.30	~0	0.263
Total S/N per pixel at peak luminance				
	22.6	20.5	34	40

CHAPTER 3

PSYCHOPHYSICAL EVALUATION OF IMAGE DISPLAYS

3.1 Introduction

Since the ultimate purpose of any image display device is to provide useful and appropriate information to observers using the display device, any evaluation of the image display device without considering the human visual system is certainly incomplete. The subject of the human visual system is very broad and complex. Although a complete discussion of the human visual system is not attempted in this dissertation, a brief review will be given in the next section with an emphasis on how the human visual system responds to the displayed luminance, contrast and spatial frequencies. Detailed overviews and analyses of the human visual system can be found in the works of Cornsweet [40], Graham [41], and Campbell and Robson [42]. The concepts of the threshold contrast and the Just-Noticeable-Difference (JND) are introduced and analyzed in §3.3. These two parameters serve as a bridge that connects the physical characteristics of the image display device and the perceptual response of the human visual system. The visual response function which relates the displayed luminance and the perceived brightness is derived from the threshold contrast function in §3.4. With the knowledge of the visual response function, we are able to calculate the perceived dynamic range, which is a better measure of the contrast resolution of the display device, and design an optimum display function to match the human visual system.

3.2 Human Visual System

The Human visual system is a complex system which has not yet been understood completely. There is a tremendous amount of literature about this system. As far as the psychophysical evaluation of display devices is concerned, we are more interested in the characteristics of the human visual system's response to various light stimuli. Color vision will not be included here because we are only dealing with the evaluation of monochrome display devices.

It has been known that the human visual system is a non-linear system which has a very large dynamic range and a property of a band-pass spatial filter. Hall and Hall [43] have proposed a non-linear system model of the human visual system which is composed of a low-pass filter followed by a logarithmic transformation plus a high-pass filter. The band-pass property of the human visual system is best characterized by the so-called Contrast Sensitivity Function (CSF) which describes the sinusoidal response of the human visual system. A typical contrast sensitivity function (usually the reciprocal of the CSF, the threshold contrast, is measured) is shown in Fig. 3.1. The contrast sensitivity of the human visual system is a function of both the spatial frequency of the sine-wave pattern and the average luminance of the pattern. It also depends slightly on the orientation of the pattern [45]. From Fig. 3.1 it can be seen that the human visual system is most sensitive at high luminance levels to sine-waves of the spatial frequency about $3 \sim 5$ cycles/degree. The lower and the upper limit of the spatial frequencies are about 0.1 cycles/degree and 50 cycles/degree, respectively. Ginsburg [46] has suggested

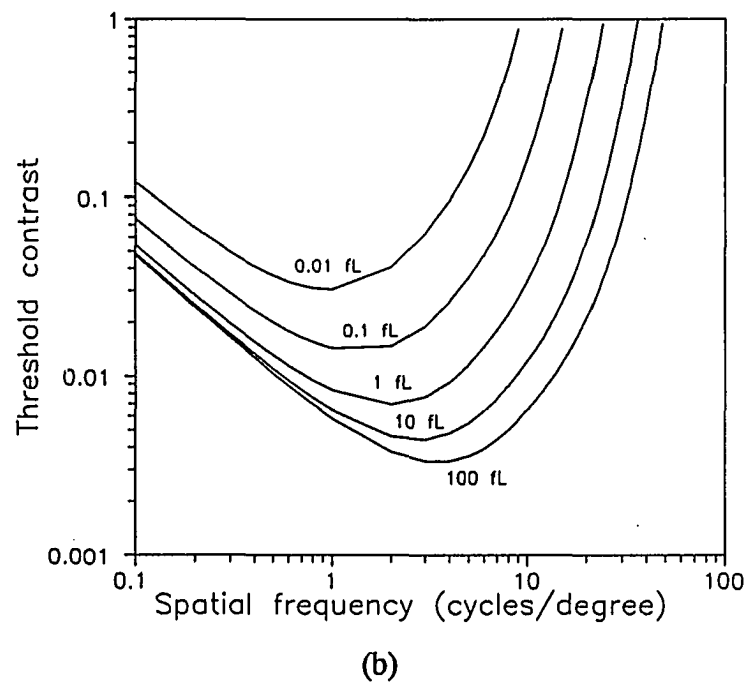
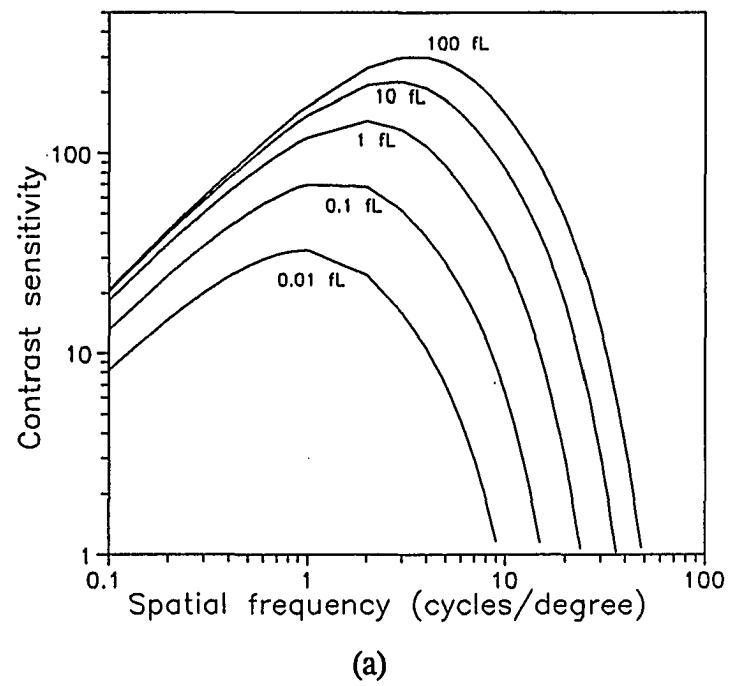


Fig. 3.1 (a) Contrast sensitivity function and (b) its reciprocal at various luminance levels. (These curves are calculated based on the mathematical model given by Barten [43]).

that the human visual system implements the band-pass filtering through many quasi-independent narrow-band channels, tuned to spatial frequencies over the entire bandwidth of the human visual system. He demonstrated that it is the lower spatial frequencies rather than the higher spatial frequencies that convey major information about objects in an image. It seemed that objects containing only 4 cycles per object size are already adequate to be identified from each other, and the higher spatial frequencies of objects are redundant and only make objects look sharper. This conclusion seems to have been supported by the psychophysical study conducted by Seeley *et al.* [47] who found no significant difference for radiologists to make diagnosis with the limiting resolution of chest images varying from 1.25 lp/mm to 6 lp/mm although the average subjective rating of the image quality improved with increasing spatial resolution.

The threshold contrast as a function of background luminance (that is the vertical cross-section in Fig. 3.1(b) for the case of the sinusoidal patterns) is also of great interest because it characterizes the non-linear property of the human visual system. How the human visual system responds to the displayed luminance is particularly important for evaluating the display function, the contrast resolution, and the perceived dynamic range of image display devices.

3.3 Threshold Contrast and Just-Noticeable-Difference (JND)

While the contrast is a physical quantity of the visual stimulus, the threshold contrast is a psychophysical quantity which is not only dependent on the physical

property of the stimulus, but also on human perception. The threshold contrast is defined as the smallest contrast that can be just perceived by the human observer with a certain detection probability. It is immediately clear that the threshold contrast measured with a large detection probability will be smaller than the one measured with a small detection probability. Blackwell [13], for instance, used a detection probability of 50% to define all the threshold contrasts he measured in his well-known experiment conducted many years ago. Without external noise, the threshold contrast depends mainly on two factors: the background luminance and the object size, as shown in Fig. 3.1(b). For a given size of the object within the pass-band of the human visual system, the dependence of the threshold contrast C_T on the background luminance L is referred to as the Weber-Fechner relationship [14]. As shown in Fig. 3.2, the curve can be divided into two asymptotic regions. The region where C_T decreases as the background luminance increases is often called the Rose-de Vries region (also termed scotopic region), because in this region the light level is so low that the threshold contrast is limited by the photon noise (i.e., the fluctuation of the number of photons) and hence C_T is proportional to $1/L^{0.5}$. The region in which C_T is independent of the background luminance is known as the Weber region (also termed photopic region), since in this region Weber's law holds:

$$C_T = \Delta L_T / L = C_0 = \text{constant} .$$

The lowest threshold contrast C_0 is limited only by the internal noise of the human observer's eye-brain system. The value of C_0 can be as low as 0.003 for a detection probability of 50% according to Blackwell's measurement. However, since an image display system will introduce extra noise as analyzed in the previous chapter, one can

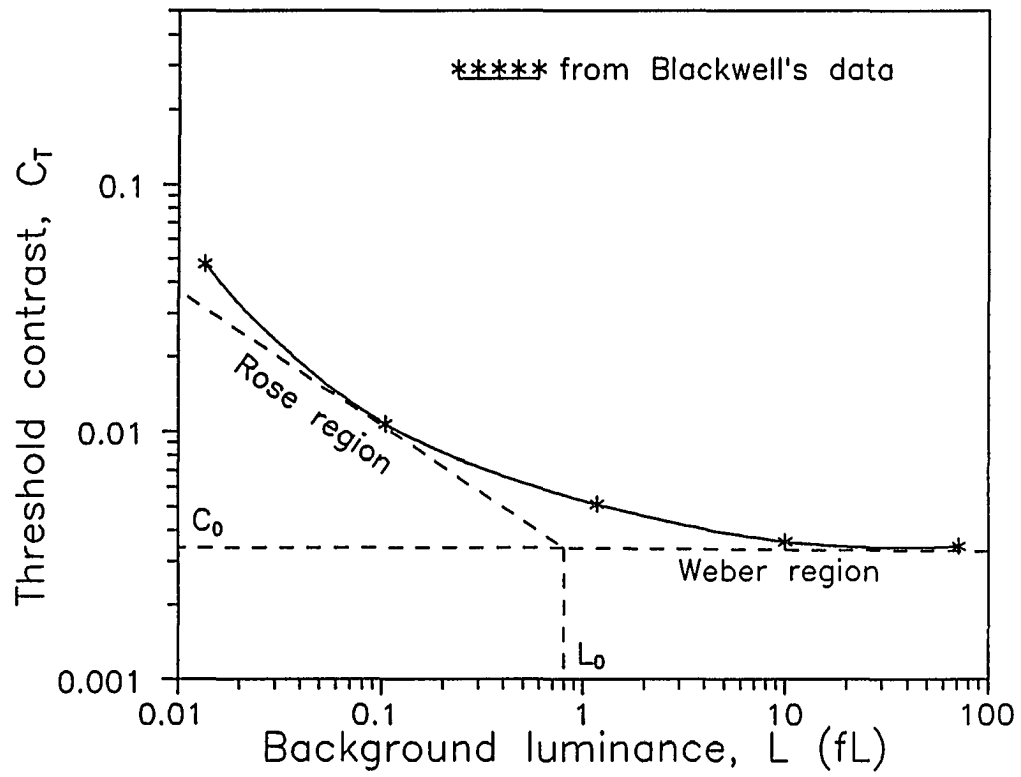


Fig. 3.2 A typical threshold contrast curve (Weber-Fechner relationship).

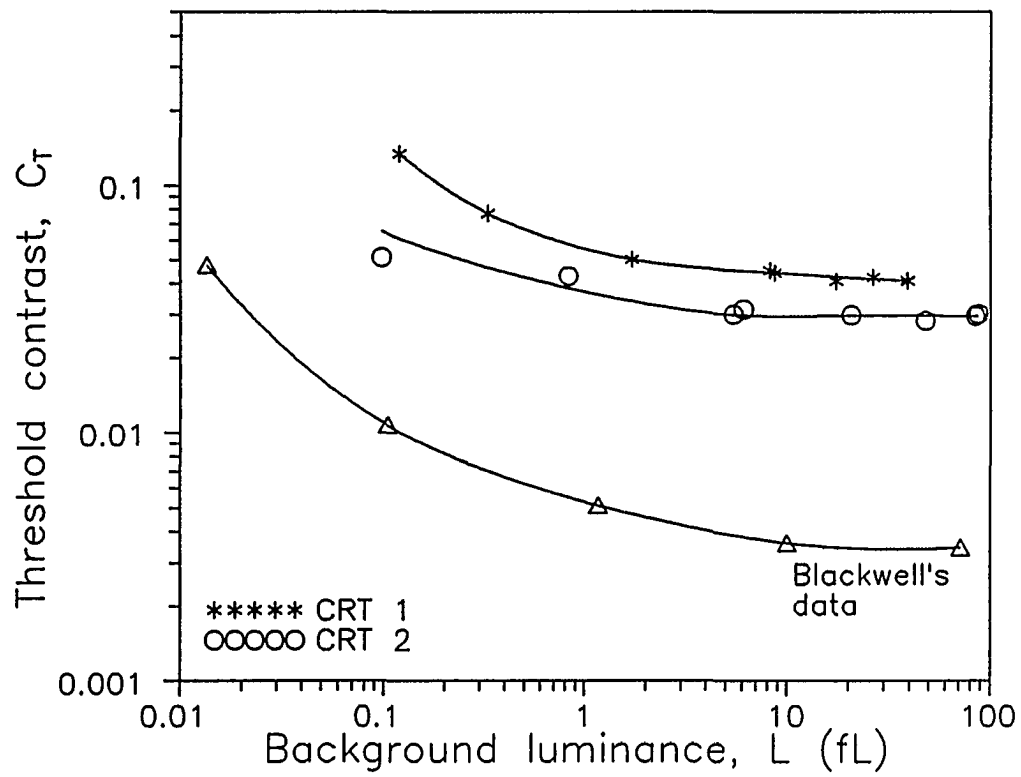


Fig. 3.3 Measured threshold contrast for two CRTs as a function of background luminance for detection of a square object of 0.3 degree of arc, compared with the Blackwell's data.

expect a larger value of C_0 . Fig. 3.3 shows the measured threshold contrast curve for two state-of-the-art CRTs. For comparison, the threshold contrast curve with only the effect of the internal noise of the eye-brain system is also plotted for the same size of the object, based on Blackwell's data [13]. It is clear from the figure that the threshold contrast for CRTs is much higher than blackwell's curve, indicating that for these CRTs, the threshold contrast is not limited by the internal noise of the human visual system, rather by the noise of the display device. It is also interesting to note that although the source of the dominant noise is totally different, the shape of the threshold contrast curve for CRTs is very similar to that given by the Weber-Fechner relationship; that is, the C_T curve is constant for high background luminance levels and a decreasing function of L for low background luminance levels.

Based on the Weber-Fechner relationship and all available experimental results, an analytical equation which can best fit the above features of the threshold contrast curve is proposed:

$$C_T(L) = C_0 \left(1 + (L_0/L)^{pm} \right)^{1/m} \quad (3.1)$$

where C_0 is the limiting threshold contrast which depends on the noise level and the object size, L_0 is the "turning" point between the two regions of the C_T curve, $-p$ is the slope of the curve in the decreasing region in a double-log plot with $0 < p < 1$ (in particular, if C_T is limited by the photon noise then $p = 0.5$). m is the parameter that controls the sharpness of the transition region around L_0 with values in the range $1 \leq m < \infty$. For a finite m , we have

$$C_T(L) = \begin{cases} C_0 \left[\frac{L_0}{L} \right]^p & L \ll L_0 \\ C_0 & L \gg L_0 \end{cases} .$$

If $m \rightarrow \infty$, $C_T(L)$ has no transition part and becomes a two-piece curve:

$$\lim_{m \rightarrow \infty} C_T(L) = \begin{cases} C_0 \left[\frac{L_0}{L} \right]^p & L < L_0 \\ C_0 & L \geq L_0 \end{cases} .$$

The Just-noticeable-difference (JND) is the smallest difference that a human observer can just detect with a certain detection probability when an object of a certain size is displayed at a given uniform background. According to this definition, the threshold contrast $C_T(L)$ and the luminance JND are simply related by

$$C_T(L) = JND/L . \quad (3.2)$$

Unlike the threshold contrast, JND has its own unit, depending on the quantity of the difference. We may have a luminance JND, a density JND or a grey level JND. From now on, we put a subscript L , D or GL to distinguish among these JNDs. Thus,

$$JND_L = L \cdot C_T(L) . \quad (3.3)$$

For CRTs, the luminance JND is usually used, while for laser film printers, the density JND, or JND_D , is more convenient. The grey level JND, JND_{GL} , is the input JND corresponding to the output JND_L or JND_D . Since JNDs are always positive and are relatively small, we have

$$JND_L \approx |f_L'(GL)| \cdot JND_{GL} \quad \text{or} \quad JND_D \approx |f_D'(GL)| \cdot JND_{GL}$$

where $f_L'(\cdot)$ and $f_D'(\cdot)$ are the first derivatives of the display functions for a CRT and a laser film printer, respectively.

Since the density and the transmission of the film are related by $D = \log_{10}(1/T)$, we have

$$dD = -\frac{dT}{T \cdot \ln 10} \quad (3.4)$$

The luminance L and the film transmission T are equivalent (i.e., $L \propto T$). So,

$$dT/T = dL/L \approx C_T(L) \quad (3.5)$$

Substituting (3.5) into (3.4) and noting that $JND_D \approx |dD|$, we obtain

$$JND_D = \frac{C_T(L)}{\ln 10} \quad (3.6)$$

With a simple derivation, JND_D can also be expressed in terms of optical density D :

$$JND_D = \Delta D_0 \left(1 + 10^{(D-D_0)pm} \right)^{1/m} \quad (3.7)$$

where $D_0 = -\log_{10} T_0$ is the counterpart of L_0 for film and $\Delta D_0 = C_0 / \ln 10$.

The threshold contrast (or JND) can be acquired from well designed psychophysical experiments. Either the two-alternative forced choice or adjustment method can be used to determine the threshold []. For the two-alternative forced choice method, an object (stimulus) of a certain size with a given contrast against a uniform background is randomly present or absent on the display device under the test. The

observers are asked to make the decision of whether or not the object is present. For the case when the object is present, the detection probabilities are calculated and averaged for all observers. Due to the spurious response probability P_S resulting from guesses, the raw data of the detection probability, P_R , has to be corrected to obtain the true detection probability P_D using the formula [48]:

$$P_D = (P_R - P_S)/(1 - P_S) . \quad (3.8)$$

Sometimes a multiple-choice method is used instead of the two-alternative method. An object is randomly located at one of n pre-specified positions. The observer is asked to report the object position whether he could see it or not. In this case $P_S = 1/n$. Varying the contrast of the object, one can obtain a curve of the contrast versus detection probability. The threshold contrast for the given object size is then defined as the contrast with which the object is found with a detection probability of 50%. The same procedure is repeated for various backgrounds. Tremendous amounts of time are generally needed for obtaining just one curve of $C_T(L)$.

For the adjustment method, the contrast of the object can be adjusted by the observer. The object of a given size is initially displayed to the observer with zero contrast. The observer increases the contrast so that he or she can just see the object, then decreases the contrast to make the object just undetectable. Averaging the two contrast values yields the threshold contrast which is considered to correspond to a detection probability of 50%. Again, the procedure needs to be repeated for different backgrounds. The second method takes less time than the first method, but is still very time-consuming.

In general, the threshold value obtained by different observers varies. Even for the same observer, the obtained results from repeated experiments are not likely a constant. This is due to the criterion fluctuation of each observer and the criterion variation between different observers. Hence, such psychophysical experiments as described above require many observers and several repeated observations for each observer to reduce the variance of the result. Some researchers have been investigating for years the ideal observer model to replace real observers for this tedious task [49]-[51]. If it is realized in a good agreement with real observers, the determination of the threshold contrast can be done much faster and in a more repeatable manner. In the next chapter, it will be shown that there is a relationship between the threshold contrast and the signal-to-noise ratio of the display device. Therefore, one can derive the threshold contrast (or JND) from the noise measurement which is less time-consuming than the above described psychophysical experiment.

The threshold contrast $C_T(L)$ (or JND) is the most important psychophysical parameter. It is the basis for the derivation of the relationship between the perceived brightness of the human visual system and the luminance displayed by a display device. The real contrast resolution characterized by the perceived dynamic range (PDR) (discussed in the next section) cannot be determined without the knowledge of the threshold contrast (or JND).

3.4 Visual Response Function and Perceived Dynamic Range (PDR)

As mentioned before, the human visual system responds non-linearly to the displayed luminance. It is commonly known that for high light levels the perceived brightness is a logarithmic function of the displayed luminance. That means, when the displayed luminance changes by an order of magnitude, the human visual system only perceives a brightness change of a factor of two. The reason for this will be given in this section. If we denote the perceived brightness by S , the relationship between S and the displayed luminance L is referred to as the visual response function. Let us use $h(L)$ to represent this non-linear function. If we define the perceptual threshold difference in S as ΔS and let ΔS equal one perceptual unit (i.e., $\Delta S = 1$), then when S changes by ΔS , the luminance L should correspondingly change by one JND_L which is a function of L . Therefore, we can derive every discrete point L_i which corresponds to $i \cdot \Delta S$ in the following way: Assume that the luminance range of a display device extends from L_{\min} to L_{\max} , starting at $L_0 = L_{\min}$, then $L_1 = L_0 + JND_L(L_0)$, $L_2 = L_1 + JND_L(L_1)$, etc. until the maximum luminance is reached: $L_n = L_{\max}$. To derive an analytical expression for $S = h(L)$, let us consider that within any one JND the function $h(L)$ can be approximated to be linear because one JND is generally very small compared to the entire display range $[L_{\min}, L_{\max}]$. The slope of this small linear piece of $h(L)$ is

$$\frac{\Delta S}{JND_L(L)} = \frac{1}{JND_L(L)} \quad ,$$

which approximately corresponds to the first derivative of $h(L)$, i.e.,

$$\frac{dS}{dL} = \frac{1}{JND_L(L)} \quad . \quad (3.9)$$

Integrating (3.9) leads to the expression for $h(L)$:

$$h(L) = \int_{L_{\min}}^L \frac{dL}{JND_L(L)} = \int_{L_{\min}}^L \frac{dL}{LC_T(L)} \quad . \quad (3.10)$$

From (3.10) we can see that (1) the visual response function is completely determined by the JND or the threshold contrast, and (2) $S = h(L)$ is actually a relative function because (3.10) has set $h(L_{\min}) = 0$ no matter what the absolute value of L_{\min} is. In other words, we set the lowest luminance corresponding to the origin of the perceived brightness for a given display device. Sometimes, it is more convenient to use the normalized visual response function which is simply given by

$$h_N(L) = \frac{h(L)}{S_{\max}} \quad (3.11)$$

where $S_{\max} = h(L_{\max})$. Fig. 3.4 plots an example of $h(L)$ with $C_T(L)$ given by (3.1). Fig. 3.5 - Fig. 3.8 show how $h_N(L)$ changes as a function of each parameter in (3.1) with all other parameters fixed. Two things should be noted about these figures. First, $h_N(L)$ does not depend on C_0 , which is what we expect from (3.10) and (3.11). This is a very important feature of $h_N(L)$ and suggests that the normalized visual response function is independent of those parameters which only influence C_0 . This point will be discussed later in more detail. Secondly, the effect of the parameter m which controls the transition sharpness of $C_N(L)$ is much smaller than the other two curve shape parameters,

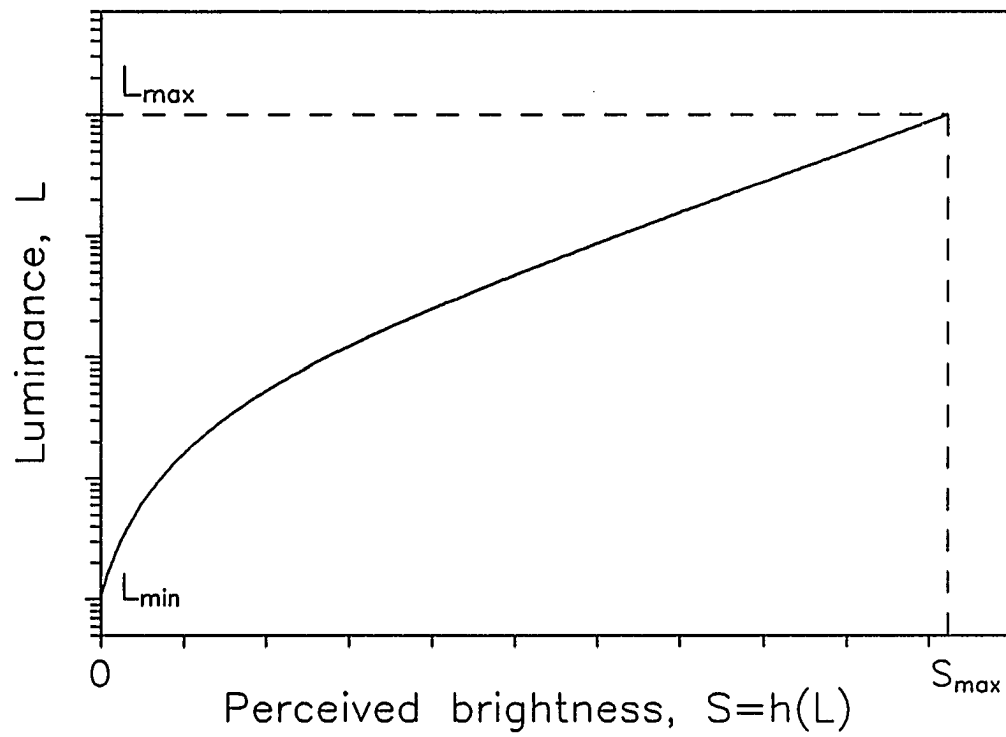


Fig. 3.4 *Illustration of the visual response function $S = h(L)$.*

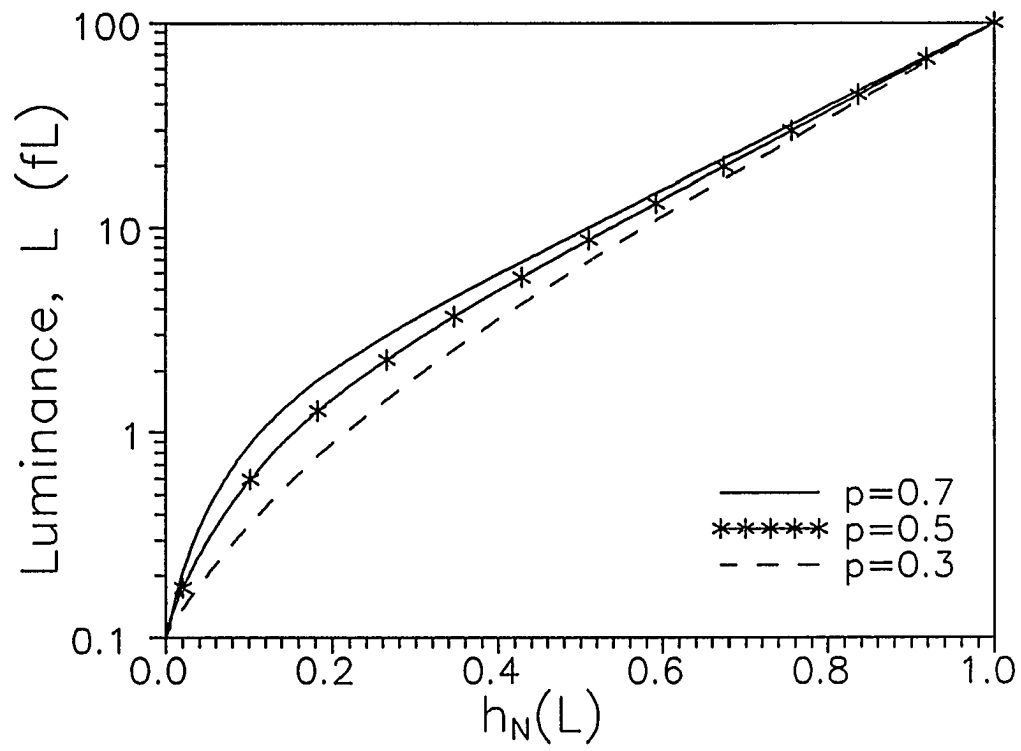


Fig. 3.5 Normalized visual response functions $h_N(L)$ for different values of p with other parameters fixed: $C_0 = 0.02$, $L_0 = 3 \text{ fL}$, $m = 4$. (luminance range: 0.1 - 100 fL).

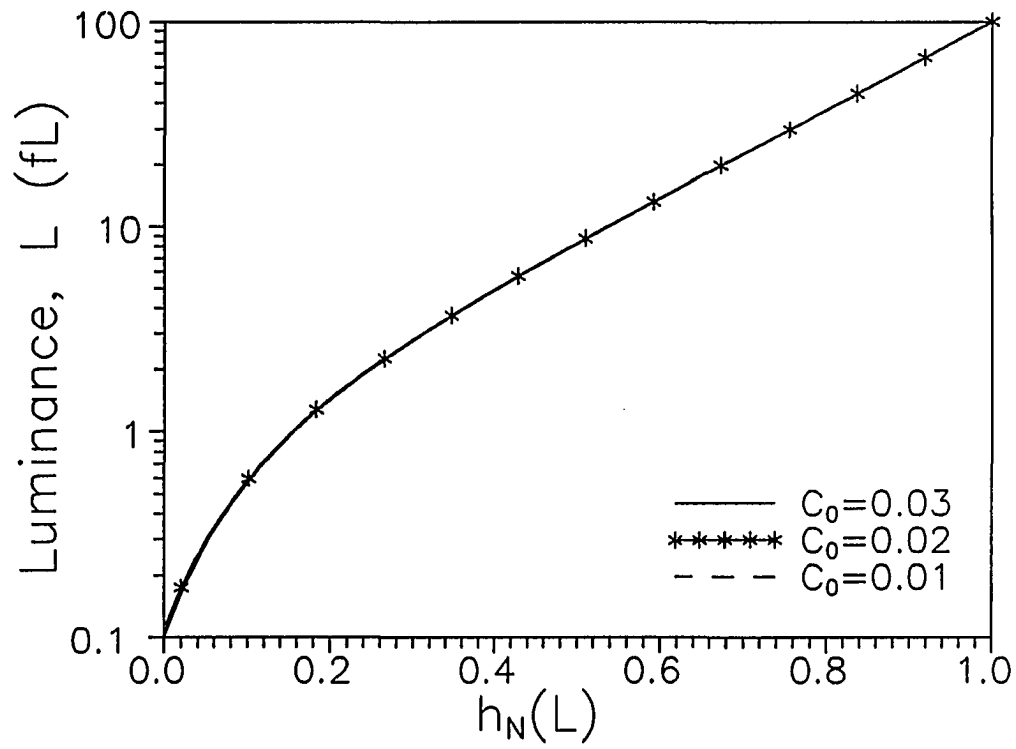


Fig. 3.6 Normalized visual response functions $h_N(L)$ for different values of C_0 with other parameters fixed: $L_0 = 3 \text{ fL}$, $p = 0.5$, $m = 4$. (luminance range: 0.1 - 100 fL).

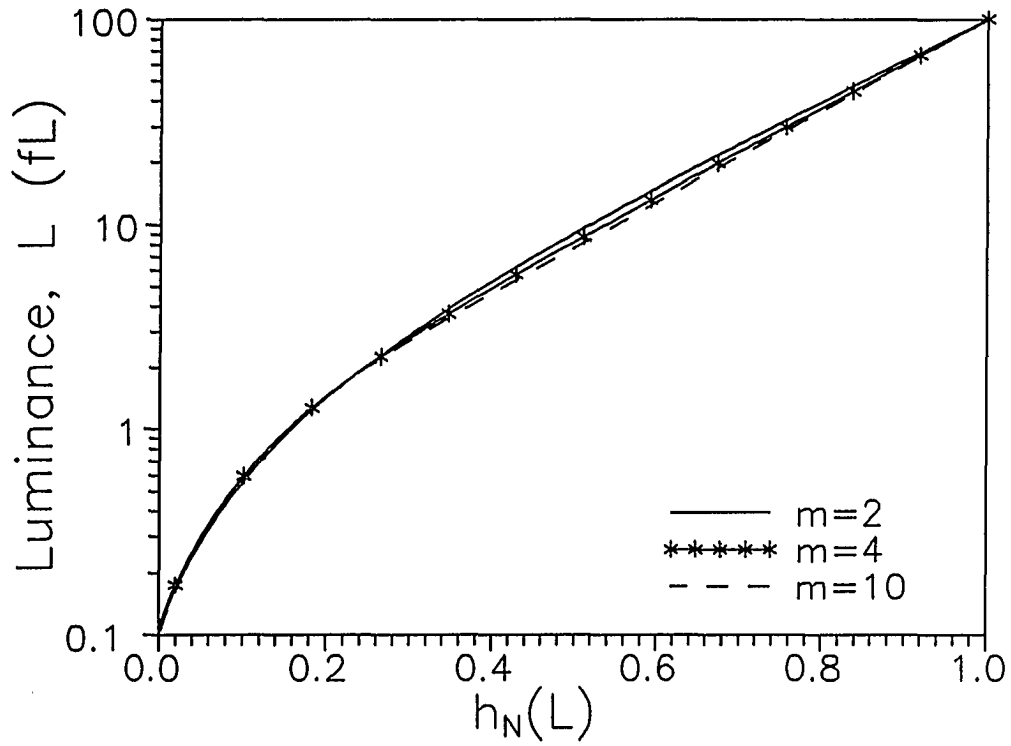


Fig. 3.7 Normalized visual response functions $h_N(L)$ for different values of m with other parameters fixed: $C_0 = 0.02$, $L_0 = 3 \text{ fL}$, $p = 0.5$. (luminance range: 0.1 -100 fL).

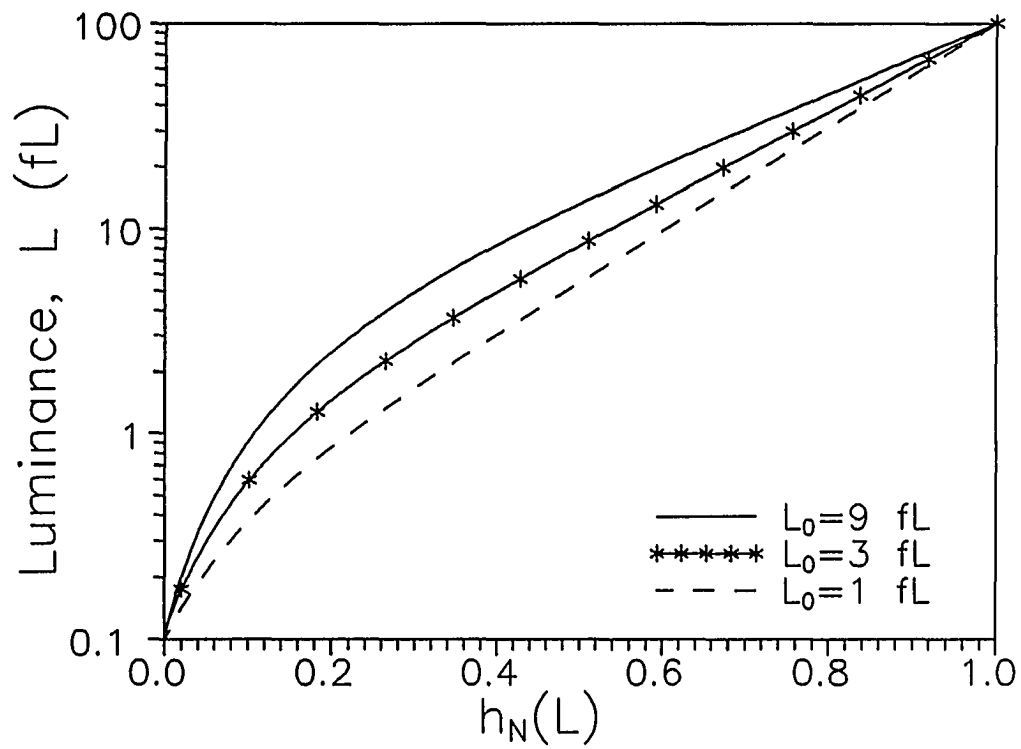


Fig. 3.8 Normalized visual response function $h_N(L)$ for different values of L_0 with other parameters fixed: $C_0 = 0.02$, $p = 0.5$, $m = 4$. (Luminance range: 0.1 - 100 fL).

p and L_0 . Thus, if we let $m \rightarrow \infty$, the limiting result of $h_N(L)$ will not be significantly different from the real $h_N(L)$. Specifically, using (3.1) we have

$$\lim_{m \rightarrow \infty} h(L) = \begin{cases} \int_{L_{\min}}^L (L^{p-1}/C_0 L_0^p) dL & L < L_0 \\ \int_{L_0}^L (C_0 L)^{-1} dL + \int_{L_{\min}}^{L_0} (L^{p-1}/C_0 L_0^p) dL & L \geq L_0 \end{cases} \quad (3.12)$$

This expression can be simplified to the form

$$\lim_{m \rightarrow \infty} h(L) = \begin{cases} \frac{1}{C_0 p} [(L/L_0)^p - (L_{\min}/L_0)^p] & L < L_0 \\ \frac{1}{C_0} [\ln(L/L_0) + \beta] & L \geq L_0 \end{cases} \quad (3.13)$$

where $\beta = [1 - (L_{\min}/L_0)^p]/p$. Hence, the perceived brightness is approximately a power function of L for $L < L_0$ and becomes a logarithmic function of L for $L > L_0$.

Clearly, when the upper limit of the integral (3.10) reaches L_{\max} , we obtain the total number of JNDs for the entire luminance range $[L_{\min}, L_{\max}]$. Pizer and Chan [15] have defined this quantity as the perceived dynamic range (PDR), that is

$$PDR = S_{\max} = \int_{L_{\min}}^{L_{\max}} \frac{dL}{JND_L(L)} \quad (3.14)$$

PDR is a more useful figure of merit for the contrast resolution of a display device because it really tells us how many distinguishable shades of gray can be achieved with the given display device for a given object. It can be seen from (3.14) that PDR is

influenced by both the physical dynamic range and the luminance JND curve.

The application of the above analysis to the laser printer/film combination is straightforward. The counterparts of (3.10) and (3.14) are

$$h(D) = \int_D^{D_{\max}} \frac{dD}{JND_D} \quad (3.15)$$

and

$$PDR = \int_{D_{\min}}^{D_{\max}} \frac{dD}{JND_D} \quad (3.16)$$

It is noted that $h(D)$ is a decreasing function of the film density D with $h(D_{\max}) = 0$ and $S_{\max} = h(D_{\min})$. If we apply the same approximation to $h(D)$ as we did to $h(L)$ (let $m \rightarrow \infty$), the result is

$$\lim_{m \rightarrow \infty} h(D) = \begin{cases} h(D_0) + (D_0 - D)/\Delta D_0 & D \leq D_0 \\ \frac{10^{D_0 p}}{\Delta D_0 p \ln 10} (10^{-D p} - 10^{-D_{\max} p}) & D > D_0 \end{cases} \quad (3.17)$$

where

$$h(D_0) = \frac{10^{D_0 p}}{\Delta D_0 p \ln 10} (10^{-D_0 p} - 10^{-D_{\max} p}) \quad (3.18)$$

D_0 and ΔD_0 are the same as in (3.7). $h_N(D)$ can also be defined the same way as $h_N(L)$.

(3.17) suggests that the perceived brightness is approximately a linear function of the film density for $D < D_0$ and exponentially decreases with the density D for $D > D_0$.

3.5 Summary

As discussed in this chapter, the threshold contrast, or equivalently the just-noticeable-difference, is the key parameter for the psychophysical evaluation of display devices. This is due to the fact that the visual response function which is particularly important for designing the optimum display function, and the perceived dynamic range which is the measure of the contrast resolution of the display device, can be derived from the threshold contrast. The establishment of an analytical expression for the threshold contrast based on the measured data enables us to acquire explicit formulas for the visual response function and the perceived dynamic range. Generally speaking, the visual response function is approximately a power function of the displayed luminance at low luminance levels and gradually changes to a logarithmic function of the luminance when the luminance level increases. With the mathematical expression for the visual response function, the effects of the display device noise on the human visual response can be more easily analyzed quantitatively, as long as the relation between the display device noise and the threshold contrast are known. This will be the central topic of the next chapter.

CHAPTER 4

CORRELATION BETWEEN PHYSICAL AND PSYCHOPHYSICAL EVALUATIONS

4.1 Introduction

A basic question of interest is: How is the perception of the displayed information influenced by the physical characteristics of a display device? Some correlations are obvious and simple, such as flicker and the refresh rate for a CRT, or the sharpness of the displayed image and the MTF of the display device. In this chapter, however, the attention is concentrated on the relationship between the threshold contrast and the display device noise because the contrast information is extremely important for the detection of objects. It must be emphasized that this relationship has not been investigated adequately in the past.

It is intuitive to think that the detection of small luminance differences by the human observer is dependent on the presence of noise. de Vries [16] and Rose [17] were the first to propose that it is the displayed signal-to-noise ratio that determines the human observer's ability to detect a signal in a noisy environment. They postulated that for the signal to be detected, the displayed signal-to-noise ratio should exceed some threshold value k_0 .

In the Rose model, it is assumed that the background has an average photon fluence Φ (photons per unit area) which is proportional to the luminance, and an object

with the area A is embedded in the background. Rose also assumes that both the object and the background areas contribute to the photon fluctuation observed by the human eye, and the eye perceives the difference between the photon count from the object area and the photon count from an area of equal size in the background. Therefore, the displayed signal is considered to be the photon count difference $(\bar{n}_o - \bar{n}_b)$, (or $(\bar{n}_b - \bar{n}_o)$, if $\bar{n}_b > \bar{n}_o$), where \bar{n}_o and \bar{n}_b are the average numbers of photons from the object area and the background area, respectively. The contrast of the object is defined as $C = \Delta\Phi/\Phi$. Obviously, $\bar{n}_o - \bar{n}_b = \Delta\Phi \cdot A = C\Phi A$. Since the photon numbers are Poisson random variables, their variances are \bar{n}_o and \bar{n}_b . The total RMS noise is given by

$$\sqrt{\bar{n}_o + \bar{n}_b} = \sqrt{A \cdot (2\Phi + \Delta\Phi)} \approx \sqrt{2A\Phi} \quad .$$

The approximation holds because $\Delta\Phi \ll \Phi$ for a small contrast considered here. The displayed signal-to-noise ratio SNR_D of the object with a small contrast is then given by

$$SNR_D = C \cdot \sqrt{\frac{A \cdot \Phi}{2}} \quad . \quad (4.1)$$

In fact, the Rose model treated the human eye as an averaging spatial integrator. Mathematically, its output is

$$\Phi = \frac{1}{A} \int_A \phi dA$$

which is the average photon fluence within the integrated area A . Thus, the signal is the

average difference in photon fluence $\Delta\Phi$. The variance of the photon noise is

$$\begin{aligned}
 \text{Var} \left(\frac{1}{A} \int_A \phi_o dA + \frac{1}{A} \int_A \phi_b dA \right) &= \frac{1}{A^2} \left[\text{Var} \left(\int_A \phi_o dA \right) + \text{Var} \left(\int_A \phi_b dA \right) \right] \\
 &= \frac{1}{A^2} (\bar{n}_o + \bar{n}_b) \\
 &= \frac{A(2\Phi + \Delta\Phi)}{A^2} \\
 &\approx \frac{2\Phi}{A} .
 \end{aligned}$$

Hence, the displayed signal-to-noise ratio is $\Delta\Phi/(2\Phi/A)^{1/2}$, which is exactly the same as (4.1). For the object to be detectable, SNR_D must exceed some threshold value k_0 , i.e.

$$SNR_D = C \cdot \sqrt{\frac{A}{2} \cdot \Phi} \geq k_0 .$$

Hence, the threshold contrast is given by

$$C_T = \frac{\sqrt{2} k_0}{\sqrt{A \cdot \Phi}} . \quad (4.2)$$

Eqn.(4.2) indicates that the threshold contrast is inversely proportional to the square-root of the area of object and the square-root of the photon fluence.

Experimental results from earlier investigations have shown that the probability of detection is directly related to the displayed signal-to-noise ratio as shown in Fig. 4.1 [52]. Rosell and Willson [52] have found a value of $k_0 = 3.1$ for a probability of detection of 50%, while a probability of detection of almost 100% required a k_0 value of in excess of 5. Since k_0 depends on the probability of detection, the threshold contrast

is also a function of the probability of detection.

If there were only photon noise, as assumed in the Rose model, any arbitrarily small threshold contrast could be reached as long as the photon fluence is large enough. However, practical threshold contrast can never be arbitrarily small because there are other noise sources which limit the threshold contrast. In real display/human observer modalities the noise mainly arises from three sources characterized by their standard deviation: photon noise (σ_{ph}), noise of the display device (σ_D), and the internal noise (neural noise) coming from the observer's eye-brain system (σ_{eye}). With the assumption that these noise sources are independent, the total noise standard deviation is given by

$$\sigma_{tot} = (\sigma_{ph}^2 + \sigma_D^2 + \sigma_{eye}^2)^{1/2} .$$

Depending on which noise is the dominant noise, σ_{tot} may be approximated by any one of the above three noise sources. It is clear that the Rose model holds only for photon-noise-limited case where the photon noise is the dominant noise. From (4.2) we see that C_T is inversely proportional to the square-root of the background luminance (noting that the luminance is proportional to Φ). This is why the $C_T(L)$ curve in a double-log plot has a slope of -0.5 in the Rose region (see Fig. 3.2 in Chapter 3). On the other hand, the internal noise is considered to be proportional to the displayed background luminance. Hence, σ_{eye}/σ_{ph} is proportional to $L^{1/2}$. As background luminance increases, the effect of photon noise becomes smaller and smaller, and the internal neural noise gradually becomes the dominant component, if there is no display device noise. Therefore, the threshold contrast C_T follows the well known Weber-Fechner relationship introduced in §3.3, that is, the threshold contrast decreases with background luminance, following the

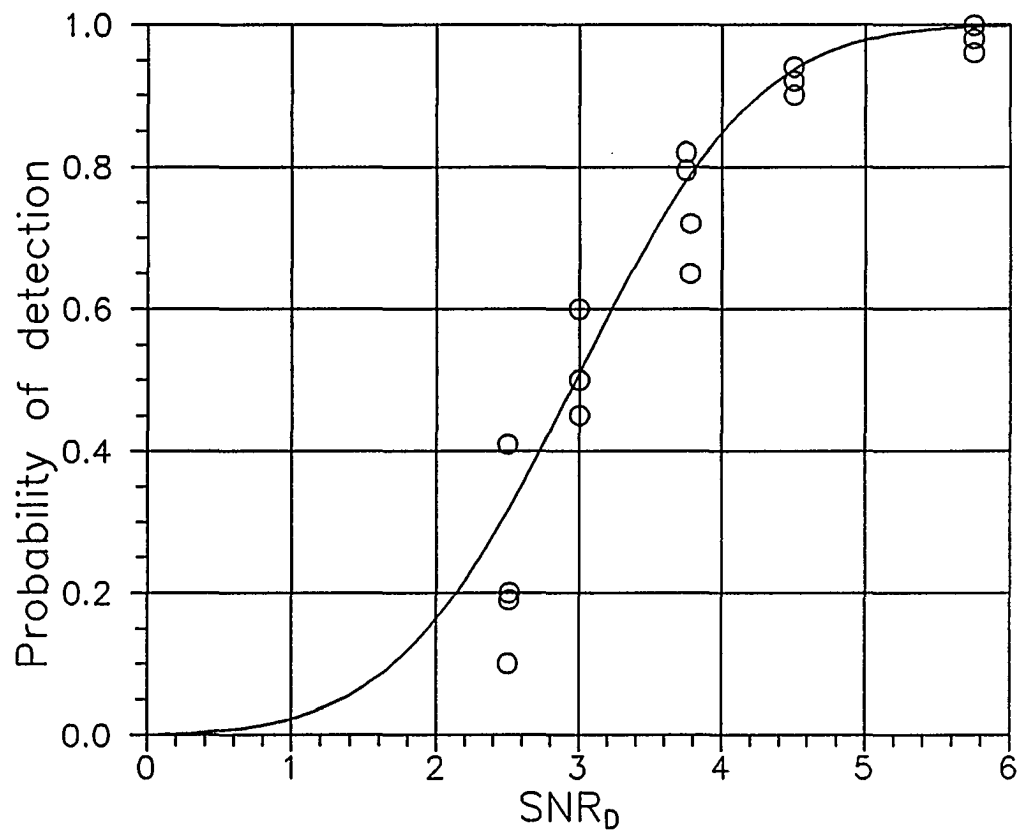


Fig. 4.1 *Probability of detection as a function of displayed signal-to-noise ratio.*

Rose model, at low background luminance levels and becomes a constant at high background luminance levels. This typical threshold contrast curve has been verified by Blackwell's experimental results [13].

4.2 Effect of Display Device Noise on the Threshold Contrast

Now let us consider the case where the object detection is limited by the display device noise which, as indicated in §2.6, mainly comes from the phosphor or film granularity. That is, σ_D is the dominant noise compared with the other two types of noises. It is known from §2.6 that the total display device noise, presented as the luminance fluctuation, is generally a function of the background luminance. Assume that the object consists of n pixels. We denote by σ_p the RMS value of the luminance noise per pixel at the background luminance L_b . As the Rose model suggested, the eye performs both spatial integration and averaging. The luminance sensed by the eye is an average luminance over the luminance values from all pixels that the eye integrates. Let $L_{j,o}$ and $L_{j,b}$ denote the luminance of the j th pixel in the object and the background areas, respectively. The displayed signal is

$$AVG \left[\frac{1}{n} \sum_{j=1}^n L_{j,o} - \frac{1}{n} \sum_{j=1}^n L_{j,b} \right] = L_o - L_b = \Delta L \quad (4.3)$$

where ΔL is the average luminance difference between the object and the background. Also, it is clear that $C_T = \Delta L/L_b$. The total luminance noise in terms of its variance is

given by

$$\sigma^2 = \text{Var} \left[\frac{1}{n} \sum_{j=1}^n L_{j,o} - \frac{1}{n} \sum_{j=1}^n L_{j,b} \right] . \quad (4.4)$$

It is shown in §2.6 that the luminance fluctuations are generally uncorrelated among the pixels. So (4.4) can be written as

$$\sigma^2 = (\sigma_o^2 + \sigma_b^2)/n \quad (4.5)$$

where $\sigma_o^2 = \text{Var}(L_{j,o})$ and $\sigma_b^2 = \text{Var}(L_{j,b})$. Since at a small contrast, the luminance levels between the object and the background are very close, approximately $\sigma_o \approx \sigma_b \approx \sigma_p$. Hence, the displayed signal-to-noise ratio SNR_D of the object with a small contrast will be

$$SNR_D = \frac{\Delta L}{\sigma} = C \cdot \sqrt{\frac{n}{2}} \cdot \frac{L_b}{\sigma_p} . \quad (4.6)$$

The similarity between (4.6) and (4.1) should be noted. The area A in (4.1) is replaced by the number of pixels n in (4.6). The major difference is that $\Phi^{1/2}$ is replaced by $SNR_p = L_b/\sigma_p$, defined as a single-pixel signal-to-noise ratio. Following the same procedure as used in the Rose model, we get

$$C_T = \frac{\sqrt{2} k_0}{\sqrt{n} SNR_p} . \quad (4.7)$$

Since

$$C_T = JND_L/L_b ,$$

it follows that

$$JND_L = \frac{\sqrt{2} k_0 L_b}{\sqrt{n} SNR_p} = \frac{\sqrt{2} k_0 \sigma_p}{\sqrt{n}} \quad (4.8)$$

Eqn.(4.8) relates JND_L to the RMS of the display device noise and indicates that JND_L is inversely proportional to the square-root of the object area (in terms of the number of pixels). In effect, we can treat $\sigma_p/n^{1/2}$ as the RMS noise of the object, σ_{obj} , because when the eye sees an object, it spatially integrates over all pixels within the object, the resultant noise will be decreased by a factor of $n^{1/2}$ if the noise is uncorrelated. Hence, for a given object, JND_L can be written as

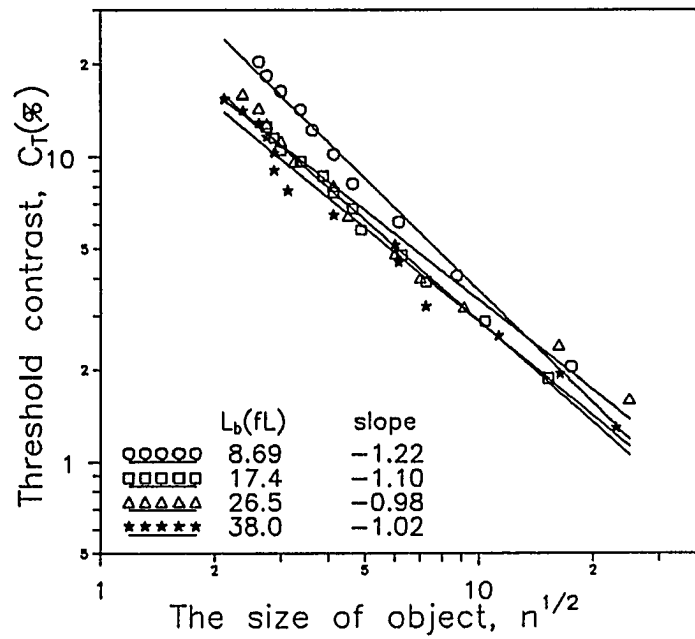
$$JND_L = \beta_0 \cdot \sigma_{obj} \quad (4.9)$$

where $\beta_0 = 2^{1/2} k_0$. SNR_p is usually a function of the background luminance L_b , and from (4.7) we notice that the dependence of C_T on L_b is identical to that of $1/SNR_p$.

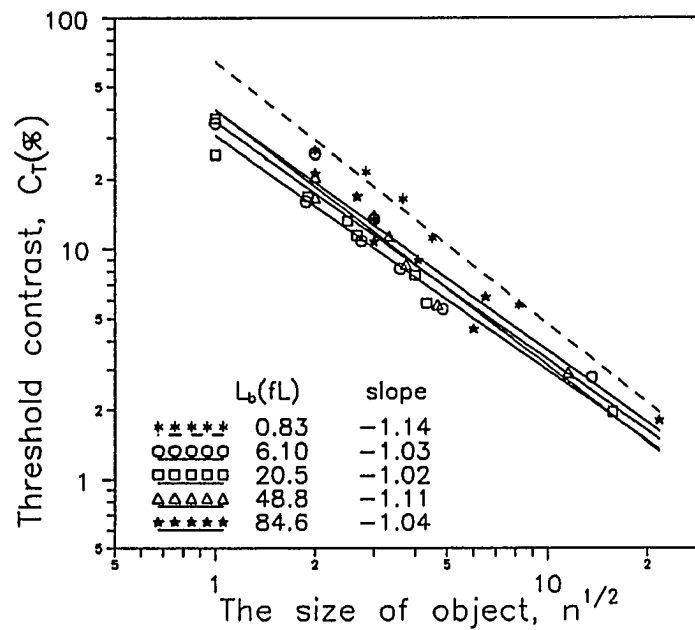
It should be remembered that there are certain limitations to the Rose model and its extension derived above. First, the size of the displayed object has to be limited within a certain range. If the object size is too small, the detection of the object will be limited by the eye's limiting spatial resolution set by the finite size of rods and cones, rather than by the displayed signal-to-noise ratio [53]. Secondly, the object area over which the eye can integrate spatially is limited by an angular field of about 0.5 degree [52]. Hence, a further increase of the angular size of object beyond the limit will not increase the received signal-to-noise ratio any more, and consequently the threshold contrast will no longer decrease with the increase of the object size.

Our physical measurements of CRT noise and psychophysical experiments showed that within a certain range of the object size, the measured threshold contrast follows well the value predicted by (4.7). The measured threshold contrast as a function of the linear size of object is shown in Fig. 4.2 for two different CRTs. Fig. 4.3 illustrates the measured threshold contrast as a function of the single pixel signal-to-noise ratio SNR_p . The straight lines in these figures are the best fitting curves. These double-log plots clearly show that the threshold contrast is inversely proportional to the square-root of the object area and the square-root of SNR_p (the slopes of the fitting curves in the double-log plots are all close to -1). The measured SNR_p of one CRT is plotted as a function of the background luminance L_b in Fig. 4.4 where we can see that the display device noise consists of the combined contribution from the spatial and temporal noise as mentioned in §2.6. For large background luminance values, spatial noise dominates the total SNR_p , and consequently results in a constant SNR_p . As the background luminance decreases, the SNR_p is increasingly determined by the temporal noise and is no longer independent of the background luminance. Fig. 4.5 shows how the measured threshold contrast varies with the background luminance for the same CRT. As we expected, it follows $1/SNR_p$. It is interesting to note that the threshold contrast curve in this case has almost the same behavior as the Weber-Fechner relationship. Hence, we can still use (3.1) to describe the C_T curve. However, it should be emphasized that the dominant noise sources are different.

These results verify that the Rose model, which was developed for the case where the human observer is limited by photon noise, can be modified to describe the human



(a) CRT 1



(b) CRT 2

Fig. 4.2 Threshold contrast as a function of the side length (in pixels) of a square object with background luminance as a parameter, measured for two CRTs.

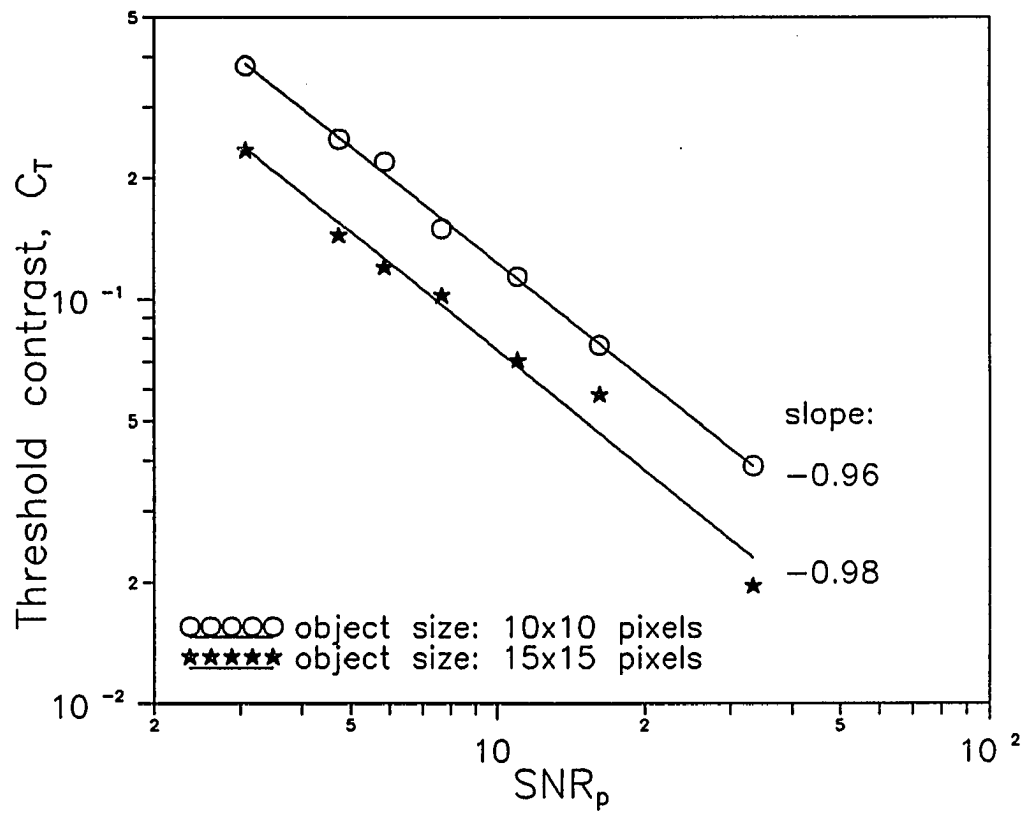


Fig. 4.3 Threshold contrast as a function SNR_p for two different object sizes.

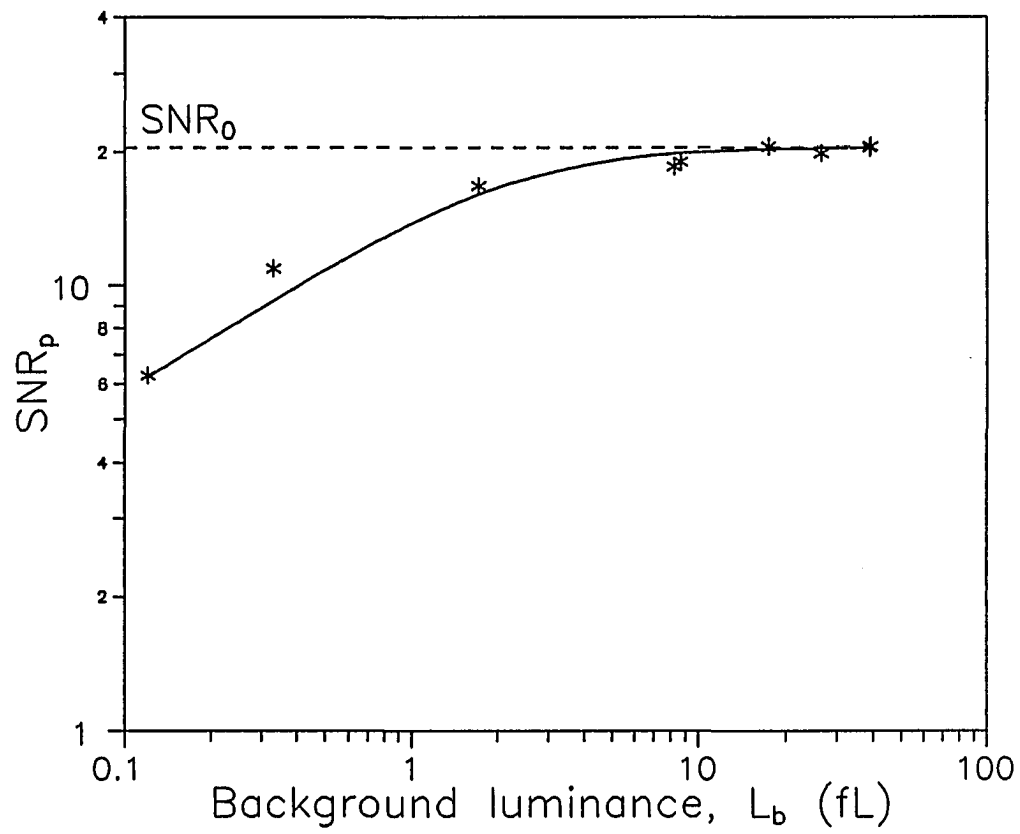


Fig. 4.4 *The measured SNR_p of a CRT as a function of the background luminance.*

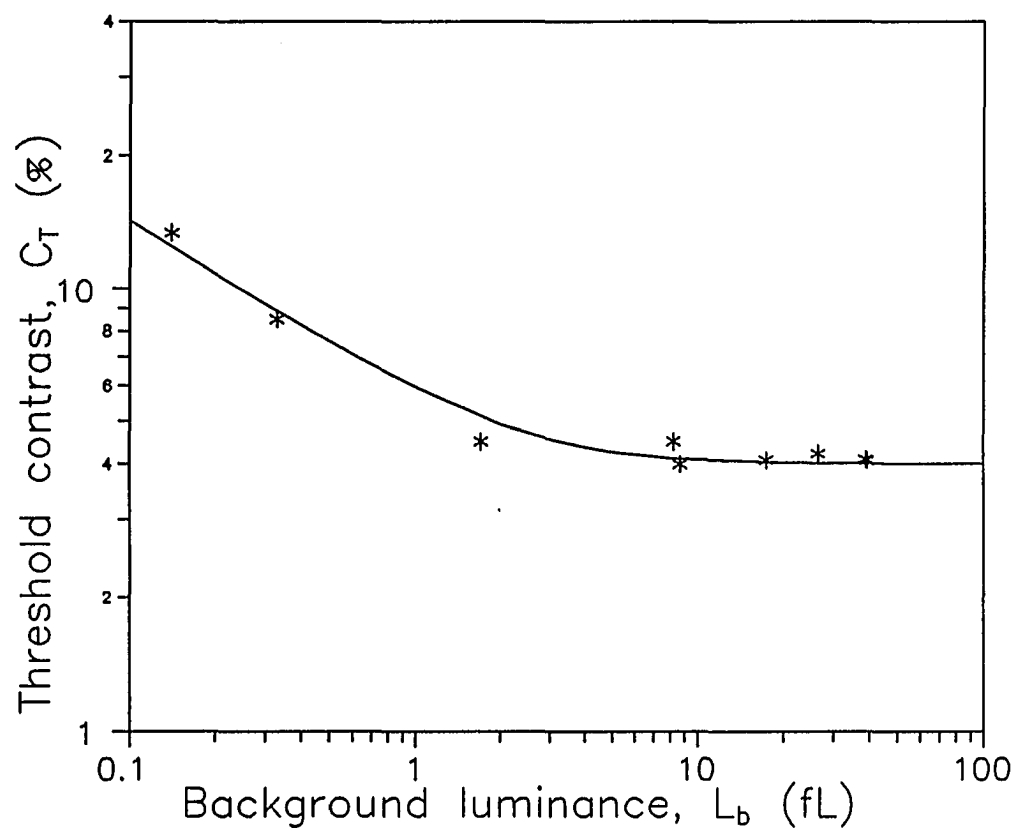


Fig. 4.5 The measured threshold contrast of the same CRT in Fig 4.4 plotted as a function of the background luminance for a square object of 15×15 pixels.

observer's perception performance in the presence of dominant display device noise. It can be seen from these results and the results presented in Chapter 3 that the threshold contrast determined by the SNR_p is much higher than the minimum threshold contrast limited only by the internal noise of the human visual system, indicating that the display device noise is the limiting noise of the entire system in most cases. With the measured SNR_p for a given display device, the threshold contrast function C_T (or equivalently JND_L) can be derived from (4.7), which is just a bridge to connect the physical and the psychophysical evaluations of a display device. In other words, by measuring the display device noise, one can predict the threshold contrast which would have been obtained with the very time-consuming psychophysical experiment. The visual response function and the perceived dynamic range of a display device can then be calculated as described in the previous chapter. Comparing (4.7) and (3.1), we can see that SNR_p , the only quantity that depends on background luminance L_b , must have the following form

$$SNR_p = \frac{SNR_0}{\left(1 + (L_0/L_b)^{pm}\right)^{1/m}}, \quad (4.10)$$

where SNR_0 represents the maximum value of SNR_p when it reaches its constant level. It follows that C_0 in (3.1) includes all other parameters that are independent of L_b , i.e.,

$$C_0 = \frac{\sqrt{2} k_0}{\sqrt{n} SNR_0}. \quad (4.11)$$

Recalling that the normalized visual response function $h_N(L)$ derived in §3.4 is independent of C_0 , we conclude that $h_N(L)$ will not be influenced by the size of the

displayed object and the maximum luminance signal-to-noise ratio per pixel.

4.3 Dependence of PDR on Physical Parameters

As introduced in Chapter 3, the perceived dynamic range is a psychophysical figure of merit described in terms of the number of JNDs achievable within the luminance range of a display device for a particular object. It is more meaningful in representing the contrast resolution of a display device, because it takes into account the effect of display device noise on discrimination of small luminance differences. If two display devices have the same physical dynamic range, the display device that has smaller noise will have a larger PDR, and consequently a better contrast resolution. Although the perceived dynamic range is a psychophysical parameter, it is directly related to physical parameters, such as physical dynamic range and luminance signal-to-noise ratio. According to (3.14), $PDR = S_{\max} = h(L_{\max})$. Hence, using (3.13) we have

$$PDR \approx \frac{1}{C_0} [\ln (L_{\max}/L_0) + \beta] \quad (4.12)$$

where $\beta = [1 - (L_{\min}/L_0)^p]/p$. Several observations can be made at this point. First, (4.11) and (4.12) indicate that PDR increases linearly with the SNR_0 of a display device. Secondly, PDR increases with the actual physical dynamic range at a much slower rate than linearly. In particular, if $p = 0$, i.e., the luminance range is completely within the Weber region (see Fig. 3.2), then $PDR = (1/C_0) \cdot \ln(L_{\max}/L_{\min})$. Hence, PDR is a logarithmic function of the physical dynamic range L_{\max}/L_{\min} . Since $\Delta D = \log(L_{\max}/L_{\min})$

for film, PDR is proportional to the density range of film in this case.

The other extreme case is that the luminance range is completely within the Rose-de Vries region. In this case,

$$PDR = \frac{(L_{\min}/L_0)^p}{C_0 p} \left[(L_{\max}/L_{\min})^p - 1 \right] . \quad (4.13)$$

Therefore, PDR is a power function of the luminance dynamic range with the power $p < 1$. If both L_{\min} and L_0 are fixed, PDR is also a logarithmic function of L_{\max} . In practice, the luminance range of the CRT usually covers both regions. The increase of luminance dynamic range may be achieved with increasing the maximum luminance L_{\max} and decreasing the minimum luminance L_{\min} . Also, the turning point L_0 remains nearly unchanged and its value is often observed to be close to the geometric average of the maximum and minimum luminance levels. i.e., $L_0 = (L_{\max}L_{\min})^{1/2}$. Then (4.12) can be reduced to

$$PDR \approx \frac{1}{C_0} \left[\frac{1}{2} \ln (L_{\max}/L_{\min}) + \frac{1}{p} \left(1 - (L_{\min}/L_{\max})^{p/2} \right) \right] , \quad (4.14)$$

which is plotted in Fig. 4.6 as a function of the physical dynamic range with various p 's.

Clearly, it is more effective to increase the single pixel signal-to-noise ratio of a display device than to increase either the physical dynamic range or the maximum luminance level of the display device for improving its PDR. It suggests that finer phosphor grain may be more favorable, among other considerations, for CRTs because it can increase SNR_0 by decreasing the spatial noise.

Another factor that greatly affects PDR of a CRT is the internal scatter (veiling

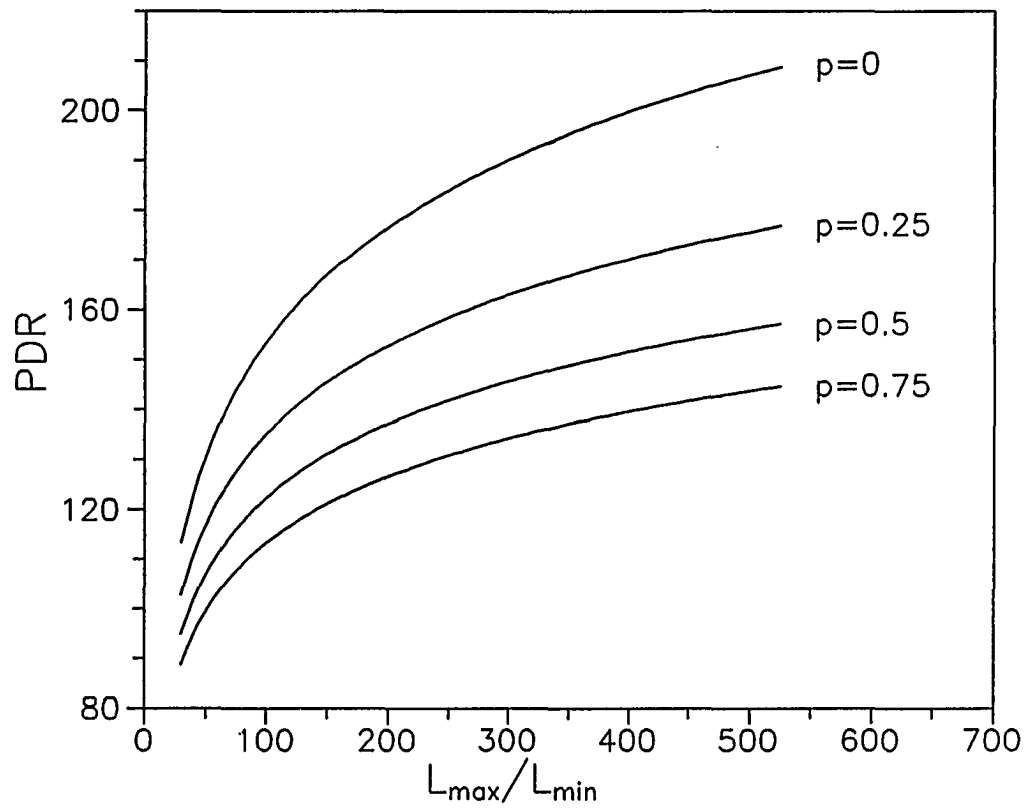


Fig. 4.6 The perceived dynamic range (PDR) as a function of the physical dynamic range. (Conditions are $L_0 = (L_{\max}L_{\min})^{1/2}$ and $C_0 = 0.03$. Note that $p = 0$ corresponds to the case where $PDR = (1/C_0) \cdot \ln(L_{\max}/L_{\min})$).

glare), because the internal scatter, if not well controlled, can significantly enhance the minimum luminance level L_{\min} , as we have shown in §2.4. It should also be noted that increasing the absolute luminance will not necessarily increase the PDR unless measures are taken to reduce the internal scatter. Since PDR also depends on the size of the object, in order to compare PDRs of two display devices one must compare them for the same object size.

It should be noted that the PDR given by (4.12) represents the maximum number of JNDs achievable for a particular object with an optimum display function (which will be discussed in the next chapter). With the original display function of a display device which is often inappropriate, the actually obtained number of JNDs for a particular object may very well be less than the PDR given by (4.12). In this sense, the perceived dynamic range is also dependent on the display function of a display device.

4.4 Summary

In this chapter, the relations between such psychophysical characteristics as the threshold contrast, JND and the perceived dynamic range, and many physical parameters of a display device are developed. The Rose model, which was originally proposed for predicting the visual performance in the photon noise limited case, is modified and applied to the determination of the threshold contrast in the case where the display device noise is the dominant noise. The experimental results have shown a good agreement between the threshold contrast values predicted by the modified model and the measured

data with some limitations to the range of the object size. It is found that the threshold contrast of a display device is inversely proportional to the luminance signal-to-noise ratio. The determination of this relationship makes the threshold contrast curve as a function of background luminance to be obtained relatively easily by noise measurements instead of a time-consuming psychophysical experiment. It is also found that the normalized visual response function is independent of the displayed object size and the maximum luminance signal-to-noise ratio, SNR_0 , and that increasing SNR_0 of a display device is a more effective method to improve its PDR than increasing the physical dynamic range or the absolute luminance.

CHAPTER 5

OPTIMIZATION OF THE DISPLAY FUNCTION OF IMAGE DISPLAYS

5.1 Introduction

The display device is usually the last, but by no means the least, important part of an imaging system. Much effort has already been invested in developing display devices for demanding applications such as digital radiography where people hope to replace the conventional film/view box combination by CRTs. The efforts of the display industry have mainly focussed on the matrix size, spatial resolution, and brightness, but not so much on the display's contrast capability. Although the spatial resolution of a display device does affect the sharpness of the displayed image, the display function has a dramatic effect on the appearance of the image. In Fig. 5.1, for example, the same image is displayed by three different display functions with the same dynamic range. As can be observed, some features which can be seen in the image in Fig. 5.1(c) cannot be seen at all in the image in Fig. 5.1(a) and Fig. 5.1(b) simply because of their inappropriate display functions. Obviously, how to optimize the display function of a display device is critical for observers to extract the maximum amount of information out of the displayed image. Since a display device presents an image to human observers, one should consider the display device and the human visual system as a whole system. Researchers concerned with the use of display systems have investigated how a display device can give the best perceptual performance by taking into account the human visual



(a)



(b)



(c)

Fig. 5.1 *The same digital image displayed with three different display functions.*

response. The perceptual linearization technique has been proposed by many authors to implement a specific display function for a display device such that equal changes in digital input level result in luminance changes that are perceptually equal throughout the entire luminance range of the display device [19]-[22]. Thus, with perceptual linearization, one grey level increment at the display's input will correspond to a linear change in perceived brightness.

Since displaying an image to an observer, whether by softcopy or hardcopy, can be treated as a contrast information transfer process, it makes sense to investigate the optimum display function from the viewpoint of information theory. In this chapter, we will show that perceptual linearization maximizes the contrast information transfer through the display/human observer system. Hence, utilizing the analysis of the human visual response function introduced in Chapter 3, we can derive the optimum display function.

5.2 Some Concepts from Information Theory

Before proceeding to derive the optimum display function for display devices we need to review some basic information theory [54]. Let us start with a few definitions:

Entropy of a Discrete Random Variable:

Let X be a random variable taking values in the set $\{x_1, x_2, \dots, x_n\}$ with probability $\{p_1, p_2, \dots, p_n\}$. Then, the entropy of X is defined as

$$H(X) = - \sum_{i=1}^n p_i \log p_i \quad . \quad (5.1)$$

The physical meaning of the entropy $H(X)$ is the uncertainty about the random variable X , or the information contained in the random variable X . The unit of entropy depends on the base of the logarithm. Usually the base is 2 and consequently the unit of entropy is bits.

Average Conditional Entropy of Two Discrete Random Variables:

Let X, Y be two random variables taking values $\{x_i, i=1, 2, \dots, N\}$ and $\{y_j, j=1, 2, \dots, M\}$, respectively, with a joint probability distribution $P(x_i, y_j)$. Then, the average conditional entropy of X given Y is defined as

$$H(X|Y) = - \sum_{i=1}^N \sum_{j=1}^M P(x_i, y_j) \log P(x_i|y_j) \quad . \quad (5.2)$$

An alternative form for this equation is

$$H(X|Y) = \sum_{j=1}^M H(X|y_j) P(y_j) \quad (5.3)$$

where $H(X|y_j)$ is defined by

$$H(X|y_j) = - \sum_{i=1}^N P(x_i|y_j) \log P(x_i|y_j) \quad . \quad (5.4)$$

Similar to $H(X)$, $H(X|y_j)$ is the uncertainty about X after y_j has been revealed. Therefore, it is clear from (5.3) that $H(X|Y)$ is the average uncertainty about X after Y

has been known.

Some useful properties of entropy are given here without proof:

- 1) The entropy of a discrete random variable is non-negative, i.e., $H(\cdot) \geq 0$.
- 2) $H(X|Y) \leq H(X)$, with equality if and only if X and Y are independent.
- 3) $H(X,Y) = H(Y) + H(X|Y) = H(X) + H(Y|X)$.

Average Mutual Information:

The average mutual information between X and Y is defined as

$$I(X,Y) = H(X) - H(X|Y) \quad (5.5)$$

It follows from the third property of entropy given above that

$$I(X,Y) = I(Y,X) \quad (5.6)$$

i.e., the information contained in X about Y is equal to the information contained in Y about X , which is why $I(X,Y)$ is called the average mutual information. It is obvious from the first and the second properties that $0 \leq I(X,Y) \leq H(X)$.

Discrete Memoryless Channel (DMC):

A DMC is a communication channel which has an input random variable X taking values in $\{x_1, x_2, \dots, x_N\}$, an output random variable Y taking values in $\{y_1, y_2, \dots, y_M\}$ and a set of transition probabilities $P(y_j|x_i)$ for $j=1, 2, \dots, M$ and $i=1, 2, \dots, N$, and has the property that the present output depends only on the present input.

For a DMC, the information transferred through the channel is given by (5.5)

where $H(X)$ and $H(X|Y)$ are considered as the original information and the information loss of the channel, respectively. Obviously, a perfect lossless channel must have $I(X,Y) = H(X)$.

5.3 Optimizing the Display Function Based on Human Visual Response

A display device transmits contrast information of a recorded digital image to human observers. Fig. 5.2 schematically shows the display device/human observer system. In Chapter 3 and Chapter 4 it was shown that the luminance noise, stemming from the phosphor granularity for CRTs or from film grains, is the major factor which limits the transmission of contrast information of an image. An optimum display function minimizes the information loss of the transmission process. It is well known that in order to maximize the transmission of information in a communication system, the characteristics of the transmitter and the receiver should be matched. Thus, to maximize the transmission of the displayed contrast information to the human observer, the characteristics of the display system must match those of the human visual system. The entire system in Fig. 5.2 can be treated as a discrete memoryless channel (DMC) in which the input is grey level GL and the output is the discrete level of the perceived brightness S corresponding to that grey level. Recall that the perceived brightness is defined by the visual response function as $S = h(L)$. As analyzed in §5.2, the information transferred through the above system can be expressed as

$$I(S, GL) = H(GL) - H(GL|S) . \quad (5.7)$$

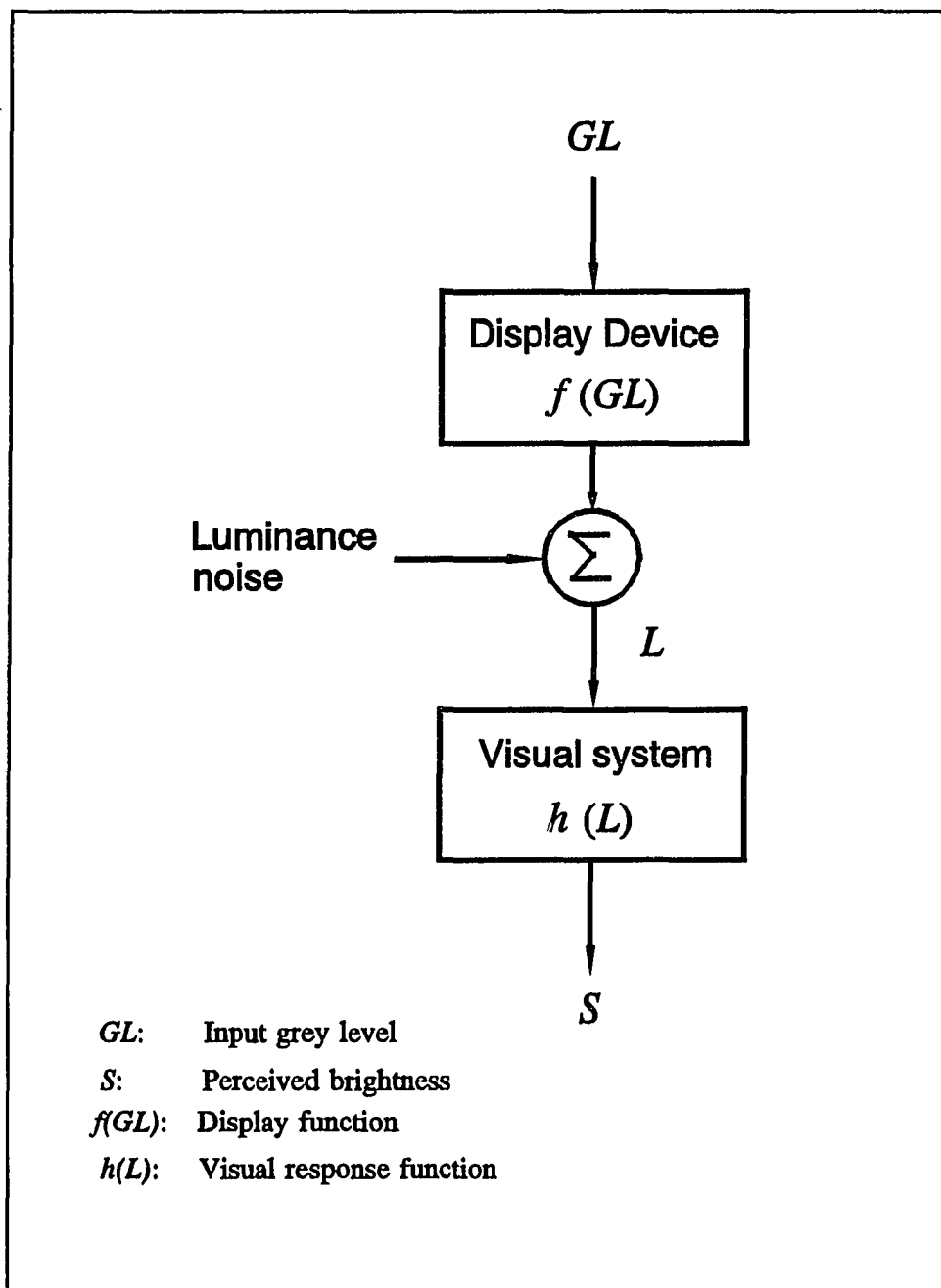


Fig. 5.2 *Display/Human observer system.*

Assume that GL ranges from 0 to GL_{\max} , and S ranges from 0 to S_{\max} . The entropy $H(GL)$ and the average conditional entropy $H(GL|S)$ in (5.7) are accordingly defined as

$$H(GL) = - \sum_i P(GL_i) \log_2 P(GL_i) \quad , \quad (5.8)$$

$$H(GL|S) = \sum_k H(GL|s_k) P(s_k) \quad (5.9)$$

where

$$H(GL|s_k) = - \sum_i P(GL_i|s_k) \log_2 P(GL_i|s_k) \quad , \quad (5.10)$$

$GL_i = i \cdot \Delta GL$, $i = 0, 1, \dots, M$. (Here we choose $M = GL_{\max}$, so $\Delta GL = 1$.)

Correspondingly, $s_k = k \cdot \Delta S$, $k=0, 1, \dots, M$, and $\Delta S = S_{\max}/M$. The units of these entropies are bits.

Due to the luminance noise of the CRT, for every input grey level a distribution of luminance values is produced. If the adjacent luminance distributions overlap significantly, the number of discernable grey levels is reduced. $H(GL)$ and $H(GL|S)$ can be interpreted as the original information and the information loss of the system, respectively. Since the entropy $H(\cdot)$ is non-negative, theoretically the best result we can get is $H(GL|S) = 0$, which requires that $H(GL|s_k) = 0 \quad \forall k$ because $P(s_k) > 0 \quad \forall k$. In reality, $H(GL|S)$ approaches zero at best. According to Bayes rule,

$$P(GL_i|s_k) = \frac{P(s_k|GL_i) P(GL_i)}{P(s_k)} = \frac{P(s_k|GL_i) P(GL_i)}{\sum_i P(s_k|GL_i) P(GL_i)} \quad . \quad (5.11)$$

If for any s_k ,

$$P(s_k | GL_i) = \begin{cases} 1 & \text{for } i=k \\ 0 & \text{for } i \neq k \end{cases} \quad (5.12)$$

then

$$P(GL_i | s_k) = \begin{cases} 1 & \text{for } i=k \\ 0 & \text{for } i \neq k \end{cases} . \quad (5.13)$$

It follows from (5.10) and (5.13) that

$$H(GL | s_k) = 0 \quad \forall k .$$

Hence, in order to minimize $H(GL | s_k) \quad \forall k$, we may attempt to satisfy (5.12).

Now let us examine $P(s_k | GL_i)$. $P(s_k | GL_i)$ is the conditional probability that given the input $GL_i = i \cdot \Delta GL$, the outcome $s_k = k \cdot \Delta S$ is obtained. The event $\{S = s_k\}$ is equivalent to the event that the fluctuation of the output luminance L due to the luminance noise falls into a certain range $\mathbf{R}_{L,k}$ whose center is at $L_k = h^{-1}(s_k)$, where $h^{-1}(\cdot)$ is the inverse function of $h(\cdot)$. Specifically,

$$\mathbf{R}_{L,k} = [L_k - \Delta L_k/2, L_k + \Delta L_k/2]$$

where ΔL_k corresponds to the ΔS centered at s_k . Hence,

$$P(s_k | GL_i) = P\{ L \in \mathbf{R}_{L,k} | GL_i \}$$

which is given by

$$P(s_k | GL_i) = \int_{L_k - \Delta L_k/2}^{L_k + \Delta L_k/2} p(L | GL_i) dL \quad (5.14)$$

where $p(L | GL_i)$ is the probability density function of L , given the input grey level GL_i .

According to the measurement of luminance noise, $p(L | GL_i)$ can be well characterized

as a Gaussian distribution density function with a mean of $f(GL_i)$ and an input-dependent standard deviation σ_i . Thus, (5.14) can be rewritten as

$$P(s_k | GL_i) = \Phi \left[\frac{h^{-1}(s_k) + \Delta L_k / 2 - f(GL_i)}{\sigma_i} \right] - \Phi \left[\frac{h^{-1}(s_k) - \Delta L_k / 2 - f(GL_i)}{\sigma_i} \right] \quad (5.15)$$

where

$$\Phi(\xi) = \int_{-\infty}^{\xi} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

It follows from (5.15) that condition (5.12) requires no overlap among $p(L | GL_i)$'s, and for any $i = k$, the distribution curve $p(L | GL_k)$ is completely centered in the range $\mathbf{R}_{L,k}$. To achieve this, we need

- 1) $\sigma_k \rightarrow 0$
- 2) $f(GL_k) = h^{-1}(s_k)$.

Unfortunately, the luminance noise cannot be arbitrarily small in practice. Therefore, there will always be some information loss. Since $k = GL_k / \Delta GL$, $s_k = k \cdot \Delta S = (\Delta S / \Delta GL) \cdot GL_k = (S_{\max} / GL_{\max}) \cdot GL_k$. Hence, the second condition above becomes

$$f(GL_k) = h^{-1} \left[\frac{S_{\max}}{GL_{\max}} GL_k \right] .$$

Thus, the optimum display function which maximizes $I(S, GL)$ is

$$f_{\text{opt}}(GL) = h^{-1}(b_1 \cdot GL) \quad (5.16)$$

with the scaling factor $b_1 = S_{\max} / GL_{\max}$.

Let us define $h_s(L) = (1/b_1) \cdot h(L)$, which is a scaled visual response function.

Then,

$$f_{\text{opt}}(GL) = h_s^{-1}(GL) . \quad (5.17)$$

Eqn.(5.17) says that the optimum display function should be the inverse of the scaled visual response function. Note that $h_s(L) = (1/b_1) \cdot h(L) = (GL_{\text{max}}/S_{\text{max}}) \cdot h(L) = GL_{\text{max}} \cdot h_N(L)$. In general, GL_{max} is known (e.g., for an 8-bit display system, $GL_{\text{max}}=255$). Hence $f_{\text{opt}}(GL)$ is completely determined by $h_N(L)$ which can be derived from $JND(L)$ or $C_T(L)$, as described in §3.4. The same result can be derived for film if one replaces luminance by density, and uses $h_s(D) = GL_{\text{max}} \cdot h_N(D)$ instead of $h_s(L)$.

Eqn.(5.17) is also consistent with the perceptual linearization of the display device because with the optimum display function we have

$$S = h[f_{\text{opt}}(GL)] = h[h^{-1}(b_1 \cdot GL)] = b_1 \cdot GL .$$

Hence, an incremental change in input signal, ΔGL , produces a linear change in perceived brightness, ΔS . The grey level JND, JND_{GL} , becomes a constant over the entire grey level range with this perceptual linearization. If the original display function is different from the optimum display function $f_{\text{opt}}(GL)$, we can easily design a non-linear mapping function $g(\cdot)$ such that $f[g(x)] = h_s^{-1}(x)$. Thus, $g(x) = f^{-1}[h_s^{-1}(x)]$.

As an example, the derived $f_{\text{opt}}(GL)$ for a specific CRT is shown in Fig. 5.3 along with the CRT's original display function. In this example, the luminance dynamic range of the CRT was 0.14 fL to 39.3 fL. The measured threshold contrast curve $C_T(L)$ for this monitor was already given in Fig. 4.5, and could be fit best by the parameters

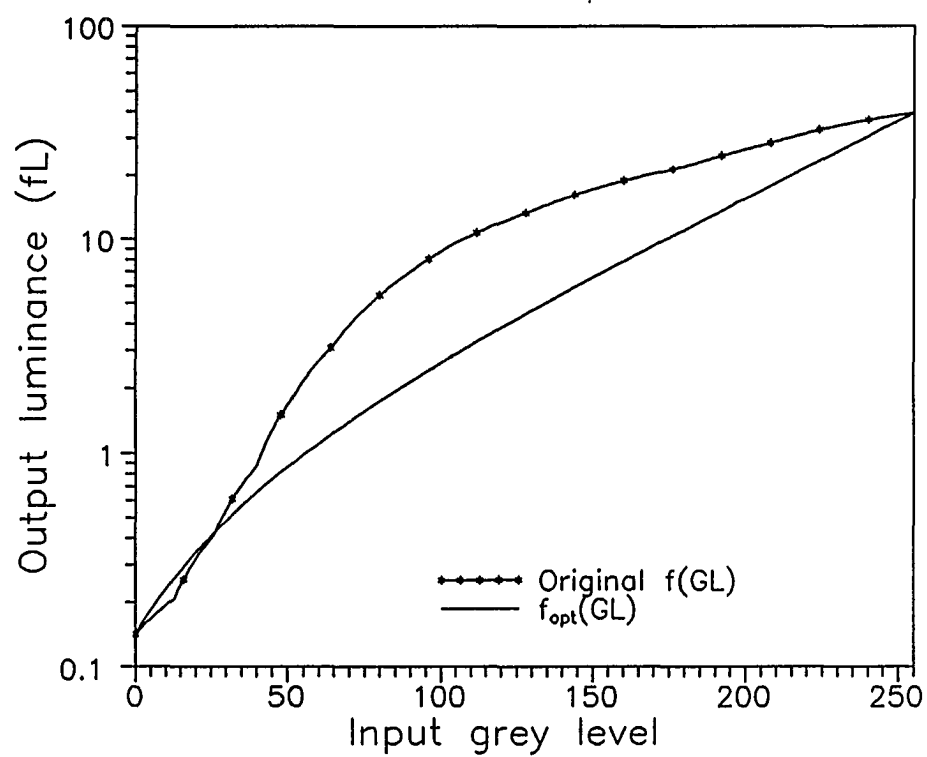


Fig. 5.3 *The original and the optimum display functions of a CRT.*

$C_0 = 0.04$, $L_0 = 2.36$ fL, $p = 0.4$, and $m = 4$. The *PDR* for a square object of 15×15 pixels was improved from 86 to 106 *JNDs* by implementing the optimum display function.

5.4 Experimental Results

Receiver Operating Characteristics (ROC) Analysis

In signal detection theory, the basic problem is to determine whether a signal is present or absent in a given noisy environment-[55]. For any selected decision threshold, there are four possible responses from the observer: (1) True positive (TP) that the observer correctly detects the signal when a signal is present, (2) False positive (FP) that the noise is mistaken for a signal, (3) True negative (TN) that the observer reports no signal when no signal exists and (4) False negative (FN) that the observer misses detection of a signal in noise. Accordingly there are four probabilities associated with each of these responses, termed as TPF (true positive fraction), FPF, TNF and FNF. These four probabilities are not independent as

$$\text{TPF} + \text{FNF} = 1$$

$$\text{FPF} + \text{TNF} = 1 \text{ .}$$

What we are interested in measuring is the ability to discriminate between signal and noise for a given system/observer modality. For instance, an interesting question is whether modality *A* can give a better discriminability between signal and noise than modality *B*. In general, a high TPF and a low FPF are desired, but unfortunately, they

are contradictory desires. With a loose decision threshold, one can obtain a high TPF, but this also results in a high FPF and vice versa. To compare two modalities, one needs to compare them at the same decision threshold which is very difficult, if not impossible, to determine. A way of overcoming this problem is to let the observer adopt his(her) decision threshold, which may not be quantitatively known, in order to obtain several pairs of true positive and false positive fractions. The curve of the TPFs versus FPFs can be plotted and is called the receiver operating characteristic (ROC) curve. A typical ROC curve is shown in Fig. 5.4. With ROC curves, one can compare TPFs for two modalities at the same FPF, or compare FPFs for two modalities at the same TPF, or more directly, compare the two ROC curves of the two modalities. It is easy to understand that better detection performance is indicated by an ROC curve that is close to the top left-hand corner. In fact the area under the ROC curve is a measure of the detection performance although it does not define the curve uniquely.

ROC analysis has been widely used in medical applications to evaluate the diagnostic performance for various modalities since the task of detecting known abnormalities can be regarded simply as a signal detection problem which can be well described by the ROC analysis [56]-[60].

Experimental Results

An ROC analysis has been conducted using real clinical images to test the effectiveness of the optimum display function. We selected sixty chest radiographs of which half were normal cases, and the other half were abnormal cases. All radiographs

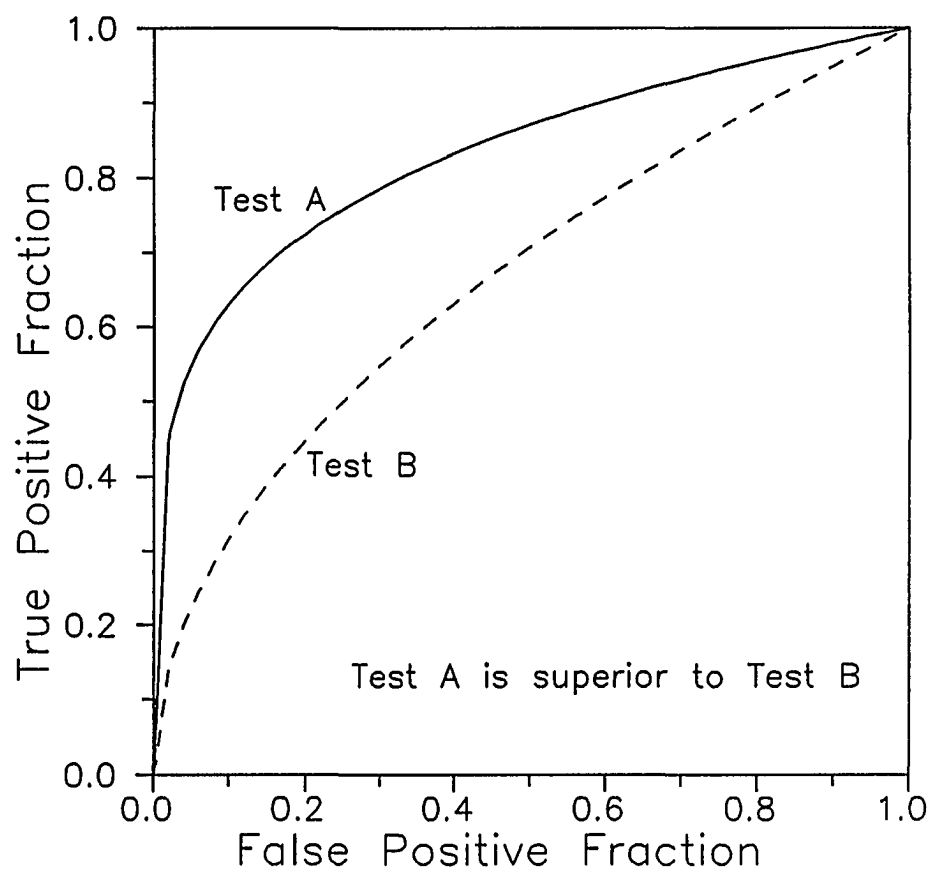


Fig. 5.4 Typical receiver operating characteristic (ROC) curves.

were digitized in 12 bits/pixel using a Lumisys DIS-1000 film digitizer. The digitized images were about $4k \times 4k$ pixels and had to be subsampled to fit the pixel matrix of the display monitor which is 1024×1536 US-Pixel CRT, controlled by an in-house-built 8-bit frame buffer/display controller. The digitized images were low-pass filtered before the subsampling to avoid possible aliasing. The abnormalities included examples of airspace disease, interstitial disease, pleural disease, small nodules or masses, hilar enlargement, and pneumothorax. Eight radiologists participated in the experiment. They were divided into two four-person groups. Each radiologist in the first group evaluated all 60 images, one by one in random order, displayed on the CRT with its original display function (see Fig. 5.5, the marked curve). In the same way, each radiologist of the second group evaluated the same set of images displayed on the CRT with a look-up table $g(\cdot)$ which corrected the original display function to the optimum display function (see Fig. 5.5, the solid line). The correction was an approximate one due to unavoidable round-off errors in digitizing the look-up table.

For each viewed image, the radiologists completed a separate report sheet consisting of two sections. On the first section, they recorded the diagnosis (normal or abnormal) and certainty (just guessing, relatively certain, absolutely certain). For abnormal diagnoses, the radiologists were instructed to list every abnormality they detected. The second section of the data sheet included a rating scale for subjective image quality (1 = "not adequate", 2 = "poor", 3 = "fair", 4 = "good", 5 = "excellent"). The observers rated the overall diagnosis quality of the spatial and contrast resolutions. No restrictions were placed on the viewing distance, so that observers could

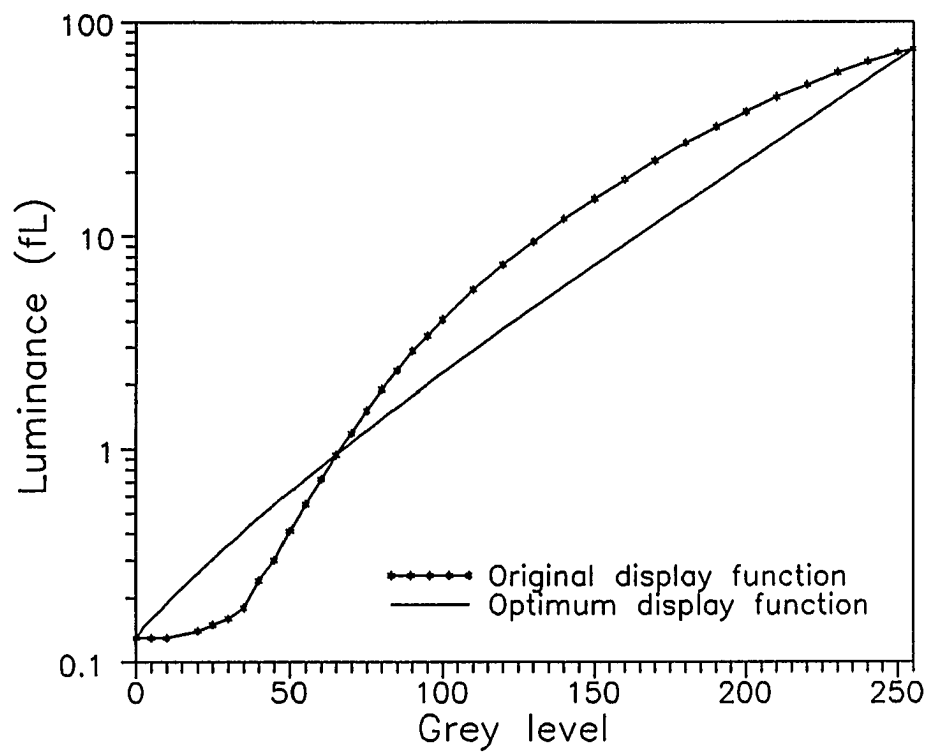


Fig. 5.5 *The original and the optimum display functions of the CRT used in the ROC study described in §5.4.*

adjust their viewing distance to obtain the most sensitive response to the spatial frequencies of the viewed objects. The observers were allowed to select their comfortable viewing positions and spend as much time as they wished for each image. The ambient light was set similar to the ambient light in the ordinary reading room.

The results are presented by receiver operating characteristics curves (ROC) in Fig. 5.6 and average subjective ratings of spatial and contrast resolutions in Table 5.1. The area under the ROC curves for the original display function is $A_z = 0.839$ with a standard deviation of 0.026 and, for the optimum display function, $A_z = 0.878$ with a standard deviation of 0.025.

TABLE 5.1 Mean adequacy ratings of spatial and contrast resolutions

	Spatial Resolution	Contrast Resolution
Original display function	2.39	2.94
Optimum display function	3.26	3.04

5.5 Discussion

From the above analysis, it can be inferred that the optimum display function $f_{\text{opt}}(\cdot)$ is completely determined by the normalized visual response function, $h_N(\cdot)$. In Chapter 3, it has been shown that $h_N(\cdot)$ is independent of C_0 . It has also been shown

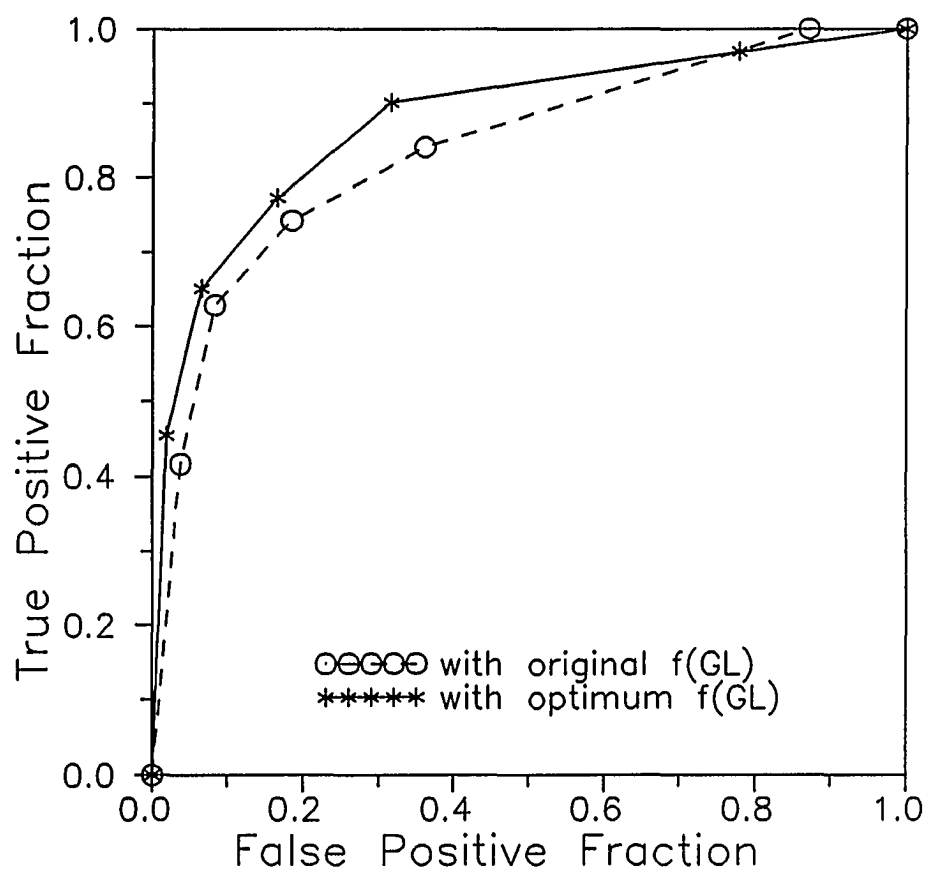


Fig. 5.6 ROC Curves for diagnosing thorax radiographs when presented on a CRT operated with either the original or the optimum display function given in Fig. 5.5.

in Chapter 4 that a variation of the object size mainly causes the threshold contrast curve to move vertically, but does not significantly change its *shape*, except for very large or very small objects near the perception limits of the human observer. Specifically, increasing the object size will mainly lead to a smaller C_0 , and vice versa. Hence, the normalized visual response function $h_N(L)$ is approximately independent of the object size and so is $f_{\text{opt}}(\cdot)$. The same conclusion can be drawn for a change in the noise level, because C_0 , the smallest threshold contrast, is linearly proportional to the maximum luminance signal-to-noise ratio per pixel of the display device (decreasing the noise of the display device will linearly reduce C_0). Therefore, the optimum display function of a display device is also independent of the noise level of the display device. It also suggests that to determine the optimum display function, we do not need to know C_0 accurately (we can even arbitrarily let $C_0 = 1$) and only the shape of the threshold contrast curve is important. The shape of the $C_T(L)$ curve is controlled by the parameters m , p and L_0 . With the typical threshold contrast curve modelled by (3.1), it follows that the visual response function is, more or less, proportional to the logarithm of the luminance at high luminance levels ($L > L_0$), and to $(L)^p$ at low luminance level ($L < L_0$) (see (3.13)). Hence, the optimum display function $f_{\text{opt}}(GL)$ is a power function of GL ($\propto (GL)^{1/p}$) at small GL s and smoothly transits to an exponential function at large GL s.

Recall that an important cause of degradation of the contrast capability of display devices, especially CRTs, is internal scatter, by which the luminance of one area is increased by the luminance of the surrounding areas. Internal scatter not only

significantly reduces the luminance dynamic range L_{\max}/L_{\min} (usually by about an order of magnitude) and consequently reduces the PDR, but also changes the actual display function, as illustrated in Fig. 2.3.

Since the internal scatter is a function of the average luminance of the displayed image, the actual display function is also image-dependent. In order to obtain the optimum display function for every image, ideally the mapping function $g(\cdot)$ should be adaptable. In practice, however, the average luminance of a class of images, e.g. chest images, does not vary too much. One mapping function may be adequate to nearly optimize the display device for a class of images. Several such mapping functions may be pre-stored and applied to the specific tasks, respectively.

Even though the optimum display function guarantees the best utilization of the information transfer capabilities of a given display system, it is not necessarily the best display function for a specific display application. The best display function for an application depends on the distribution of the needed information which may spread over only part of the input data range for one specific image. Hence in many practical display tasks, parts of the input data range need to be assigned over a relatively wide portion of the available display range so that many *JNDs* can be perceived while other parts of the input data range are relatively compressed. In addition, the distribution of the needed information varies from image to image. Hence, it is generally impossible to have one *best* display function for the display of all images or even all parts of one image. An implication in the derivation of the optimum display function is that the needed information is distributed equally well over the entire display dynamic range. However,

the superiority of using the optimum display function over the original display function of an image display is that the optimum display function makes full use of the greatest potential of the contrast display capability of the display device. For a given luminance dynamic range, the optimum display function facilitates detection of the maximum number of contrast differences (JNDs), which can be easily demonstrated by displaying, for example, a grey-scale checker board pattern of 100 equally spaced grey levels with the optimum display function and with the original display function. It was also demonstrated by the previous example given in §5.3 where 15×15 pixel square objects were displayed on the given CRT monitor and an improvement from 86 to 106 JNDs was achieved. Because the distribution of the needed information varies from image to image, adaptive image processing methods may be required to account for the specialty of each image. Furthermore, the processed image should also be displayed on a display device with its optimum display function. This subject will be discussed in more detail in the next chapter.

The results of the ROC experiment with the specific task and the given CRT monitor appear to be a little disappointing. Although an improvement in observer performance (given by the increased area under the ROC curve) was obtained as indicated in Fig. 5.6, it was not statistically significant (as evaluated by paired-difference t-test [61]). This is not too surprising, however, considering that several factors may have affected the experiment results: a) The original display function of the CRT monitor was not very different from the optimum display function (see Fig. 5.5), b) The coarse digitization (only 8 bits per pixel) limited by the hardware of the CRT (i.e., D/A

converter) kept us from accurately implementing the optimum display function, and c) For the given detection task, clinically relevant information in a chest radiograph was very likely not distributed uniformly over the entire input data range. Due to these factors, neither the optimum nor the original display function of the monitor were probably the best display functions under the given conditions. For most of the chest images presented in the experiment, the input data range of the image parts which the radiologists were interested in was above about grey level 40. Consequently, the overall contrast of images displayed by the original and the optimum display functions was almost the same, which can be seen from Fig. 5.5. This may explain why the subjective ratings of contrast resolution did not improve when displaying the thorax radiographs with the optimum display function as compared with the original display function because the radiologists somehow took the impression of overall image contrast as the contrast resolution. On the other hand, it is interesting to note that the subjective ratings of the spatial resolution was increased, as shown in Table 5.1, even though the MTF of the CRT did not change. It suggests that the overall performance of perception did indeed improve with the optimum display function. It is common that high-spatial-frequency details are typically weak in chest radiographs due to intrinsic small density modulation of the anatomy as well as the rather poor high-frequency detective quantum efficiency (DQE) of the image acquisition system. The improvement in contrast for high-spatial-frequency details with the optimum display function brought up the easier perception of these details compared to the original display function, which gave an impression of improved spatial resolution.

5.6 Summary

It has been shown in this chapter that the optimum display function of a display device should match the human visual response function in order to maximize the transferred contrast information of an image. Specifically,

$$f_{\text{opt}}(GL) = h_s^{-1}(GL) .$$

Given a display device, $f_{\text{opt}}(GL)$ provides the maximum perceived dynamic range available to human observers. This result is equivalent to perceptual linearization of image displays. For a display device, its optimum display function can be directly derived from the JND curve or the threshold contrast curve, $C_T(L)$, and is approximately independent of the size of objects in the image, i.e., independent of the spatial frequency. $f_{\text{opt}}(GL)$ is also independent of the noise level of the display device. A non-linear mapping, which can be derived based on $f_{\text{opt}}(GL)$ and $f(GL)$, is needed to correct the original display function. For this purpose, the effect of internal scatter on $f(GL)$ must be considered.

By default, a display device should always present images by this standard, generalized optimum display function. The optimum display function is the best display function for a given application when the useful contrast information is uniformly spread over the input data range. As long as the optimum display function has been implemented, it should be kept all the time (hence, the contrast/brightness knobs for a CRT should not be used for arbitrary adjustments). For the best visualization of arbitrary images, adaptive methods (e.g. adaptive contrast enhancement techniques) may

be required to process images before displaying them.

CHAPTER 6

ADAPTIVE IMAGE CONTRAST ENHANCEMENT BASED ON HUMAN VISUAL PROPERTIES

6.1 Introduction

The purpose of image contrast enhancement is to improve the appearance of the image, allowing observers to extract from the image as much information as possible. This is particularly important in some medical applications, for example chest radiography where radiologists view the patient's chest radiographs to make a primary diagnosis. Abnormal structures on radiographs are often presented by small luminance differences caused by small differences in the x-ray path attenuation and in object thickness. In addition, current softcopy display systems are incapable of displaying as many different discernible levels of luminance as can be recorded in a digital image. These factors are very likely to cause some subtle, but probably crucial, contrast information loss when a digital image is displayed on a CRT. The diagnostic accuracy can be reduced due to the poor contrast in the displayed image. Contrast enhancement is certainly one way to compensate for the decrease of contrast in the displayed image.

Numerous contrast enhancement methods exist, ranging from the simplest linear contrast stretch (also called windowing-and-leveling) to some very sophisticated adaptive enhancement techniques [25], [26]. The following brief review attempts to outline a few representative techniques and to summarize some existing common problems which need

to be overcome.

Linear Contrast Stretch

In linear contrast stretch, a sub-range of the input grey levels (a window) is selected and mapped linearly into the entire range of display driving intensities (output grey levels), as illustrated in Fig. 6.1. Since the output range is greater than the input range, the global contrast of the image is enhanced. The input-to-output relationship can be mathematically given by

$$I_o(x,y) = \begin{cases} 0 & I(x,y) \leq GL_1 \\ \frac{M}{GL_2 - GL_1}(I(x,y) - GL_1) & GL_1 < I(x,y) < GL_2 \\ M & I(x,y) \geq GL_2 \end{cases} \quad (6.1)$$

where $I(x,y)$ and $I_o(x,y)$ are the input and the output grey levels of the pixel at location (x,y) , respectively, $[GL_1, GL_2]$ is the input grey level range and $[0, M]$ is the output grey level range. This method is so simple that users can interactively choose the wanted input grey level range to stretch the image contrast in real time. The disadvantages of this global technique are: (i) only a partial grey level range of the image is enhanced, while grey levels outside the selected range are mapped either to the minimum or the maximum output grey levels, which means a loss of image information, and (ii) the requirement for manual interaction with the image adds more tasks to the user to search for a possible optimum grey level window.

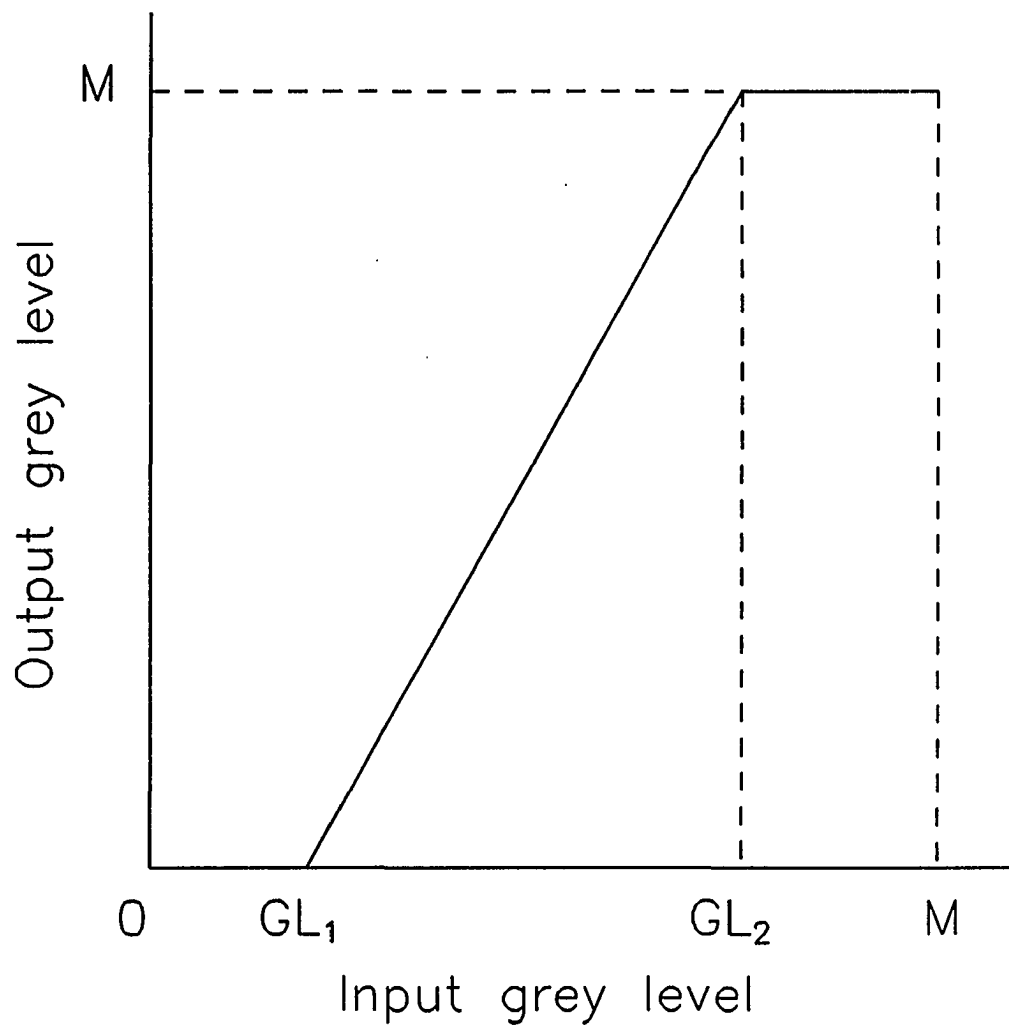


Fig. 6.1 *Illustration of linear contrast stretch.*

Histogram Equalization

Histogram equalization is a contrast enhancement technique which attempts to assign the output grey levels based on the statistical distributions of the input grey levels of the image. The input-to-output mapping is determined from the cumulative distribution function (CDF) of the image as

$$I_o(x, y) = K_1 \cdot CDF[I(x, y)] + K_2 \quad (6.2)$$

where K_1 and K_2 are two constants that scale the output image to the desired output grey level range. If this range is the total grey level range of the display device, $[0, GL_{\max}]$, which is a common case, then

$$K_1 = GL_{\max}/N_p \quad \text{and} \quad K_2 = 0$$

where N_p is the total number of pixels in the image. Since the histogram of the image is the first derivative of its CDF and the contrast gain is just the first derivative of the input-to-output mapping function, the output grey level range is proportional to the pixel intensity distribution of the image. The result is that those grey levels where there are large pixel populations are expanded while other grey level ranges where there are fewer pixels are compressed. This technique is a global method and its biggest advantage is that it automatically manipulates the input image without the need for any input parameter, as it utilizes only the statistics of the input image. The disadvantage of histogram equalization is that it tends to over-enhance the image contrast if there are high narrow peaks in the histogram of the image. Because these narrow peaks in the histogram correspond to large smooth regions of the image, the over-enhancement in

these regions often results in a harsh appearance of the output image and an unacceptably high noise.

Spatial Filtering Methods

Since human vision is more sensitive to the intensity change of the mid to high spatial frequency components of the image than to that of low spatial frequency components, enhancing the mid to high spatial frequency components relative to low spatial frequency components will have the effect of contrast enhancement, which is known as high spatial frequency emphasizing. In spatial domain, the input-to-output mapping is expressed as

$$I_O(x,y) = \sum_{k=1}^n G_k I_k(x,y) \quad (6.3)$$

where $I_k(x,y)$'s are components of different spatial frequencies of the original image and G_k 's are gain coefficients for these component images. This technique is also called multichannel filtering since the whole spatial frequency range of the input image is divided into several spatial frequency bands. Unsharp masking [62], for instance, is a special case of multichannel filtering with $n = 2$ in the above formula. Specifically, the output image resulting from unsharp masking is given by

$$I_O(x,y) = G_B I_B(x,y) + G_S (I(x,y) - I_B(x,y)) \quad (6.4)$$

where $I_B(x,y)$ is a heavily smoothed version of the original image (low-pass filtered image), which is subtracted from the original image in order to obtain a structured

image, denoted by $I_S(x,y)$, that contains the mid to high frequency components of the original image, and G_B and G_S are gain factors for the low and high frequency components, respectively. For contrast enhancement, it is required that $G_B \leq 1$ and $G_S > 1$.

One example of multichannel filtering was given by Tahoces *et al.* [63] who designed a three-channel filtering algorithm to enhance chest and breast radiographs. They also used a post-processing (a non-linear transformation) to re-scale the processed image so that the grey level range was within the range of the display device. The spatial filtering does effectively enhance the image contrast. But the selection of coefficients G_k 's is a difficult task. They are obviously image content dependent. Unless G_k 's change adaptively with the local image characteristics, it is unlikely that a fixed set of G_k 's can work well for all the images. Another problem with this technique is the occurrence of ringing artifacts around sharp edges.

All the techniques discussed above are global methods, i.e., one transformation is applied to all the pixels of the image. This type of methods may work well for some images. However, in the real world there are more complex situations where an image may have enough global contrast with considerable low-contrast local details, or the contrast in some parts of the image is poor while the contrast in other parts of the image is adequate. A better way to deal with this problem is the adaptive contrast enhancement (ACE). Generally, ACE techniques involve a mapping of the form

$$I_O(x,y) = f[I(x,y), O_\Omega(x,y)] \quad (6.5)$$

where $O_\Omega(x,y)$ represents some local characteristics within the neighborhood Ω around

the location (x,y) . Hence, the mapping f changes adaptively with the local characteristics of the image. Let us briefly outline a few typical adaptive techniques.

Local Range Modification (LRM)

Fahnestock and Schowengerdt [64] have developed a very efficient ACE method called Local Range Modification (LRM) which is a linear contrast stretch with variable coefficients. The mapping in this case is given by

$$I_o(x,y) = \frac{C_1}{\max_{\Omega}(x,y) - \min_{\Omega}(x,y)} (I(x,y) - \min_{\Omega}(x,y)) + C_2 \quad , \quad (6.6)$$

where $\min_{\Omega}(x,y)$ and $\max_{\Omega}(x,y)$ are the local minimum and the maximum in the neighborhood Ω around (x,y) . $C_1 + C_2$ is the desired maximum output grey level and C_1 is just the output range. Instead of calculating the local minimum and the maximum for each pixel, the LRM method first calculates these values for contiguous blocks of the image and associated them with the central pixel of each block, and then applies a two-dimensional interpolation to estimate $\min_{\Omega}(x,y)$ and $\max_{\Omega}(x,y)$ for any other pixels of the image. This technique substantially reduces the computation burden and makes the algorithm highly efficient. Furthermore, this technique can certainly be extended for the calculation of other statistics of the image as we will see later. The parameter space of LRM is small, because in most cases the block size is the only parameter to be determined, whereas C_1 is usually set equal to the full display range and C_2 is set to zero. As the authors have indicated [64], since LRM is based on the maximum and the minimum values, it is very sensitive to spike noise. Besides, ringing artifacts are also

evident on either side of the transitions.

Adaptive Filtering

Peli and Lim [65] have proposed an adaptive filtering algorithm based on the structure of unsharp masking. The algorithm first separates an image into low-pass and high-pass components. The high-pass component is then enhanced with a gain which depends on the local mean, and the low-pass component is also transformed by a non-linear function which is image-dependent. These two processed components are then combined to give the output:

$$I_O(x,y) = G_B(I_B) \cdot (I(x,y) - I_B(x,y)) + T(I_B) \quad (6.7)$$

where $G_B(\cdot)$ and $T(\cdot)$ are both functions of the local mean $I_B(x,y)$. The particular characteristic of the image used for adaptive filtering is the local mean. Peli and Lim provided three different image samples processed successfully by this algorithm with different non-linear functions $G_B(\cdot)$ and $T(\cdot)$. However, the success of their method requires customized tailoring of the non-linear functions for each image to be enhanced. The best choice of the non-linear function has to be tried out for specific applications.

Adaptive Histogram Equalization (AHE)

An extension of the use of histogram equalization in local areas of the image was developed first by Ketcham *et al.* [66], and then continued by others [67]-[69], to improve the local contrast of the image. A sliding window centered at each pixel is used

in which the local histogram is calculated and equalized, thus altering the value of that pixel. For an image of size $N \times N$ pixels, N^2 local histograms need to be calculated. This time-consuming procedure was improved by Pizer with a method of reducing the computation burden for calculating local histograms [68]. His method is similar to that of LRM. The image is divided into blocks and the histogram equalization mapping is calculated only for each block and assigned to its central pixel. Then the mapping function for any other pixel is bilinearly interpolated from the mapping functions of the pixel's four surrounding blocks. The only parameter to be determined in this method is the block size. Some empirical results indicated that 8×8 blocks were adequate for most images, and that a 50% overlap between blocks could smooth the transition of mapping between blocks. It has been shown by experiments that there is virtually no difference between the results of applying true AHE and its interpolated version for the same effective block size. It has also been reported that this efficient AHE worked well for many medical images, including chest radiographs and CT images [70]. However, the problem of noise over-enhancement associated with the histogram equalization procedure is still not solved. To attack this problem, Pizer *et al.* [71] introduced the following modification to the AHE, called CLAHE.

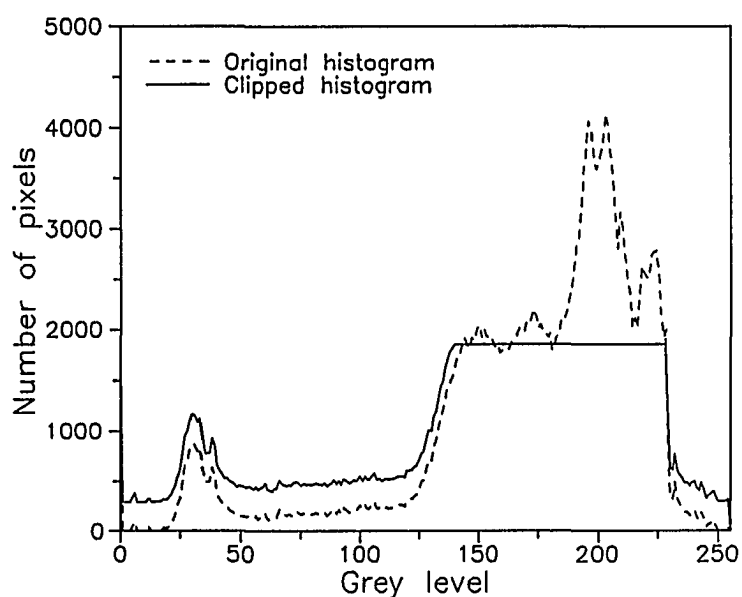
Contrast-Limited Adaptive Histogram Equalization (CLAHE) and Its Extension

This method is based on the interpolated version of AHE with the modification of limiting the contrast gain by restricting the height of local histograms. Recall that in histogram equalization the contrast gain is given by the first derivative of the CDF, or

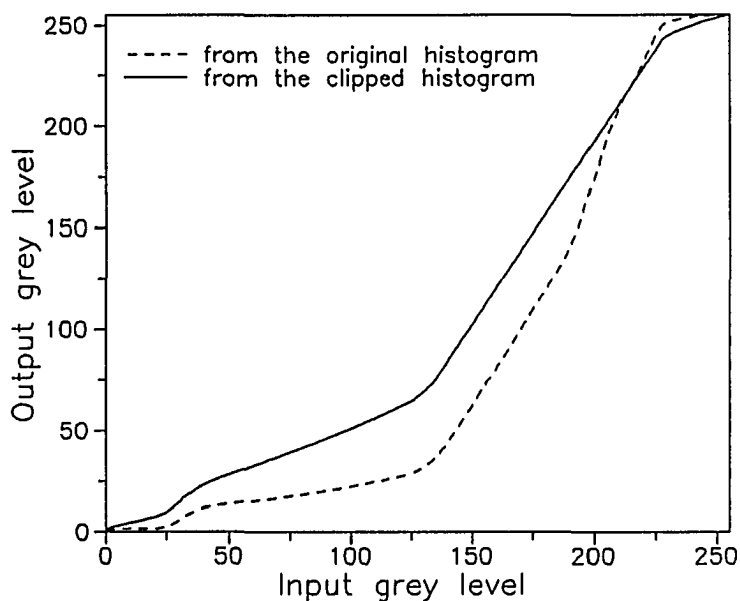
the histogram. If the local histogram is too high in some area of the image, the corresponding contrast gain may exceed what is needed and cause over-enhancement. Hence, one can limit the contrast gain by clipping the height of the local histogram. Specifically, if the desired contrast gain is G_C , the maximum height of the histogram has to be restricted by

$$H_{\max} = \frac{N_p}{GL_{\max}} G_C$$

where N_p is the number of pixels in the region where the local histogram is calculated, and GL_{\max} , as before, is the display range. Since the number of pixels in the region should remain unchanged, the clipped histogram has to be re-distributed and re-normalized so that the area under the clipped histogram curve is still equal to the area under the original histogram curve. For simplicity, a uniform redistribution is used to reallocate the clipped pixels to all grey level bins. Note that this step will push the clipped histogram up again above the height limit H_{\max} , and hence it needs to be clipped again. A simple iterative algorithm has been designed to implement this re-distribution and re-normalization. After this procedure, the mapping function is derived from the clipped histogram in the same way as ordinary histogram equalization. Fig. 6.2 illustrates, as an example, the difference between the ordinary histogram equalization and the clipped histogram equalization. As we can see, the contrast gain (i.e. the slope of the mapping curve) is limited in the grey level range where the histogram is clipped and the rest of grey level range is uniformly expanded relative to ordinary histogram equalization.



(a)



(b)

Fig. 6.2 (a) The histogram of a chest image and its clipped version. Note that the area under the histogram curves remains unchanged as the total number of pixels in the image is a constant, (b) The input-to-output transformation (look-up table) derived from the original histogram and its clipped version given in (a).

An extension of CLAHE is Artifact-Suppressed Adaptive Histogram Equalization (ASAHE) [72]. It was developed by Rehm who divided the image into structure and background, applied CLAHE to the structure image only, and then recombined the enhanced structure image with the background.

CLAHE and ASAHE proved to be superior to the original AHE for their improved noise performance although they do not completely eliminate noise enhancement in smooth regions. The selection of contrast gain limit G_C is also an image-dependent task.

In summary, the image contrast enhancement methods can be classified into two categories: global methods and adaptive methods. Generally speaking, the adaptive techniques provide a better performance than the global ones at the expense of more complex algorithms and more computations. From an examination of these existing techniques one can find some common problems. First, there is not a guideline that tells us how much enhancement is adequate at each location of the image. It is likely for a processed image to have one region where the local contrast is already too much, but have another region where the local contrast is still not enough to allow observers to see the details. Secondly, the over-enhancement of noise is distracting, especially in relatively uniform regions of the image. The second problem is partly associated with the first problem because of not knowing where the contrast needs to be enhanced and the amount of enhancement. Finally, the ringing artifacts commonly appear around the sharp transitions within the image. Based on this analysis, we propose the following desired goals for the new ACE method that will be reported in this chapter: 1) A

quantitative determination of both the global and the local contrast gains. 2) A good control of noise enhancement. 3) Suppression of ringing artifacts. 4) Fast implementation.

6.2 JND-Guided Adaptive Contrast Enhancement (JGACE)

It is clear from the previous discussion that for any ACE method, some local characteristics of the image have to be utilized to determine the varying local contrast gain. Since the ultimate purpose of ACE is to make the information contained in an image more visible, some characteristics of the human vision system must also be considered. The new method proposed in this section uses the basic structure of unsharp masking as mentioned above, based on the fact that most of the information in an image is conveyed by its mid to high frequency components although they contain only a very small fraction of the total energy of the image, and that the human visual system is very sensitive to these frequencies. The human vision also has a property of being more sensitive to random noise in smooth areas than in "busy" areas where there are more details. Netravali and Prasada [73] have shown that the visibility of noise decreases monotonically with increase in the spatial activity, which they defined as the rate of spatial change in image intensity from one pixel to another. Hence, any over-enhancement of noise is mainly due to the unnecessary enhancement in low spatial activity regions where there is little useful information. To overcome this undesirable feature, the key is to classify low spatial activity regions (or smooth regions) and high

spatial activity regions (or detail regions) and treat them differently. One should keep the local contrast unchanged in smooth regions in order to avoid noise enhancement, and only enhance local contrast in detail regions.

The next question is then how to determine the adaptive contrast gain G_S in detail regions. It is reasonable to determine G_S based on the actual local contrast C_A and the human visual response to the contrast, i.e., the contrast sensitivity function discussed in §3.2 (see Fig. 3.1). We assume that the required contrast in detail regions should be several times higher than the threshold contrast C_T (i.e. $G_S \cdot C_A \geq \gamma \cdot C_T$ with $\gamma > 1$) in order to ensure a reliable and comfortable detection or recognition of low contrast objects in an image. If the local contrast is less than $\gamma \cdot C_T$, the local contrast should be increased to that value; if the local contrast is already larger than or equal to $\gamma \cdot C_T$, no contrast enhancement is needed. This treatment can also eliminate unnecessary contrast enhancement around sharp edges because the local contrast there is already high enough. Consequently, the ringing artifacts can be hopefully avoided. Recall that C_T is a function of the spatial frequency and the background luminance, which would require calculating both local mean grey level and average grey level change, and converting them to luminance values through the non-linear display function of the display device. This complication can be simplified if we perceptually linearize the display device as discussed in Chapter 5, since with a perceptually linearized display device, the JND in grey level, JND_{GL} , is independent of the grey level of the background and is only a function of the local spatial frequency. Furthermore, it can be shown easily that JND_{GL} follows the same spatial frequency dependence as that of $C_T(u)$, where u is the angular spatial

frequency in units of cycles/degree. Therefore, the above stated requirement for the local contrast ($G_S \cdot C_A \geq \gamma \cdot C_T$) can be replaced by adjusting G_S such that the average grey level change is larger than $\gamma \cdot JND_{GL}$ in each detail region. Thus, the computation of local contrast is replaced by the calculation of the average grey level deviation from the local mean. Hence, the required local characteristics of the image are the local average grey level deviation and the local spatial frequency to be used in order to determine JND_{GL} .

Based on the above considerations, we have developed the new ACE algorithm, called JND-Guided Adaptive Contrast Enhancement (JGACE), whose schematic diagram is shown Fig. 6.3. The core structure is based on unsharp masking with variable contrast gain G_S . To reduce the computational burden, the bilinear interpolation method used in LRM is also applied here. The input image is first divided into small contiguous blocks. Then, the contrast gain G_S is determined according to the local characteristics of each block and JND_{GL} , and is associated with the central pixel of the block. For any other pixel in the image, the contrast gain is calculated, using bilinear interpolation, from the G_S 's of the four neighboring blocks. In the following, we discuss each part of JGACE in more detail.

6.2.1 Low-pass filtering

The low-pass filtering is performed in spatial domain by convolving the input image with an equally weighted square window, since the spatial filtering can be done very fast with the use of the so-called "box-car" algorithm. Its purpose is to separate the

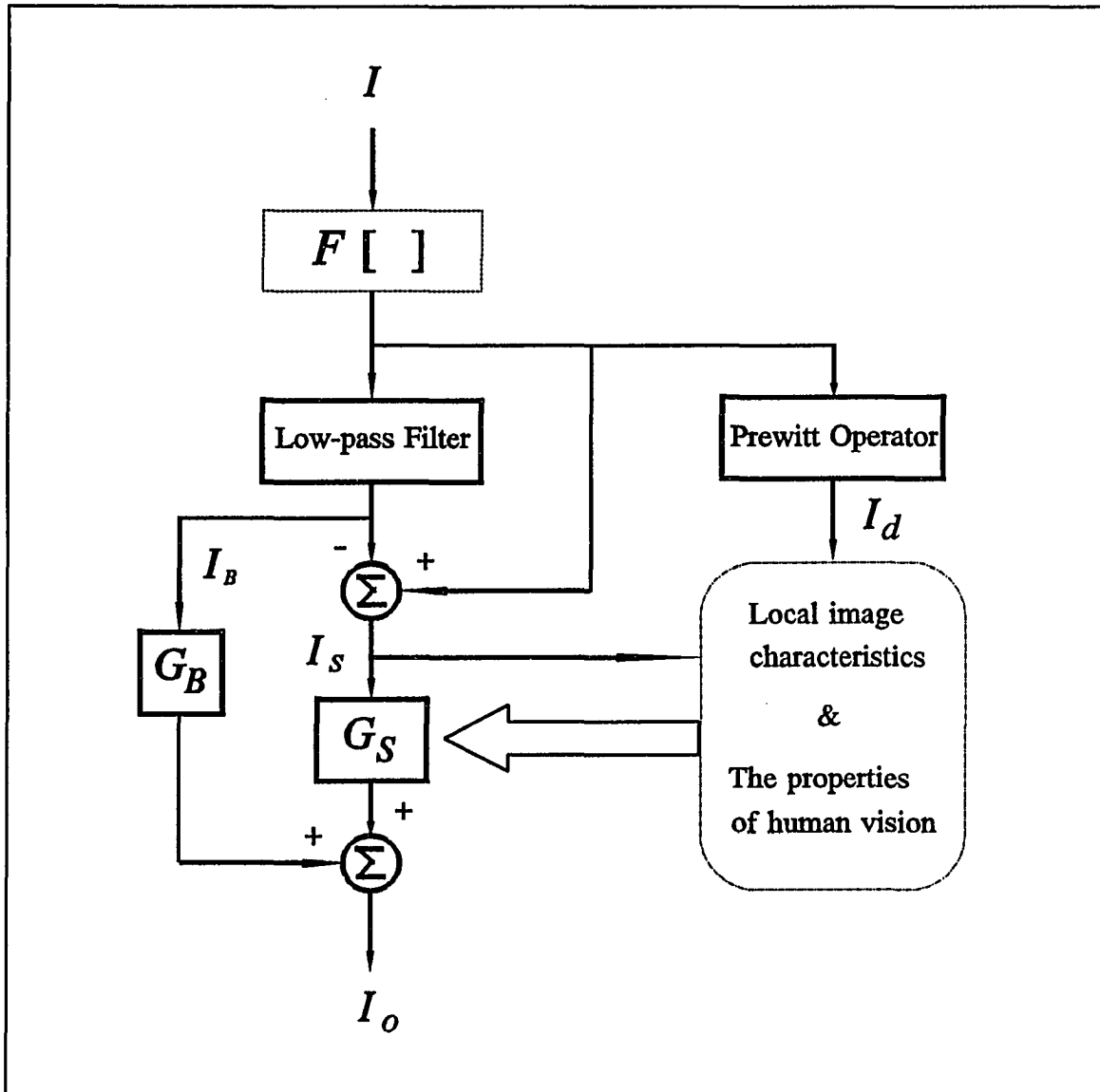


Fig. 6.3 A schematic diagram of JGACE algorithm.

very low frequency component $I_B(x,y)$, which contains only gradually-changing shadows from the more detailed structure component $I_S(x,y)$ of the original image $I(x,y)$. Thus,

$$I(x,y) = I_B(x,y) + I_S(x,y) \quad .$$

It is very important to determine an appropriate window size, and consequently the cutoff frequency of the low-pass filter. Ginsburg [46] has done an experiment showing how the spatial information of an image is distributed in various frequency bands. Based on his observation, it seems that a cutoff frequency of about 16 cycles per picture width is high enough for the filtered image to show the basic shape of the object presented in the image and also low enough to exclude all detail structures. Fig. 6.4 illustrates, as an example, the results of a low-pass filtering experiment with various window sizes. Based on this, a window size of $w/16$ is selected where w is the width of image. In order to speed up the filtering, the input image is first sub-sampled down to a sub-image of smaller size and then filtered, and the filtered result is then bilinearly interpolated to obtain the full-size filtered image. Since the energy in a real-world image is generally concentrated around low spatial frequencies, the aliasing caused by the sub-sampling has little effect on the low frequency components, and therefore the result of using the above procedure is almost identical to the result achieved by full-size low-pass filtering.

6.2.2 Separating detail and smooth regions

As performed in the implementation of the LRM, the input image is divided into contiguous blocks with the same size as the window size of the low-pass filter. Whether a block belongs to a smooth or a detail region is determined on the basis of the spatial

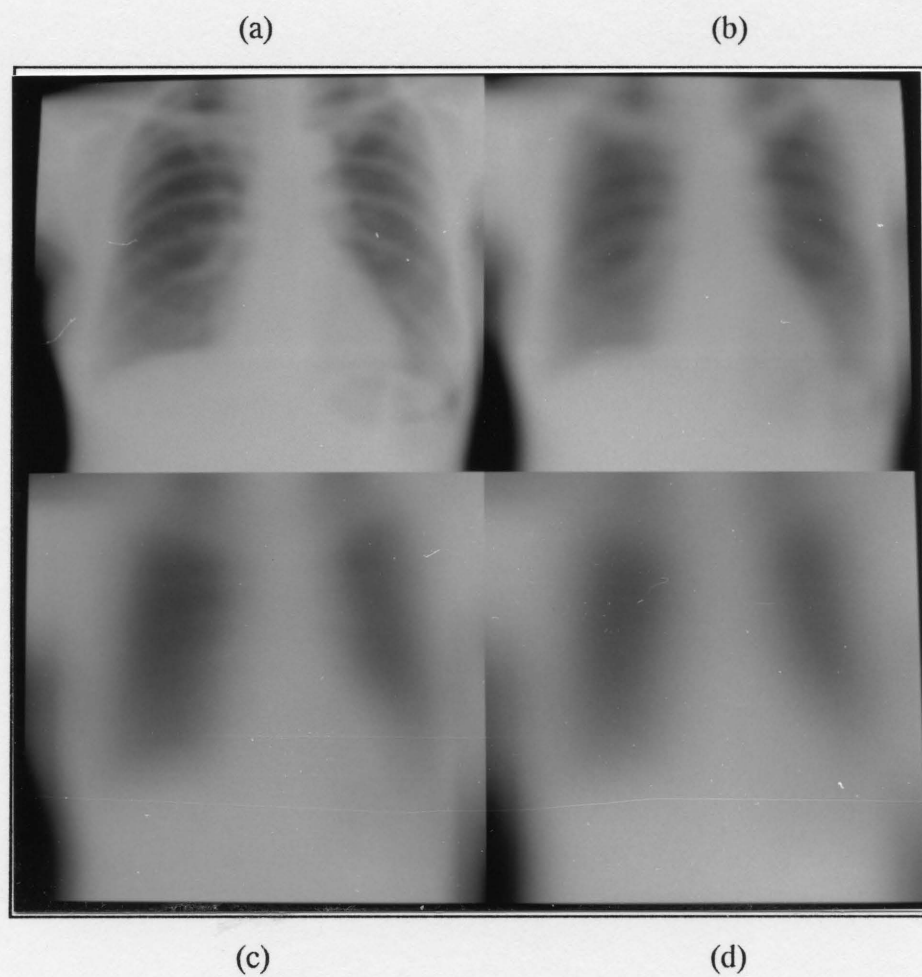


Fig. 6.4 The dependence of low-pass filtered background image on the filter's window size which is (a) $w/32$, (b) $w/16$, (c) $w/8$, (d) $w/4$, where w is the side width of the image.

activity within the block. To compute the spatial activity, we use the gradient image (i.e. the derivative image) obtained by applying two orthogonal Prewitt operators [26] to the input image $I(x,y)$. The Prewitt operator is chosen because it has a noise smoothing effect, and is also simple and fast. The gradient image $I_d(x,y)$ is given by

$$I_d(x,y) = \begin{cases} \Delta_x / 3 & |\Delta_x| \geq |\Delta_y| \\ \Delta_y / 3 & |\Delta_x| < |\Delta_y| \end{cases} \quad (6.8)$$

where

$$\begin{cases} \Delta_x = I(x+1,y-1) + I(x+1,y) + I(x+1,y+1) - I(x-1,y-1) - I(x-1,y) - I(x-1,y+1) \\ \Delta_y = I(x-1,y-1) + I(x,y-1) + I(x+1,y-1) - I(x-1,y+1) - I(x,y+1) - I(x+1,y+1) \end{cases} \quad (6.9)$$

Obviously, the amplitude of $I_d(x,y)$ is very small in smooth regions relative to detail regions. By thresholding $|I_d(x,y)|$ one should be able to separate smooth and detail regions. However, the threshold to be used varies with different images due to the differences in their noise levels. The selection of the threshold has to be made such that it is just above the noise fluctuation in smooth regions. To determine this variable threshold, we first calculate the average amplitude of $I_d(x,y)$ (i.e., the gradient mean) of each region. Notice that the gradient means of smooth regions tend to get together at lower values while the gradient means of detail regions distribute around higher values, and consequently the distribution of the gradient means generally possesses a bi-mode (sometimes multi-mode) shape. The first "dip" point of the distribution curve corresponds to a boundary between two types of regions (see Fig. 6.5). One could use this value, denoted by μ_T , as the gradient threshold. However, because the number of

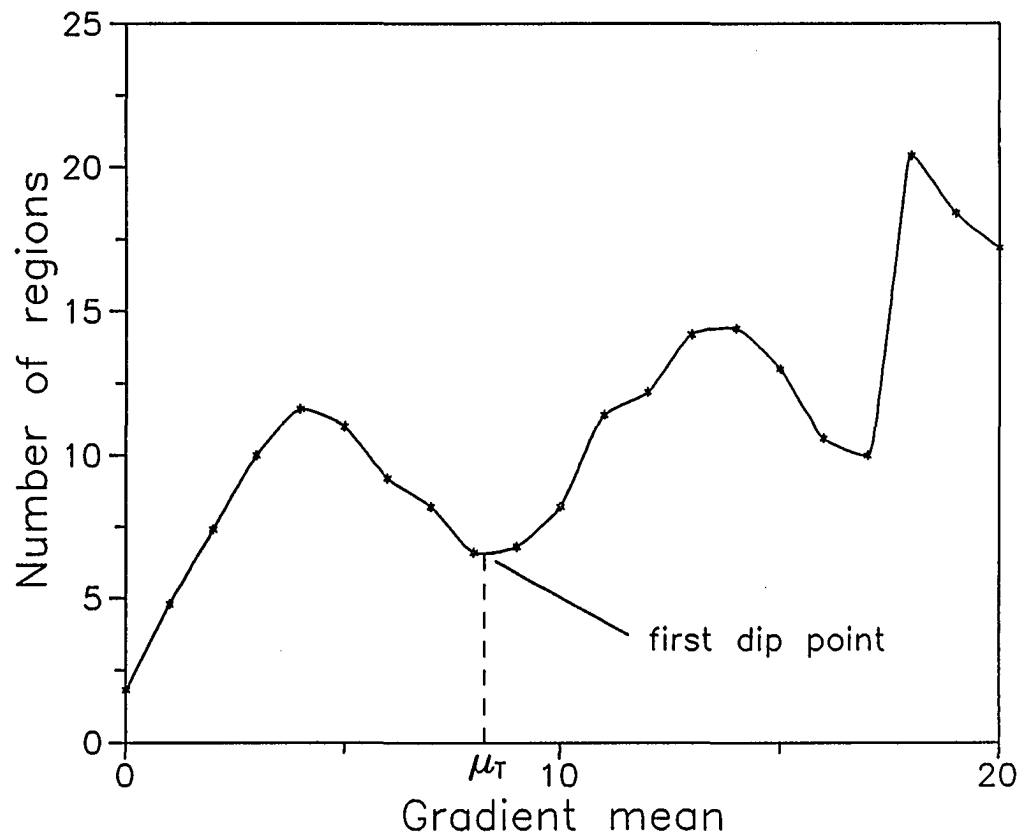


Fig. 6.5(a) *The distribution of regional gradient means of an image. Note that the initial threshold μ_T is given by the first "dip" point of the distribution curve.*

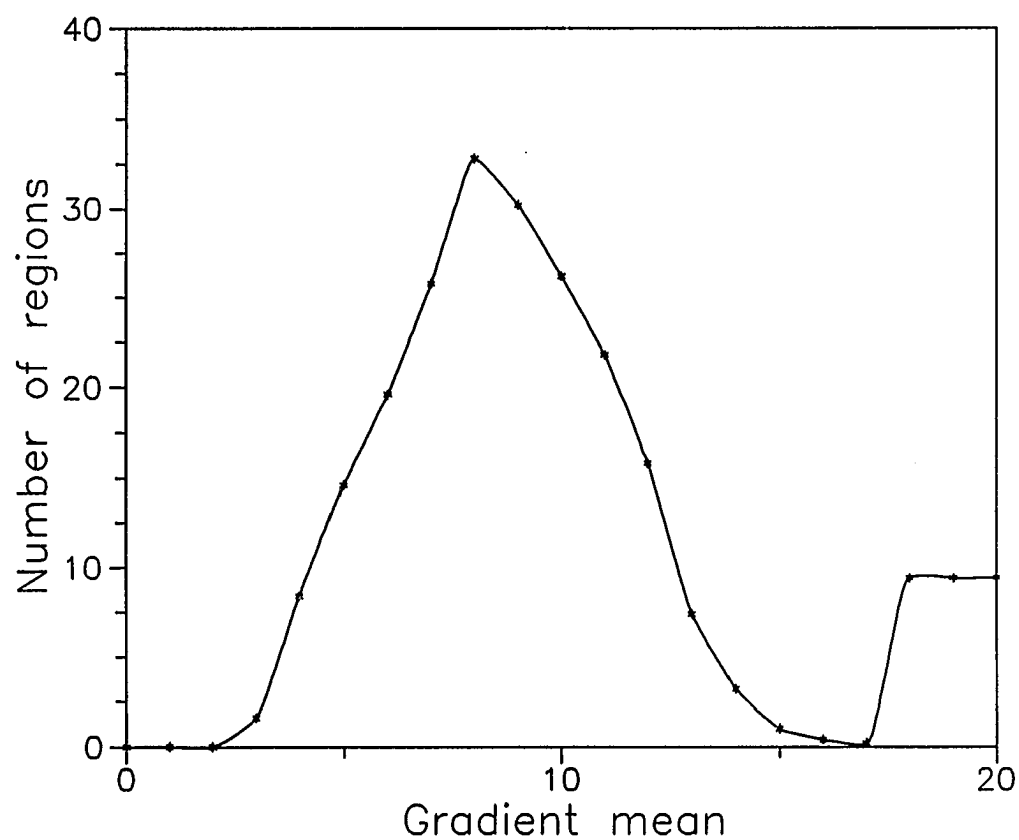


Fig. 6.5(b) *Another example of the distribution of regional gradient means. Note that the image has a large smooth area and relatively high background noise.*

regions in an image is relatively small, the distribution curve may be quite ragged and the position of the "dip" point may be in error, causing misclassification. Two approaches can be used to overcome this difficulty. The first is smoothing the distribution curve and the second is classifying those regions whose gradient mean is less than μ_T as smooth region samples, and then calculating the mean μ_g and the standard deviation σ_g of gradient amplitudes of all pixels in these smooth region samples. Note that μ_g and σ_g represent the average fluctuation in the smooth regions, attributed to background noise, and are less sensitive to a few possibly misclassified region samples. Hence, we select $T_2 = \mu_g + k_g \sigma_g$ as the final gradient threshold with $k_g = 3.5$ (it ensures that more than 99% of noise fluctuations in the smooth regions are below T_2 , assuming a Gaussian-like background noise). All contextual regions are then classified based on the following rule:

Define a spatial activity index, $p_k = N_k/N_{\max}$, where N_k is the number of pixels which satisfy $|I_d(x,y)| \geq T_2$ in the k th region, and $N_{\max} = \max\{N_1, N_2, \dots, N_k, \dots, N_K\}$. Thus p_k is actually a relative spatial activity. The region is classified as a smooth region if $p_k < p_T$, where $p_T = 0.1$, and is otherwise classified as a detail region (that is, a region whose relative spatial activity index is less than 0.1 is considered as a smooth region). One benefit of separating smooth and detail regions first is that it eliminates the unnecessary calculation of local spatial frequencies and contrasts for smooth regions.

6.2.3 Determination of local spatial frequency and contrast

To determine the local spatial frequency for each detail region, we use $I_d(x,y)$ to

calculate the ratio of the number of zero-crossings along a row (or column) to the number of pixels in that row (or column). This ratio approximately represents the local spatial frequency. If a detail region consists of n rows and m columns, one can obtain a spatial frequency as defined above for each row and each column. The average spatial frequencies in the horizontal and the vertical directions are given by

$$f_H = \frac{1}{n} \sum_{j=1}^n f_j, \quad f_V = \frac{1}{m} \sum_{i=1}^m f_i \quad (6.10)$$

where f_j and f_i are the spatial frequencies of the j th row and the i th column in the detail region, respectively. Then, the local spatial frequency of the region is defined as $f_R = \max(f_H, f_V)$. Notice that f_R is actually a normalized frequency, i.e., $f_R = 1$ corresponds to 0.5 cycles/pixel. The relationship between the angular spatial frequency u and f_R is given by

$$u = \frac{\pi d}{360 \Delta x} \cdot f_R \quad (6.11)$$

where d is the viewing distance which is assumed to be 50 cm for usual observing conditions, and Δx is the pixel size.

As discussed previously, the local contrast of each detail region is simply represented by the local average grey level deviation calculated using the structure image. With exclusion of the small fluctuations caused by the background noise, this local average grey level deviation, denoted by Δ_R , is computed by

$$\Delta_R = \frac{1}{n} \sum_{|I_S(x,y)| > \epsilon} |I_S(x,y)| \quad (6.12)$$

where ϵ is a noise threshold which can be set equal to T_2 , and n is the number of pixels satisfying the condition $|I_S(x,y)| > \epsilon$ in the region.

6.2.4 Determination of local contrast gain

In smooth regions, as we discussed before, the local contrast should be generally kept unchanged to avoid a possible over-enhancement of noise. Hence, we would like to have $G_S = 1$ for all smooth regions. However, for some images that have large background noise, it may be desirable to let the gain of the smooth region be less than one. On the other hand, for some noise-free images, or images with very high signal-to-noise ratio, it may be desirable to have G_S greater than one for smooth regions so that some subtle structures in smooth regions can be enhanced without worrying about too much noise enhancement. To meet this goal, one can design G_S as a function of the noise level in smooth regions. One approach that has been used for tailoring G_S is given by

$$G_S = G_1 + \frac{G_2}{1 + (\mu_g / \mu_{g,0})^4} \quad (6.13)$$

where $G_1 < 1$ and $(G_1 + G_2) > 1$, the mean gradient of smooth regions μ_g is used as a measure of noise level in smooth regions, and the value of $\mu_{g,0}$ is set about 1 ~ 2 percent of GL_{\max} (for 8-bit images, we have selected $\mu_{g,0} = 3.5$).

For detail regions, the local contrast gain G_S is determined by comparing the local average grey level deviation Δ_R with the calculated $\gamma \cdot JND_{GL}$. Recall that JND_{GL} is a function of the angular spatial frequency u and follows the same form of function as $C_T(u)$. For the mid to high spatial frequencies, $C_T(u)$ can be well approximated by an

exponential function: $C_T(u) = c_0 \cdot e^{0.166u}$ [44]. Hence, JND_{GL} becomes

$$JND_{GL} = JND_0 e^{0.166u} \quad (6.14)$$

where JND_0 is the JND in units of grey level at low spatial frequencies. As discussed in the earlier chapters, JND_0 depends on the noise performance of the display device. For a given display device, JND_0 can be assumed to be a known constant. The local contrast gain G_S is then given by

$$G_S = \begin{cases} 1 & \Delta_R \geq \gamma \cdot JND_{GL} \\ \frac{\gamma \cdot JND_{GL}}{\Delta_R} & \Delta_R < \gamma \cdot JND_{GL} \end{cases} \quad (6.15)$$

where γ is an user-specified constant.

6.2.5 The transformation $F[\cdot]$

The transformation $F[\cdot]$ is optional. Its purpose is to provide an adequate global image contrast if the global contrast of the input image is not sufficient. When the histogram of the original image only occupies a small portion of the entire dynamic range of the display device, a global contrast enhancement is usually desired before the adaptive local contrast enhancement. The simple linear contrast stretch described in §6.1 may not be suitable for this purpose if the histogram of the image has long "tails". Such an example is shown in Fig. 6.2(a) (dashed line) which is the histogram of a chest image, where although most pixels of the image are distributed within a relatively narrow grey level range, the histogram with a long "tail" spreads over almost the entire grey level

range. If we just apply a linear contrast stretch to that narrow grey level range, some information of the image will be lost. What is really needed is to stretch the major portion of the histogram and at the same time compress those grey level ranges where the value of the histogram is small. The clipped histogram equalization used in CLAHE can achieve this goal. But, unlike CLAHE, this clipped histogram equalization is applied only once to the whole image. The only question is how to determine the global contrast gain limit G_C . Obviously, G_C depends on the desired global contrast. We describe the global contrast of an image by its overall standard deviation of pixel values. From our experience, it may be inferred that if the image standard deviation σ is greater than one fifth of the entire display range, the global contrast is usually adequate. The standard deviation after the clipped histogram equalization is approximately $G_C \cdot \sigma$. Hence, we set $G_C = GL_{\max}/(5 \cdot \sigma)$. The derived non-linear transformation $F[\cdot]$ can be easily and efficiently implemented by a look-up table, and an adequate global contrast can be achieved before the local ACE starts.

6.3 Results and Comparison with Other Algorithms

JGACE has been applied to a variety of images, including many chest radiographs. In the following, we shall present the results obtained by applying JGACE and some other existing algorithms to two representative images: (i) a test image consisting of both high and low contrast sinusoidal patterns and two sharply-edged squares with a large uniform background and subtle white noise being added over the

entire image, and (ii) a clinical chest image with an added dark rectangle for the purpose of showing more clearly the possible ringing artifacts which may be generated by these algorithms. Both images are 512×512 pixels and are digitized to 8 bits.

Fig. 6.6 shows the original test image and the results of processing by five different algorithms. Cross profiles along the middle of each processed image are plotted in Fig. 6.7 to present more clearly the effects of different methods. The first method applied was the global histogram equalization. The second algorithm was a slight modification of the original LRM with $C_2 = 0$, and C_1 being given by

$$C_1 = I_{\max} - I_{\min}$$

and

$$I_{\min} = \text{MAX}(0, \mu_{\Omega} - G_{\Omega}(\mu_{\Omega} - \min_{\Omega}))$$

$$I_{\max} = \text{MIN}(255, \mu_{\Omega} + G_{\Omega}(\max_{\Omega} - \mu_{\Omega}))$$

where μ_{Ω} is the local mean, \max_{Ω} and \min_{Ω} are the local maximum and the minimum, respectively, and G_{Ω} is a fixed, user-specified local contrast gain. This modification, although adds one more parameter G_{Ω} to be determined by the user, prevents possible over-enhancement when $(\max_{\Omega} - \min_{\Omega})$ is too small. The contextual block size was selected such that both mid and high spatial frequencies were included within the block. According to the discussion given in §6.2.1, a block size of 32×32 pixels was used. The local contrast gain was chosen at a medium value of $G_{\Omega} = 2$. The third method used was a multichannel filtering with three channels. The window sizes of the two low-pass filters were 7×7 and 25×25 (as suggested by Tahoces *et al.* [63]). The three gain coefficients for the multichannel filtering for this test image were 2.0 for the high

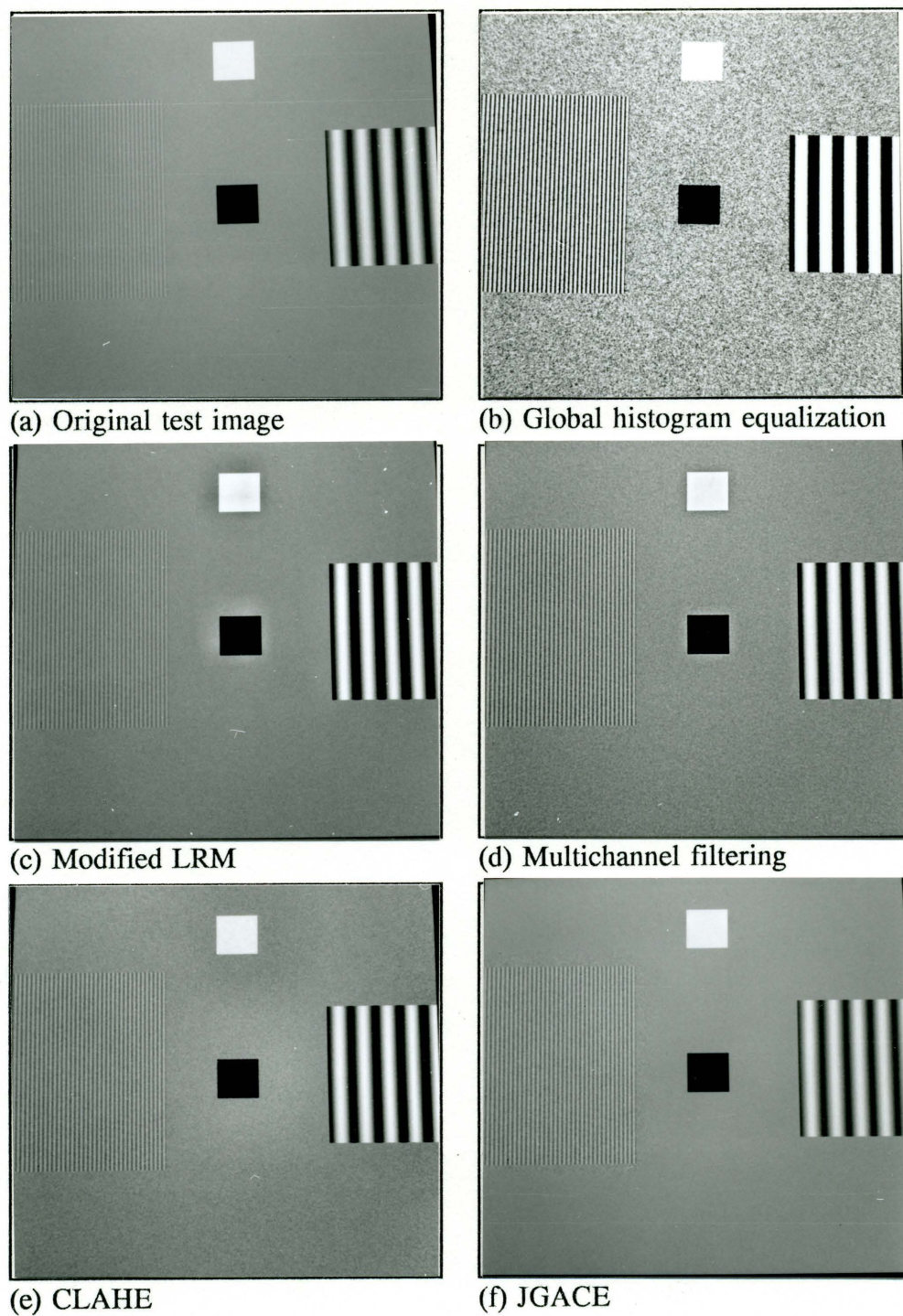


Fig. 6.6 Comparison of five different contrast enhancement algorithms for a test image.

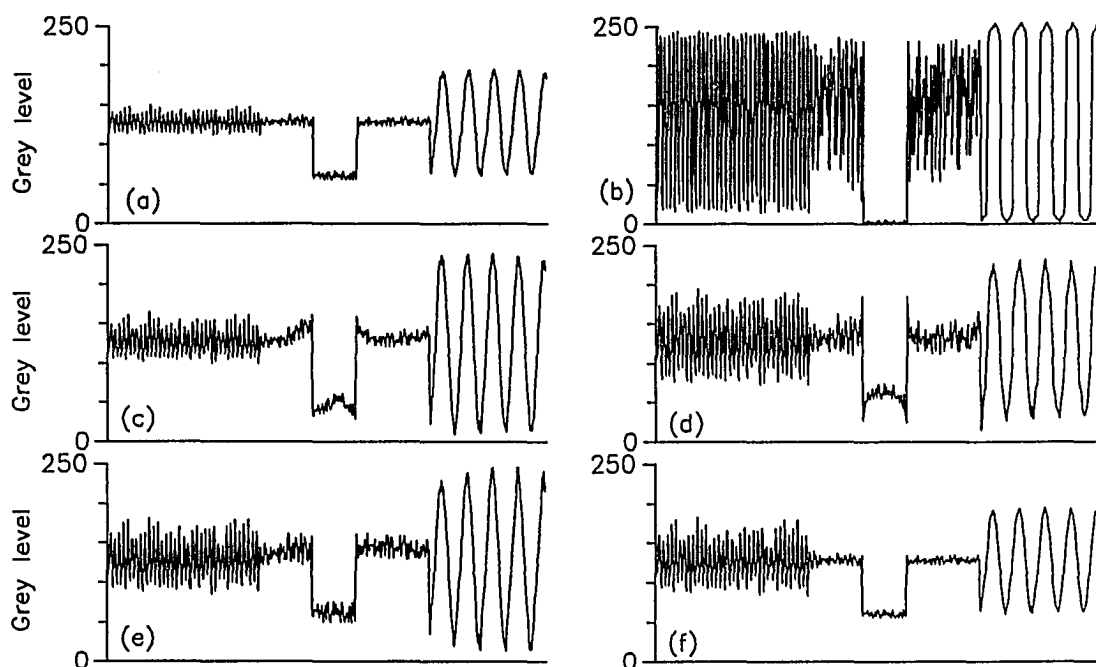


Fig. 6.7 A cross profile along the middle of the test image: the original image (a), the image processed by a global histogram equalization (b), by the modified LRM (c), by the multichannel filtering (d), by CLAHE (e), and by JGACE (f).

frequency band, 1.5 for the mid frequency band and 0.95 for the low frequency band. The fourth technique employed for comparison was CLAHE for which the contrast gain limit was 2.0, and the side length of the contextual block was 127 pixels with 50% overlapping. The fifth method was JGACE. The contextual block size was also 32×32 pixels with 50% overlapping. The parameter γ was selected as 6.

From the results shown in Fig. 6.6 and Fig. 6.7, it is clear that the histogram equalization method produced, as expected, the worst noise enhancement, and caused a saturation of the high contrast sine-wave pattern due to over-enhancement. The modified LRM enhanced both sinusoidal patterns as well as the background noise, and the ringing artifacts are strongly noticeable around the sharp-edged squares. The contrast enhancement by the multichannel filtering method appears to be stronger than the modified LRM, while resulting also in stronger ringing artifacts and noise enhancement. CLAHE produced a considerable contrast enhancement for both sine-wave patterns and noise in smooth regions, but it had less ringing artifacts. JGACE brought up an adequate contrast for the low-contrast sinusoidal pattern without enhancing noise in the uniform background, and had no ringing artifacts. It may also be noticed that those high contrast signals in the test image (the squares and the low spatial frequency sine-wave pattern) are kept unchanged by JGACE because their contrast is judged to be already adequate on the basis of the human visual JNDs.

Fig. 6.8 shows the original chest image and the results of processing by the same five contrast enhancement algorithms that were discussed in the earlier case. As in the case of the test image, cross profiles along the middle of the chest image (through the

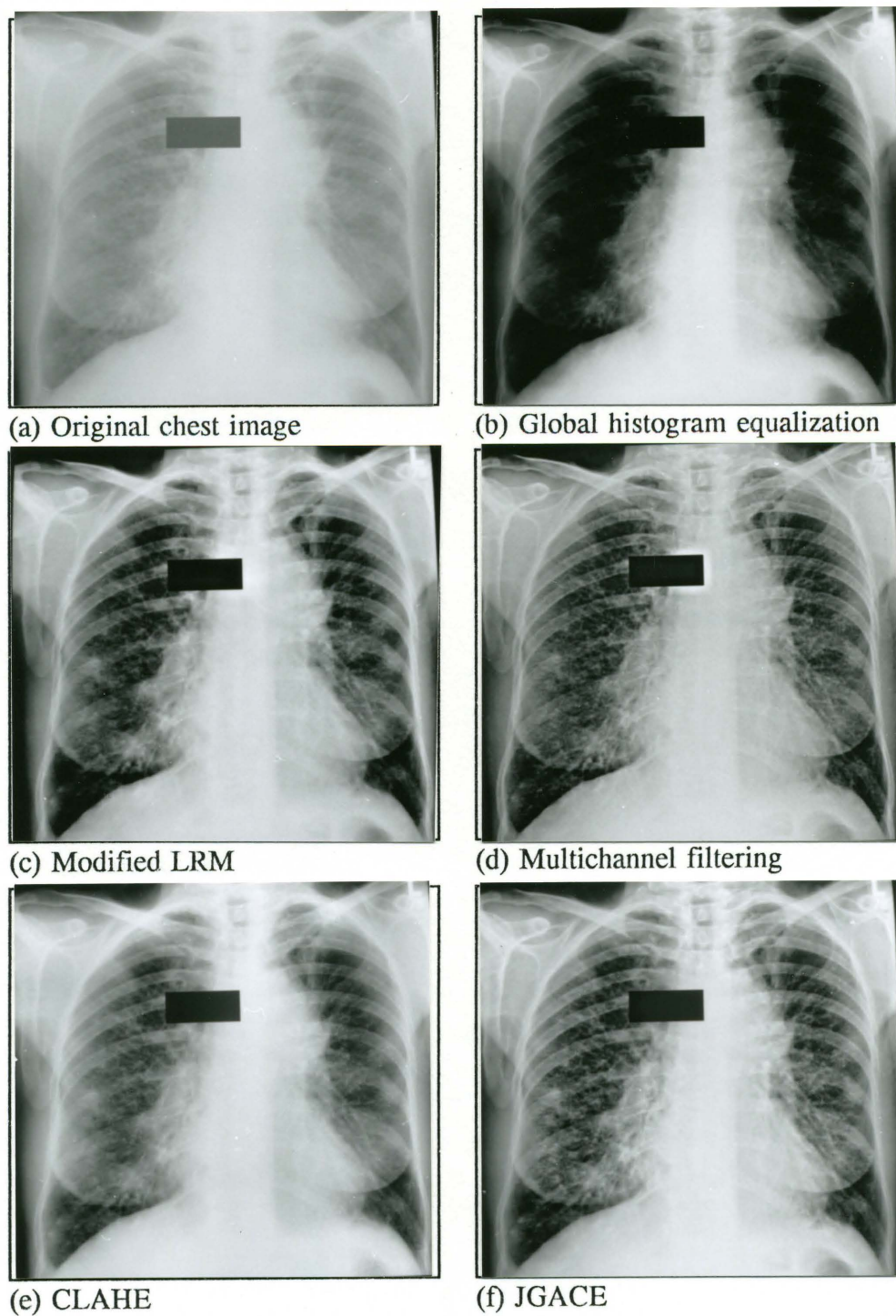


Fig. 6.8 Comparison of five different contrast enhancement algorithms for a chest image.

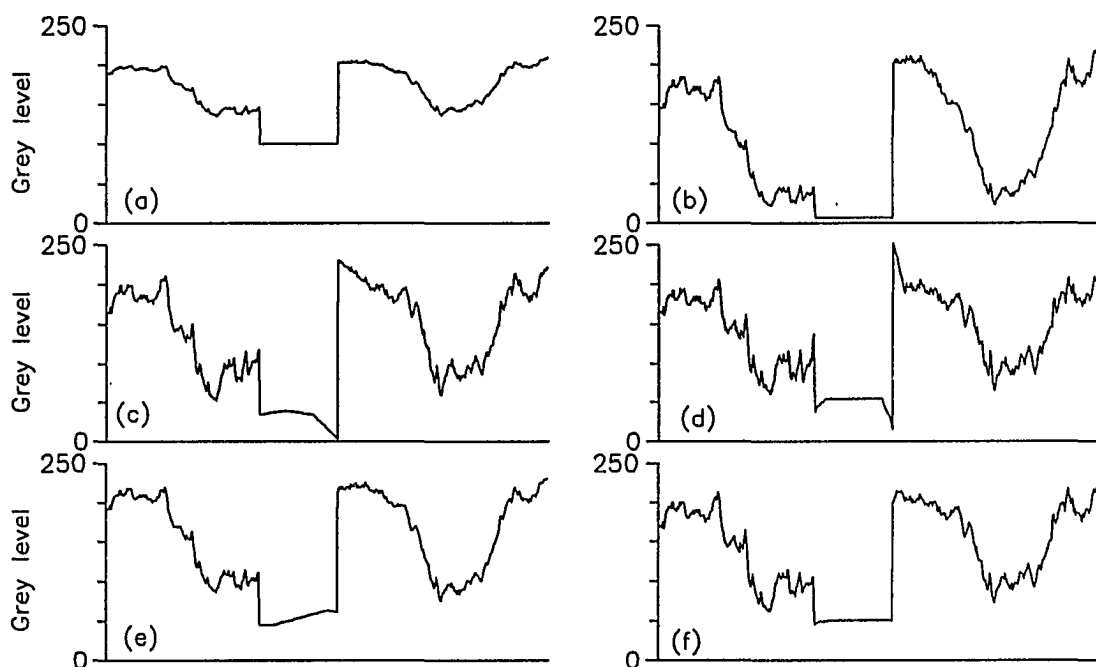


Fig. 6.9 A cross profile along the middle of the chest image (through the rectangle): the original image (a), the image processed by a global histogram equalization (b), by the modified LRM (c), by the multichannel filtering (d), by CLAHE (e), and by JGACE (f).

rectangle) are plotted in Fig. 6.9 for better inspection. The three gain coefficients of the multichannel filtering were changed a little in an attempt to obtain the best possible results with this method. These are 2.5 for the high frequency band, 1.8 for the mid frequency band and 0.9 for the low frequency band. All the other parameters were kept at the same values as before. No changes were made to the other algorithms. From Fig. 6.8 and Fig. 6.9 we again see that the results of processing with JGACE are better than with any other methods in comparison.

Besides JGACE, CLAHE appears to have the best processed image among the other algorithms compared here. Our experimental experience with chest images has shown that in most cases the results of JGACE and CLAHE are similar if background noise in the original image is low. To further compare the two methods, the computing time was compared for the same image on the same computer (VAX 8600). The results are summarized in Table 6.1. The last column of the table gives the computing time ratio of JGACE to CLAHE. The table shows that JGACE is faster than CLAHE, especially for large pixel bit depth (the number of bits per pixel).

TABLE 6.1 Comparison of computing time (minutes) for JGACE and CLAHE

Image Format	CLAHE	JGACE	Time Ratio
512 x 512 x 8	0.59	0.56	0.95
512 x 512 x 12	1.40	0.62	0.44
1024 x 1024 x 8	2.76	2.18	0.79
1024 x 1024 x 12	4.07	2.22	0.55
2048 x 2048 x 12	18.6	8.73	0.47

6.4 Discussion

In order to perform a quantitative evaluation of a given contrast enhancement algorithm, let us analyze the change in spatial activity of the test image before and after processing. As proposed by Netravali and Prasada [73], the image spatial activity can be described by a masking function defined as

$$M(x,y) = \sum_{p=x-i}^{x+i} \sum_{q=y-j}^{y+j} \alpha^{\|(x,y)-(p,q)\|} \left[\frac{1}{2} (|G_{pq}^H| + |G_{pq}^V|) \right], \quad (6.16)$$

where

$$\begin{aligned} G_{pq}^H &= I(p,q) - I(p-1,q) \\ G_{pq}^V &= I(p,q) - I(p,q-1) \end{aligned} \quad (6.17)$$

and $\|(x,y)-(p,q)\|$ is the Euclidean distance between pixel (x,y) and pixel (p,q) . α is a constant ($\alpha < 1$) controlling the rate of exponential decay of the effect of the image intensity change at pixel (x,y) on its neighbors, and i, j are constants controlling the size of the neighborhood around (x,y) . A large value of $M(x,y)$ represents a high spatial activity around pixel (x,y) . Let us divide the entire range of the masking function of an image into m equally spaced intervals:

$$[M_0, M_1), [M_1, M_2), [M_2, M_3), \dots, [M_{m-1}, M_m]$$

where M_0 and M_m are the minimum and the maximum masking values, respectively. A pixel (x,y) in the image is assigned into the spatial activity region A_k if $M(x,y) \in [M_{k-1}, M_k)$, ($k=1,2,\dots, m$). The average spatial activity in each spatial activity region is calculated as

$$\bar{M}_k = \frac{1}{n_k} \sum_{(x,y) \in A_k} M(x,y) \quad . \quad (6.18)$$

Let \bar{M}_k^b and \bar{M}_k^a denote the average spatial activities before and after contrast enhancement, respectively. Then the difference $D_k = (\bar{M}_k^a - \bar{M}_k^b)$ represents the average increase of spatial activity in each spatial activity region. By examining $M(x,y)$ and the contents of an image, one can always divide the image, based on its masking value, into three regions: (a) background noise (low spatial activities), (b) image details (medium spatial activities), and (c) strong edges (high spatial activities). It is of interest to see how D_k varies in these three spatial activity regions with different contrast enhancement algorithms. Fig. 6.10 plots D_k as a function of \bar{M}_k^b for the test image in the case of the five algorithms compared above. In calculating D_k , a 3×3 pixel neighborhood was selected (i.e., $i = j = 1$), and the value of α was taken to be 0.35, as suggested by Netravali and Prasada [73]. The spatial activity range of the image details is observed from about 35 to 70 for this test image. We can see clearly from Fig. 6.10 that only JGACE provides the largest enhancement in the image detail regions relative to the other two regions. This explains why JGACE not only enhances the detail structures in the image, but also increases the signal-to-background noise ratio of the image and suppresses the ringing artifacts around strong edges.

Image contrast enhancement is a rather old problem. Generally speaking, it can be achieved by image processing in either frequency domain or spatial domain. Since for an image the random noise usually has a very broad spectrum mixed together with

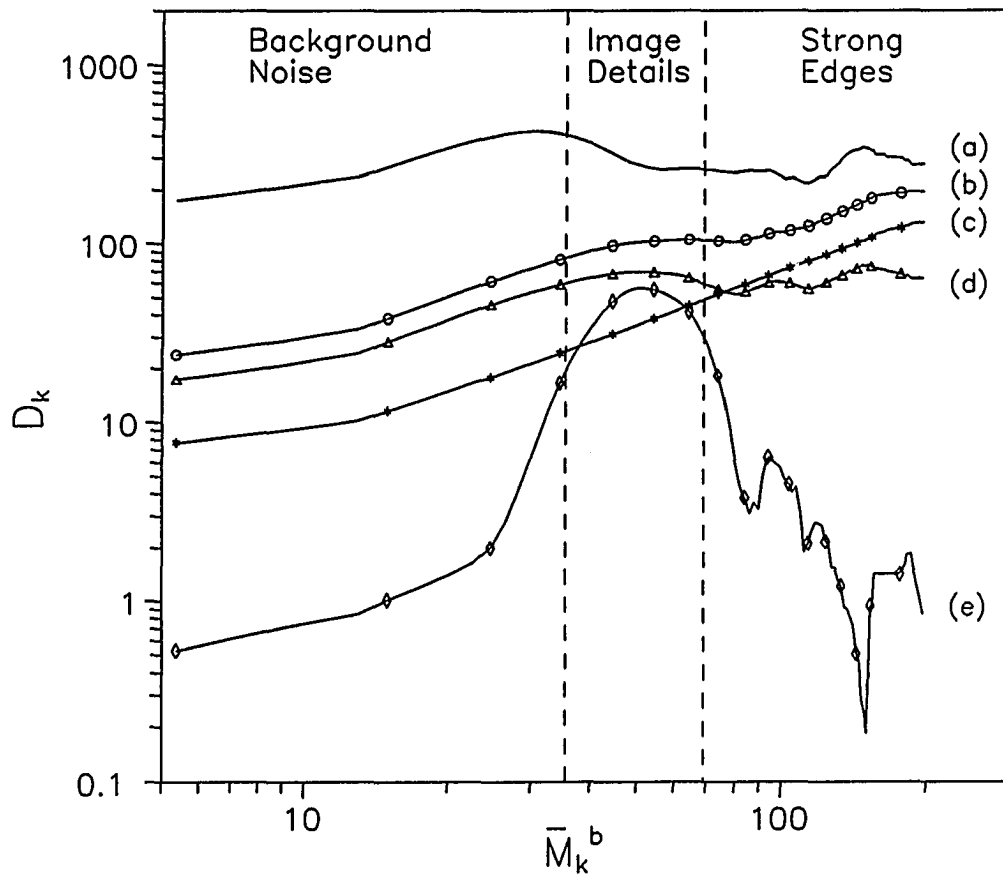


Fig. 6.10 The increase of spatial activity varies as a function of the original spatial activity of the test image in Fig. 6.6(a) for five different algorithms: (a) a global histogram equalization, (b) the multichannel filtering, (c) the modified LRM, (d) CLAHE, and (e) JGACE.

the spectrum of signal (structure details), any manipulation of the image spectrum in frequency domain (e.g., filtering) will have the same effect on both the signal and the noise if their spectra are inseparable. It is impossible, in the frequency domain, to boost the signal contrast without enhancing noise. Also, enhancing the mid to high frequency components in the frequency domain always creates ringing artifacts at sharp edges because the frequencies at a sharp edge are much higher than those at its surroundings. If the contrast of the sharp edge is small, this ringing effect just gives an appearance of enhanced contrast for the edge. But, if the contrast of the sharp edge is large, the ringing effect becomes a severe artifact. Hence, noise enhancement and ringing artifacts are two unavoidable problems for any frequency domain enhancement algorithm, such as multichannel filtering. A tradeoff has to be made between enhancing signal contrast and avoiding excessive noise enhancement and ringing artifacts.

On the other hand, in spatial domain one can certainly handle these two problems better, as shown by JGACE. Two keys for the success of JGACE are the JND-guided contrast gain selection, which permits a quantitative determination of adequate contrast enhancement for local details of various frequencies, and separation as well as different treatments of smooth and detail regions, which effectively eliminates the enhancement of noise in smooth regions. Since the noise visibility is high only in smooth regions, the consequence is that the perceived signal-to-noise ratio is increased. The possibility of ringing artifacts is automatically suppressed, because around the high contrast edges the local contrast has already been adequate as judged by the local JND, and thus no enhancement occurs there to cause the ringing artifacts. Experience with JGACE

algorithm further shows that more overlapping between the adjacent blocks makes the change of the local contrast gain G_S smoother at the expense of more computations.

Among histogram based algorithms, CLAHE provides the best processing performance without the over-enhancement problem by utilizing the clipped local histograms. The major computation of CLAHE is attributed to the calculation and the manipulation of local histograms for each block of image. Hence, the larger the number of grey level bins, the more computation time is required. Because the number of grey level bins in each histogram to be calculated increases exponentially with the pixel bit depth of the image, the computing time of CLAHE increases sharply from an 8-bit image to a 12-bit image even though the two images have the same number of pixels. This can be clearly noticed in Table 6.1. However, the speed of JGACE is almost independent of the pixel bit depth since no histogram is required in the major portion of the algorithm, except that only one global histogram may need to be calculated for the optional transformation $F[\cdot]$.

It should be noted that γ and JND_0 are system dependent parameters. Selection of values for these parameters may vary with different display devices, observers and viewing conditions. However, the product of γ and JND_0 can be actually considered as one parameter (see (6.15)). It can be adjusted to an appropriate value once for a given display device and viewing condition, and is then kept fixed. Alternately, several enhancement levels can be preset (by changing $\gamma \cdot JND_0$) for an observer to select the most appropriate one for his (her) specific application.

6.5 Summary

The presently developed JGACE algorithm is quite effective in enhancing image contrast for the best visualization. Unlike some special ACE algorithms suitable only for one type of image, this new algorithm can be applied to all kinds of images, including chest radiographs. JGACE meets the four desired goals of image contrast enhancement, as proposed in §6.1. The quantitative determination of an adequate contrast enhancement is obtained by utilizing, as a guide of local contrast gain, the human visual properties as well as the local characteristics of the image. The two common problems existing in many earlier developed ACE techniques are overcome by the present approach. The possible noise enhancement is minimized by separating smooth and detail regions of an image and processing them differently. Hence, the perceived signal-to-noise ratio is increased. The ringing artifacts are successfully suppressed because of the appropriate JND-guided control of the local contrast gain. The computation burden is considerably reduced by three factors. First, the local characteristics (the average spatial frequency, local contrast) and the contrast gain are calculated only for the center of blocks, and the entire processing of the image is then implemented by linear interpolation, as used by Fahnestock and Schowengerdt [64], and Pizer *et al.* [68]. Secondly, separating smooth and detail regions saves the unnecessary calculation of local spatial frequencies and local contrasts for smooth regions, especially if the image has a large smooth area. Finally, low-pass filtering is applied to the sub-sampled, smaller size image, and then a linear interpolation is used to retrieve the full-size filtered image.

As discussed in Chapter 5, to best visualize the contrast information presented by an image, perceptual linearization is applied to the display device to display the image processed by the JGACE algorithm, which can be easily implemented by a look-up table designed specifically for the display device in use. For application to medical images, e.g., chest radiographs, it is almost certain that radiologists will be better equipped to detect small contrast features more easily on the processed chest images. Additional observer performance studies (ROC studies) are needed in the future to evaluate, in a more rigorous manner, if this technique can provide a significant diagnostic advantage over the original images.

CHAPTER 7

CONCLUSIONS

7.1 Contributions of This Dissertation

In this section, the major contributions of this dissertation are summarized as follows.

- 1) A systematic and quantitative evaluation of both physical and psychophysical performance of image display devices is developed. This evaluation includes hardcopy display devices (e.g., laser film printers) and softcopy display devices (such as high-resolution monochrome CRTs). In view of the image quality presented by the display devices, the emphasis is placed on those display system parameters that are most closely related to the performance of image visualization. The physical performance parameters of great interest are display function, physical dynamic range, internal scatter, spatial resolution (MTF), noise characteristics (including signal-to-noise ratio and noise power spectra), contrast transfer factor and luminance uniformity. The correlations among these parameters are also investigated. For perception performance, threshold contrast, just-noticeable-difference (JND) and perceived dynamic range are the evaluated parameters. A mathematical expression for the visual luminance response function is derived in Chapter 3 which allows one to develop an optimum display function for image display devices. In the past, the determination of these

psychophysical parameters has required considerable observer involved experiments which are very tedious and time-consuming. In Chapter 4, a direct quantitative relation between the physical and psychophysical parameters is established (given by equations (4.7), (4.8) and (4.12)). One of principal benefits resulting from these studies is that very tedious and time-consuming psychophysical experiments can be replaced by more reliable and faster physical measurements. It is concluded that in the present state of modern CRTs, the spatial noise due to phosphor granularity offers the major limit to the contrast resolution of CRTs, and that trying to decrease the spatial noise of CRT is a more effective approach to increase the perceived dynamic range, among other considerations.

- 2) Optimization of the display function for image devices is developed in Chapter 5 from the point of view of contrast information transfer. The display function of a display device, having been given very little attention for a long time, is probably the most important characteristic which can dramatically influence the appearance of the displayed image. It is obvious that in order to maximize the transferred contrast information of an image, the display function of the display device must somehow match the human visual response to the displayed luminance. The result presently derived by using the display device/human observer model and concepts from information theory indicates that the optimum display function is mathematically the inverse of the scaled visual response

function which can be derived from the JND curve, and that the optimum display function is independent of the object size and the noise level (RMS) of the display device. The optimum display function perceptually linearizes the display device so that equal changes in grey level produce changes in luminance that are perceptually equal throughout the entire dynamic range of the display device. It also allows a meaningful comparison of different display devices for various images. Although the optimum display function is not necessarily the best display function for a specific display application, it fully utilizes the maximum potential of the contrast capability of a display device in the sense that the maximum perceived dynamic range (i.e., the maximum number of just-noticeable-differences) can be achieved only with the optimum display function.

- 3) A novel adaptive contrast enhancement algorithm (JGACE) is developed in Chapter 6, using human visual JND as a guideline, along with a perceptually linearized display device, for the best visualization of needed contrast information in an image. Adaptive contrast enhancement, as an effective way to improve the image display, has long been investigated by many researchers. It is the first time, however, that an adaptive contrast enhancement algorithm that quantitatively achieves an adequate amount of contrast enhancement in terms of the human visual JNDs is proposed. This scheme effectively solves two common problems associated with many other ACE techniques — ringing artifacts around sharp edges and over-enhancement of the background noise. The present algorithm can

be applied to a variety of images, including chest radiographs to facilitate a better diagnosis of abnormalities that have low contrast features.

7.2 Suggestions for Future Investigation

Optimization of image display is a complex subject because it is influenced by many variables which are related to the display system itself, to the human visual system and to the viewing environment. The basic goal of optimizing image display is to match the display system to the human visual system for the given viewing condition. Further investigations on the characteristics of the human visual system are desired to understand better how the human visual system responds to different stimuli and different viewing conditions.

A number of researchers [21], [22], [74] have proposed to utilize perceptual linearization as a display standard. Some of the advantages that have been proposed for developing and adopting perceptual linearization as a display standard are the following [74]:

- The relation between luminance and digital input grey level can be defined by mathematical functions.
- The grey scale rendition of the CRT monitor will be predictable and reproducible.
- Similarity between hardcopy and softcopy images will be facilitated.
- Valid and reliable observer performance studies on different display

modalities can be carried out and the results will have more general meaning than previously.

Notice that strictly speaking, perceptual linearization would require a non-uniform quantization of the driving signal of a display device, which may not be acceptable since all commercial A/D converters used in display devices now have uniformly spaced levels. An alternative approach may be to employ a conventional D/A converter having more than 8 bits (e.g., 10 or 12 bits) so that the non-uniform quantization required by perceptual linearization can be well approximated by a designed look-up table with finer steps of digitization. This software implementation of perceptual linearization has another advantage that it becomes easy to calibrate the display device in use by changing the look-up table based on the possible changes in the intrinsic display function caused by internal and/or external factors (e.g., the system changes, viewing condition changes, etc.).

In the presently developed JGACE algorithm, the local spatial frequency of a region is determined approximately by counting the average number of zero-crossings in the region and the local contrast is represented by the average grey level difference between the structure and the background in the region. This is only for simplicity and is based on the assumption that the local characteristics of the image are homogeneous within the region, which may not be true in all cases. More sophisticated approaches may be designed to segment the image into meaningful object regions instead of fixed size regions and to calculate the local characteristics for true objects, although more computations may be needed.

With the application of JGACE to medical images, such as digital chest radiographs, it is obvious that the low contrast features which may be important for diagnosis will be enhanced by JGACE without the need for windowing and leveling manipulations by radiologists. A definitive determination of whether it can improve the diagnostic accuracy or save doctor's diagnosing time (or both) needs to be evaluated with considerable observer performance studies in the future.

REFERENCES

- [1] S. Webb, Editor, The Physics of Medical Imaging, Adam Hilger, Bristol and Philadelphia, 1988.
- [2] T. S. Curry III, J. E. Dowdey, and R. C. Murry Jr., Christensen's Physics of Diagnostic Radiology, 4th Edition, Lea and Febiger, Philadelphia, 1990.
- [3] R. L. Arenson, D. P. Chakraborty S. B. Seshadri and H. L. Kundel, "The Digital Imaging Workstation," *Radiology*, Vol. 175, 1990, pp. 303-315.
- [4] J. E. Tannas Jr., Editor, Flat-Panel Displays and CRTs, Van Nostrand Reinhold, New York, 1985.
- [5] F. Terman, Electronic and Radio Engineering, McGraw-Hill, New York, 1955.
- [6] M. Browne, "Noise Limited Performance of a Hybrid Detector and High Resolution Display Monitors," *Ph.D Dissertation*, Optical Sciences Center, University of Arizona, 1991, pp. 68-74.
- [7] L. M. Biberman, Editor, Perception of Displayed Information, Plenum Press, New York, 1973.
- [8] W. N. Charman and A. Olin, "Tutorial: Image Quality Criteria for Aerial Camera Systems," *Photographic Science and Engineering*, Vol. 9, 1965, pp. 385-397.
- [9] H. C. Borough, R. F. Fallis, R. H. Warnock, and J. H. Britt, "Quantitative Determination of Image Quality," *Boeing Company Technical Report D2-1140581-1*, 1967.
- [10] H. L. Snyder, "Image Quality and Observer Performance," In Perception of Displayed Information, ed. L. M. Biberman, Plenum Press, New York, 1973, pp. 87-118.
- [11] H. L. Snyder, "Image Quality and Face Recognition on a Television Display," *Human Factors*, Vol. 16, 1974, pp. 300-307.
- [12] A. H. Blumenthal and S. B. Campana, "An Improved Electro-optical Image Quality Summary Measure," *Proceedings of the Society of Photographic Instrumentation Engineers, Image Quality*, Vol. 30, 1981.

- [13] H. R. Blackwell, "Contrast Threshold of the Human Eye," *Journal of the Optical Society of America*, Vol. 36, November, 1946, pp. 624-643.
- [14] M. I. Sezan, K. L. Yip, S. J. Daly, "Uniform Perceptual Quantization: Applications to Digital Radiography," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-17, No. 4, July/August, 1987, pp. 622-634.
- [15] S. M. Pizer and F. H. Chan, "Evaluation of the Number of Discernible Levels Produced on a Display," *Information Processing in Medical Imaging*, R. DiPaolo and E. Kahn, Eds., Editions INSERM, Paris, France, 1980, pp. 561-580.
- [16] H. de Vries, "The Quantum Character of Light and Its Bearing upon Threshold of Vision, The Differential Sensitivity and Visual Acuity of the Eye," *Physica*, Vol. 10, No. 7, 1943, pp. 553-564.
- [17] A. Rose, "The Sensitivity Performance of the Human Eye on a Absolute Scale," *Journal of the Optical Society of America*, Vol. 38, No. 2, February, 1948, pp. 196-208.
- [18] H. H. Barrett and W. Swindell, *Radiological Imaging*, Vol. 1, Academic Press, New York, 1981. pp. 199-201.
- [19] S. J. Briggs, "Photometric Technique for Deriving a "Best Gamma" for Displays," *SPIE Proceedings*, Vol. 199, 1979, pp. 134-145.
- [20] S. M. Pizer, "Intensity Mapping to Linearize Display Devices," *Computer Graphics and Image Processing*, Vol. 17, 1981, pp. 262-268.
- [21] R. E. Johnston, J. B. Zimmerman, D. C. Rogers, and S. M. Pizer, "Perceptual Standardization," *SPIE Proceedings*, Vol. 536, 1985, pp. 44-49.
- [22] H. Blume, H. Roehrig, M. Browne, and T-L. Ji, "Comparison of the Physical Performance of High Resolution CRT Displays and Films Recorded by Laser Image Printers and Displayed on Light-Boxes and the Need for a Display Standard," *SPIE Proceedings*, Vol. 1232, 1990, pp. 97-114.
- [23] S. Daly, "Application of Noise-Adaptive Contrast Sensitivity Function to Image Data Compression," *Optical Engineering*, Vol. 29, 1990, pp. 977-987.
- [24] J. Lubin and A. P. Pica, "A Non-uniform Quantizer Matched to Human Visual Performance," *SID 91 Digest*, 1991, pp. 619-622.

- [25] A. Rosenfeld and A. C. Kak, Digital Picture Processing, Academic Press, New York, 1976.
- [26] R. A. Schowengerdt, Techniques for Image Processing and Classification in Remote Sensing, Academic Press, New York, 1983.
- [27] H. Roehrig, W. J. Dallas, T-L. Ji, R. D. Lamoreaux, D. Oikawa, R. Vercillo, and D. Yorky, "Physical Evaluation of CRT's for Use in Digital Radiography," *SPIE Proceedings*, Vol. 1091, 1989, pp. 262-278.
- [28] T-L. Ji, H. Roehrig, H. Blume, G. Seeley, and M. Browne, "Physical and Psychophysical Evaluation of CRT Noise Performance," *SPIE Proceedings*, Vol. 1444, 1991, pp. 136-150.
- [29] H. Roehrig, T-L. Ji, M. Browne and H. Blume, "Physical Performance Data of the Tektronix GMA-202 Monitor," Technical Report, Department of Radiology and Optical Sciences Center, University of Arizona, July, 1990.
- [30] H. Roehrig, H. Blume, T-L. Ji, and M. Browne, "Performance Test and Quality Control of Cathode Ray Tube Displays," *Journal of Digital Imaging*, Vol. 3, 1990, pp. 134-145.
- [31] H. Roehrig, T-L. Ji, M. Browne, and W. J. Dallas, "Signal-to-Noise Ratio and Maximum Information Content of Images Displayed by a CRT," *SPIE Proceedings*, Vol. 1232, 1990, pp. 115-133.
- [32] H. Blume, H. Roehrig, T-L. Ji, and M. Browne, "Very-High Resolution Monochrome CRT Displays: How Good Are They Really?" *SID 91 Digest*, 1991, pp. 355-358.
- [33] H. Blume, H. Roehrig, and T-L. Ji, "Very High Resolution CRT Display Systems: Update on State of the Art of Physical and Psychophysical Performance," *SID 92 Digest*, 1992, pp. 699-702.
- [34] H. Blume, T-L. Ji and H. Roehrig, "Physical Performance Data of the Sony DDM-2802F Monitor", Technical Report, Department of Radiology and Optical Sciences Center, University of Arizona, April, 1992.
- [35] M. D. Levine, Vision in Man and Machine, McGraw-Hill, New York, 1985.
- [36] J. C. Dainty and R. Shaw, Image Science, Academic Press, New York, 1974.

- [37] J. D. Gaskill, Linear Systems, Fourier Transforms, and Optics, Wiley, New York, 1978.
- [38] J. W. Coltman, "The Specification of Imaging Properties by Response to Sine-wave Input", *Journal of the Optical Society of America*, Vol. 44, No. 6, June, 1954, pp. 468-471.
- [39] J. V. Candy, Signal Processing: The Modern Approach, McGraw-Hill, New York, 1986.
- [40] T. Cornsweet, Visual Perception, Academic Press, New York, 1970.
- [41] C. H. Graham, Editor, Vision and Visual Perception, Wiley, New York, 1965.
- [42] F. W. Campbell and J. G. Robson, "Application of Fourier Analysis to the Visibility of Gratings," *Journal of Physiology*, Vol. 197, No. 3, 1968, pp. 511-566.
- [43] C. F. Hall and E. L. Hall, "A Nonlinear Model for the Spatial Characteristics of the Human Visual System," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-7, No. 3, March, 1977, pp. 161-170.
- [44] P. G. J. Barten, "The SQRI Method: a New Method for the Evaluation of Visible Resolution on a Display," *Proceedings of SID*, Vol. 28, No. 8, 1987, pp. 253-262.
- [45] F. W. Campbell, J. J. Kulikowski, and J. Levinson, "The Effect of Orientation on the Visual Modulation of Gratings," *Journal of Physiology*, Vol. 187, 1966, pp. 427-436.
- [46] A. P. Ginsburg, "Specifying Relevant Spatial Information for Image Evaluation and Display Design: An Explanation of How We See Certain Objects," *Proceedings of SID*, Vol. 21 No. 3, 1980, pp. 219-227.
- [47] G. W. Seeley, W. J. Dallas, et al., "Spatial Resolution Requirements for Digital Radiology," Private Communication, 1991.
- [48] A. D. Schnitzler, "Analysis of Noise-Required Contrast and Modulation in Image-Detecting and Display Systems," In Perception of Displayed Information, ed. L. M. Biberman, Plenum Press, New York, 1973, pp. 119-166.

- [49] A. E. Burgess, R. F. Wagner and R. J. Jennings, "Human Signal Detection Performance for Noisy Medical Images," *Proc. IEEE Com. Soc. Int. Workshop on Medical Imaging*, Asilomar, California, March, 1982.
- [50] R. D. Fiete, H. H. Barrett, W. E. Smith and K. J. Myers, "The Hotelling Trace Criterion and its Correlation with Human Observer Performance," *Journal of the Optical Society of America*, Vol. 4, 1987, pp. 945-953.
- [51] H. H. Barrett, "Evaluation of Image Quality Through Linear Discriminant Models," *SID Digest*, 1992, pp. 871-873.
- [52] F. A. Rosell and R. H. Willson, "Recent Psychophysical Experiments and the Display Signal-to-Noise Ratio Concept," In Perception of Displayed Information, ed. L. M. Biberman, Plenum Press, New York, 1973, pp. 167-232.
- [53] A. Rose, Vision: Human and Electronic, Plenum, New York, 1974.
- [54] R. G. Gallager, Information Theory and Reliable Communication, Wiley, New York, 1968.
- [55] D. M. Green and J. A. Swets, Signal Detection Theory and Psychophysics, Wiley, New York, 1966.
- [56] M. Lassen and P. Bloch, "Measurements of the Effects of X-ray Film-Screen Characteristics on Threshold Detectability of Low-contrast Objects," *Medical Physics*, Vol. 5, 1978, pp. 152-161.
- [57] P. M. Lams and M. L. Cocklin, "Spatial Resolution Requirement for Digital Chest Radiographs: A ROC Study of Observer Performance in Selected Cases," *Radiology*, Vol. 158, 1986, pp. 11-19.
- [58] H. MacMahon, C. E. Metz, K. Doi, T. Kim, M. L. Giger, and H-P. Chan, "Digital Chest Radiography: Effect on Diagnostic Accuracy of Hard Copy, Conventional Video, and Reversed Gray Scale Video Display Formats," *Radiology*, Vol. 168, 1988, pp. 669-673.
- [59] B. S. Slasky, D. Gur, W. F. Good, M. A. Costa-Greco, K. M. Harris, L. A. Cooperstein, and H. E. Rockette, "Receiver Operating Characteristic Analysis of Chest Image Interpretation with Conventional, Laser-printed, and High-Resolution Workstation Images," *Radiology* Vol. 174, 1990, pp. 775-780.

- [60] M. Razavi, J. W. Sayre, R. K. Taira, M. Simons, H. K. Huang, K. Chuang, G. Rahbar, and H. Kangarloo, "Receiver-Operating-Characteristic Study of Chest Radiographs in Children: Digital Hard-Copy Film vs 2K \times 2K Soft-Copy Images," *AJR* Vol. 158, 1992, pp. 443-448.
- [61] C. E. Metz, "Some Practical Issues of Experimental Design and Data Analysis in Radiological ROC studies," *Investigative Radiology*, Vol. 24, 1989, pp. 234-245.
- [62] L. Levi, "Unsharp Masking and Related Image Enhancement Techniques," *Computer Graphics and Image Processing*, Vol. 3, 1974, pp. 163-177.
- [63] P. G. Tahoces, J. Correa, M. Souto, C. Gonzalez, L. Gómez, and J. J. Vidal, "Enhancement of Chest and Breast Radiographs by Automatic Spatial Filtering," *IEEE Transactions on Medical Imaging*, Vol. MI-10, No. 3, September, 1991, pp. 330-335.
- [64] J. D. Fahnestock and R. A. Schowengerdt, "Spatially Variant Contrast Enhancement Using Local Range Modification," *Optical Engineering*, Vol. 22, No. 3, 1983, pp. 378-381.
- [65] T. Peli and J. S. Lim, "Adaptive Filtering for Image Enhancement," *Optical Engineering*, Vol. 21, No. 1, 1982, pp. 108-112.
- [66] D. J. Ketcham, R. W. Lowe and J. W. Weber, "Real-Time Image Enhancement Techniques," *Seminar on Image Processing*, Pacific Grove, California, 1976, pp. 1-6.
- [67] R. A. Hummel, "Image Enhancement by Histogram Transformation," *Computer Graphics and Image Processing*, Vol. 6, 1977, pp. 184-195.
- [68] S. M. Pizer, J. B. Zimmerman and E. V. Staab, "Adaptive Grey Level Assignment in CT Scan Display," *Journal of Computer Assisted Tomography*, Vol. 8, No. 2, 1984, pp. 300-305.
- [69] R. H. Sherrier and G. A. Johnson, "Regionally Adaptive Histogram Equalization of the Chest," *IEEE Transactions on Medical Imaging*, Vol. MI-6, No. 1, March, 1987, pp. 1-7.
- [70] J. B. Zimmerman, "Effectiveness of Adaptive Contrast Enhancement," *Ph.D Dissertation*, Department of Computer Science, University of North Carolina, 1985.

- [71] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. H. Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive Histogram Equalization and Its Variations," *Computer Vision, Graphics and Image Processing*, Vol. 39, 1987, pp. 355-368.
- [72] K. Rehm, "Development and Image Quality Assessment of A Contrast-Enhancement Algorithm for Display of Digital Chest Radiographs," *Ph.D Dissertation*, Optical Sciences Center, University of Arizona, 1992.
- [73] A. N. Netravali and B. Prasada, "Adaptive Quantization of Picture Signals Using Spatial Masking," *Proceedings of the IEEE*, Vol. 65, No. 4, 1977, pp. 536-548.
- [74] H. Blume, S. Daly and E. Muka, "Presentation of Medical Images on CRT Displays: A Renewed Proposal for a Display Function Standard," *SPIE Proceedings*, Vol. 1897, 1993, pp. 215-231.